AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

A Human MSH6 Germline Variant Associated With Systemic Lupus Erythematosus Induces Lupus-like Disease in Mice

Rithy Meas,[1] 🆔 Joanne Nititham,[2] Kimberly E. Taylor,[2] Stephen Maher,[3] Kaylyn Clairmont,[3] Kelly E. W. Carufe,[3] Michael Kashgarian,[3] Timothy Nottoli,[3] Ana Cheong,[4] Zachary D. Nagel,[4] Patrick M. Gaffney,[5] 🆔 Lindsey A. Criswell,[6] and Joann B. Sweasy[1] 🆔

Objective. To determine if single-nucleotide polymorphisms (SNPs) in DNA repair genes are enriched in individuals with systemic lupus erythematosus (SLE) and if they are sufficient to confer a disease phenotype in a mouse model.

Methods. Human exome chip data of 2499 patients with SLE and 1230 healthy controls were analyzed to determine if variants in 10 different mismatch repair genes (*MSH4, EXO1, MSH2, MSH6, MLH1, MSH3, POLH, PMS2, ML3,* and *APEX2*) were enriched in individuals with SLE. A mouse model of the *MSH6* SNP, which was found to be enriched in individuals with SLE, was created using CRISPR/Cas9 gene targeting. Wildtype mice and mice heterozygous and homozygous for the *MSH6* variant were then monitored for 2 years for the development of autoimmune phenotypes, including the presence of high levels of antinuclear antibodies (ANA). Additionally, somatic hypermutation frequencies and spectra of the intronic region downstream of the V_HJ558-rearranged J_{H4} immunoglobulin gene was characterized from Peyer's patches.

Results. Based on the human exome chip data, the *MSH6* variant (rs63750897, p.Ser503Cys) is enriched among patients with SLE versus controls after we corrected for ancestry (odds ratio = 8.39, P = 0.0398). Mice homozygous for the *MSH6* variant (*Msh6$^{S502C/S502C}$*) harbor significantly increased levels of ANA. Additionally, the *Msh6$^{S502C/S502C}$* mice display a significant increase in the infiltration of CD68+ cells (a marker for monocytes and macrophages) into the lung alveolar space as well as apoptotic cells. Furthermore, characterization of somatic hypermutation in these mice reveals an increase in the DNA polymerase η mutational signature.

Conclusion. An *MSH6* mutation that is enriched in humans diagnosed with lupus was identified. Mice harboring this *Msh6* mutation develop increased autoantibodies and an inflammatory lung disease. These results suggest that the human *MSH6* variant is linked to the development of SLE.

INTRODUCTION

Systemic lupus erythematosus (SLE) is characterized by the activation of self-reactive B and T cells, which subsequently leads to multiorgan inflammation, affecting organs such as the skin, kidney, and lung. Currently, the exact etiology of SLE is unknown. Previous studies have linked 132 risk loci to SLE based on genome-wide association studies (1–3); however, many of their etiological roles in SLE are not firmly established (4). Recent data from our laboratory have shown that murine SLE can develop from a single amino acid substitution in the base excision repair

(BER) protein DNA polymerase β (Pol β)—specifically, a tyrosine to cysteine substitution at position 265 (Y265C) (5). Additionally, defects in DNA repair have been reported previously in lymphoid cells from patients with SLE. These lymphoid cells from patients with SLE are hypersensitive to hydrogen peroxide and N-methyl-N-nitrosourea and, furthermore, are defective in the repair of O^6-methylguanine and double-strand break repair (6–8). Collectively, these data provide a strong rationale to characterize the potential roles of DNA repair variants in the development of SLE.

In this study, we identify a DNA repair genetic variant in the *MSH6* gene that is substantially enriched in humans diagnosed

---

[1]Rithy Meas, PhD, Joann B. Sweasy, PhD: University of Arizona, Tucson; [2]Joanne Nititham, MPH, Kimberly E. Taylor, PhD, MPH: University of California, San Francisco; [3]Stephen Maher, MHS, Kaylyn Clairmont, Kelly E. W. Carufe, Michael Kashgarian, MD, Timothy Nottoli, PhD: Yale School of Medicine, New Haven, Connecticut; [4]Ana Cheong, PhD, Zachary D. Nagel PhD: Harvard School of Public Health, Boston, Massachusetts, USA; [5]Patrick M. Gaffney, MD: Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA; [6]Lindsey A. Criswell, MD, MPH, DSc: National Institute of Arthritis and Musculoskeletal and Skin Diseases, Bethesda, Maryland, USA.

with SLE. Under normal conditions, *MSH6* with its cognate partner, *MSH2*, form a heterodimeric complex, MutSα, that initiates mismatch repair (MMR) by recognizing and binding to DNA mismatches and one to two base nucleotide insertion and deletion loops (9). Thereafter, there are two strand-directed processes that can occur. One process involves the recruitment of exonuclease 1 (Exo1) to a nick that is 5′ of the DNA mismatch, and on activation by MutSα, Exo1 hydrolyzes the nicked strand 5′ to 3′ past the DNA mismatch (10). The second process requires MutLα, which comprises MLH1/PMS2 and can occur when the nick is either 3′ or 5′ of the DNA mismatch. In the presence of the nick, MutSα, replication factor C, and proliferating cell nuclear antigen (PCNA), the latent endonuclease activity of MutLα is activated to incise the DNA on the same strand that contains the pre-existing strand break (11). The incision results in the DNA mismatch being bracketed by a DNA break on each side, which can either be displaced by polymerase δ during DNA synthesis or hydrolytically removed via MutSα-activated Exo1 activity (12). After either of the strand-directed processes, the remaining gap requires PCNA and polymerase δ or ε, followed by DNA ligation via DNA ligase I (13,14). In the absence of MMR proteins, there can be a 2- to 90-fold increase in mutation frequencies in mammalian cells, indicating the importance of MMR in maintaining genomic integrity (15). Paradoxically, MutSα also facilitates mutagenesis at immunoglobulin loci during antibody diversification in activated B cells. Interestingly, aberrant somatic hypermutation (SHM) and class switch recombination (CSR), both antibody diversification processes that employ MMR, are associated with SLE in humans and the lupus-prone MRL-*Fas^lpr/lpr* mice by increasing the levels of antibodies that recognize self-antigens and elicit an immune response (16).

Given that aberrant DNA repair is linked with SLE, the first objective of this study is to identify specific DNA repair variants associated with SLE. Next, by developing and studying an animal model, this study aims to determine if an SLE-associated DNA repair variant is sufficient for lupus development.

PATIENTS AND METHODS

Patients. Patients with SLE were enrolled from the University of California, San Francisco (n = 394); Oklahoma Medical Research Foundation (n = 1298); and King's College Hospital in London (n = 807). All participants provided written informed consent and fulfilled American College of Rheumatology 1997 SLE classification criteria. Controls were from the Alzheimer's Disease Genetics Consortium (Radboud University, n = 1230), which included the institutions Washington University in St. Louis, Children's Hospital of Philadelphia, University of Miami, and the Feinstein Medical Institute for Medical Research in New York (17).

Human exome chip. The Infinium Exome-24 v1.1 BeadChip (Illumina) was used on the above cohorts because it provided a high-throughput workflow that allowed for the analysis of more than 240,000 markers, which included protein coding variants selected from the exome sequences of approximately 12,000 samples globally. Genotype calling was performed by Illumina Genome Studio software. First, we removed samples missing more than 20% of genotyping data. We then removed single-nucleotide polymorphisms (SNPs) missing more than 5% of data, samples missing more than 5% of data, and SNPs showing a departure from Hardy–Weinberg equilibrium ($P < 0.001$). The sample set was restricted to individuals of self-reported European descent, and then principal components analysis resulted in an additional removal of ancestry outliers at a sigma threshold of 4.5. After the quality control filtering, 2499 patients with SLE and 1230 healthy controls remained. For this project, we examined all variants in 10 MMR genes outside of the major histocompatability complex (MHC) (*MSH4*, *EXO1*, *MSH2*, *MSH6*, *MLH1*, *MSH3*, *POLH*, *PMS2*, *MLH3*, and *APEX2*) that were present in patients and controls (114 genetic variants) to identify SNPs that were enriched in patients with SLE. We performed logistic regression analyses, adjusting for the first principal component and sex. Functional predictions for candidates for follow-up were obtained using Polyphen and Sift.

Mouse construction and maintenance. The mouse was constructed using the Yale Genome Editing Center. Specifically, the single-guide RNA that was used to construct the mouse via CRISPR/Cas9 targeting is: TGGAGCAGACCGA-GACTCCAGAAATGATGGAGGCGCGATGTCGTAAGATGGCA-CACGTGTGCAAGTTTGATAGAGTGGTGAGAAGGGAGATTTG-CAGGATCATTACCAAGGGCACACAGAC. Underlined is the codon that was altered in C57BL/6 mice using CRISPR/Cas9 gene targeting. The founder mouse was then backcrossed to C57BL/6 (Charles River) and maintained in pathogen-free conditions at the Yale animal care facility with oversight by the Yale University Institutional Animal Care and Use Committee. Mice were monitored twice weekly for dermatitis and other developmental phenotypes. All mice were included during analysis, and there were no exclusions. Locations of mouse cages were randomized to control for mouse cage location confounder. The genotype of the mice was identified by using the following primers: M6F′ (GTATGTGCCTGAA-GAGTTC) and M6R′ (ACTAGCGTCCTAAATCTGG). The polymerase chain reaction (PCR) protocol used for genotyping with Taq polymerase (Thermo Scientific, EP0402) was 95°C for 3 minutes, 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 35 seconds, repeat the second to fourth steps 35 times, and 72°C for 5 minutes. Afterward, the amplified products were resolved on an agarose gel and subsequently sequenced by the Yale Keck facilities using standard technologies.

Histology and scoring of kidney disease. Hematoxylin and eosin–stained slides of kidney tissue were scored as described in the Supplementary Materials.

Immunohistochemistry. Five-micrometer sections were stained as described in the Supplementary Materials (Supplementary Table 1).

Antinuclear antibody detection and scoring. This was performed as previously described (5).

Microsatellite instability. Details regarding microsatellite instability are described in the Supplementary Materials.

***Pig-a* assay.** Details regarding the *Pig-a* assay are described in the Supplementary Materials.

SHM. The assay was performed (similarly as described in ref. 5) to characterize SHM in intronic region downstream of the $V_H J558$-rearranged J_{H4} gene. Briefly, Peyer's patches from non-immunized 4- to 6-month-old mice were placed in dissociation media (Stemcell Tech., cat. #07915), and tissue grinding between two frosted slides followed (FisherSci, cat. #12-544-5CY). Cells were then passed through a 70-µm strainer. Fc block (Biolegend, cat. #101320) was used, followed by a cocktail consisting of α-B220 and a PNA-conjugated protein. The B220+PNA[high] cells were then sorted and collected from a beckton dickinson FAC-SAria III sorter and then lysed. The DNA from the lysed cells were then amplified through two rounds of PCR. For each round, new england biolabs (NEB) Phusion Taq (NEB, cat. #M0530) was used. The first round of PCR used the following primers: SHMvF (AGCCTGACATCTGAGGAC) and SHMvR (TAGTGTGGAACAT-TCCTCAC), and the program used was 94°C for 3 minutes, 94°C for 30 seconds, 55°C for 90 seconds, 72°C for 2 minutes, and repeat the second to fourth steps for 30 cycles. Afterward, the following nested primers were used: SHMv-nestF (5′CCG-GAATTCCTGACATCTGAGGACTCTGC) and SHMv-nestR (5′C-GCGGATCCGCTGTCACAGAGGTGGTCCTG), and the program used was 94°C for 3 minutes, 94°C for 30 seconds, 60°C for 90 seconds, 72°C for 1 minute, and repeat the second to fourth steps for 30 cycles. TOPO cloning was then performed with the amplified DNA, as described by the manufacturer (Invitrogen, cat. #K4500-01SC). PCR of DNA isolated from colonies was performed with M13 forward primer and T7 reverse primers, and the PCR amplicons were subsequently sequenced at the Keck facilities at Yale University using standard technologies.

Class switching assay. Details regarding the class switching assay are described in the Supplementary Materials

Fluorescence-based mismatch repair assay. Details regarding the fluorescence-based mismatch repair assay are described in the Supplementary Materials.

RESULTS

The *MSH6^(S503C)* variant is enriched in individuals with SLE. Based on our discovery that the Y265C Pol β mice develop lupus-like phenotypes and also research suggesting that lymphoid cells from patients with SLE are defective in DNA repair (5–8), we tested the hypothesis that germline variants (GVs) in DNA repair genes are associated with human SLE. Specifically, for this project, we analyzed exome chip (eChip) data for 114 SNPs within 10 non-MHC MMR genes (see Methods) that were present in both patients and controls. In a comparison of 2499 patients with SLE with 1230 controls, corrected for ancestry, nine GVs associated with MMR with an odds ratio (OR) greater than 1.0 and association P value less than 0.05 were identified as being enriched in patients with SLE versus controls (Supplementary Table 2) and considered for our mouse model. One of the GVs that emerged from our eChip analysis was the *MSH6^(S503C)* variant. This *MSH6* GV was more than eight times more prevalent in patients with SLE versus controls after we corrected for ancestry (OR = 8.39; $P = 0.0398$) and was predicted to be a damaging variant by both the Polyphen and Sift algorithms. When comparing these carriers with patients with SLE who are not carriers for the *MSH6* variant, we determined that specific SLE phenotypes, such as lupus nephritis and auto-antibodies to double-stranded DNA, and the SLE bias for women are similar between the two groups (Table 1). Remarkably, among the patients with SLE in the cohort, patients heterozygous for the *MSH6^(+/S503C)* variant are diagnosed with SLE 11.2 years earlier than patients who are *MSH6^(+/+)* (Table 1). Further analysis using logistic regression determined that carriers of the *MSH6* variant are five times more likely to develop childhood-onset SLE (ages 0-15 years, OR = 5.31 [confidence interval = 1.41-19.75]; $P = 0.014$; ages of patients listed in Supplementary Table 3). Childhood-onset SLE is associated with increased severity of SLE phenotypes compared with adult-onset or late-onset SLE (16,18). Therefore, this variant was an

Table 1. Phenotypes of patients with SLE based on *MSH6* genotypes

	Lupus nephritis, n/total (%)	Anti-dsDNA, n/total (%)	Women, n/total (%)	Mean age at diagnosis,[a] years
MSH6^(+/+)	829/2247 (37)	719/2054 (35)	2250/2482 (91)	34.6
MSH6^(+/S503C)	5/14 (36)	3/10 (30)	14/15 (93)	23.4

Abbreviations: dsDNA, double-stranded DNA; SLE, systemic lupus erythematosus.
[a] C C alleles, n = 2188; G C alleles, n = 12; *t*-test, $P = 0.0039$.

ideal choice for the creation of a mouse model to determine if it is sufficient for development of lupus.

Creation of the *Msh6* variant mouse model. There is high conservation between human *MSH6* and mouse *Msh6*, as shown in Figure 1A. Further analysis of the protein sequence indicates that the missense mutation, C→G, results in a serine to cysteine substitution at position 503 in humans, or 502 in mice, which is located in the mismatch binding domain of the MSH6 protein (Figure 1B and C). Given the conservation of the gene and protein sequences between human and mouse, we used CRISPR/Cas9 to construct the *Msh6*[S502C] mouse model. After obtaining the mice, we crossed heterozygous mice with each other. Genotyping of the resulting pups confirmed that the expected genotypes were present (Figure 1D) at a normal Mendelian ratio (14:35:14). Additionally, the expression of the MSH6 protein is similar between all three genotypes (Figure 1E), indicating that the variant is not affecting MSH6 protein expression, which is in concordance with data from human tissue (19).

The *Msh6*[S502C/S502C] mice exhibit high levels of antinuclear antibodies and a reduced lifespan. To determine if the mice displayed evidence of autoimmunity, we monitored the development of antinuclear antibodies (ANAs). Remarkably, the *Msh6*[S502C/S502C] mice exhibited higher levels of ANAs at both 6 and 15 months of age (Figure 2A). To determine if the mice exhibited other manifestations of autoimmunity, we assessed the animals for dermatitis and glomerulonephritis. Analyses of dermatitis incidence and renal disease presence and severity in our mice, scored as described in Methods, did not reveal significant differences between wildtype and the *Msh6*[S502C/S502C] mice (Figure 2B and C, Supplementary Figure 1). However, Figure 2D shows that the MSH6 protein is minimally expressed in the kidneys, which may explain why renal disease is not detected in the *Msh6*[S502C/S502C] mice. Intriguingly, some of the *Msh6*[S502C/S502C] mice have a reduced lifespan compared with the wildtype and *Msh6*[S502C/+] mice (Figure 2E), which is a common phenotype observed in other murine SLE models (20). We also analyzed tumor development in the mice given that mutations in MMR genes can be associated with Lynch syndrome, which predisposes individuals to different types of cancer (21).
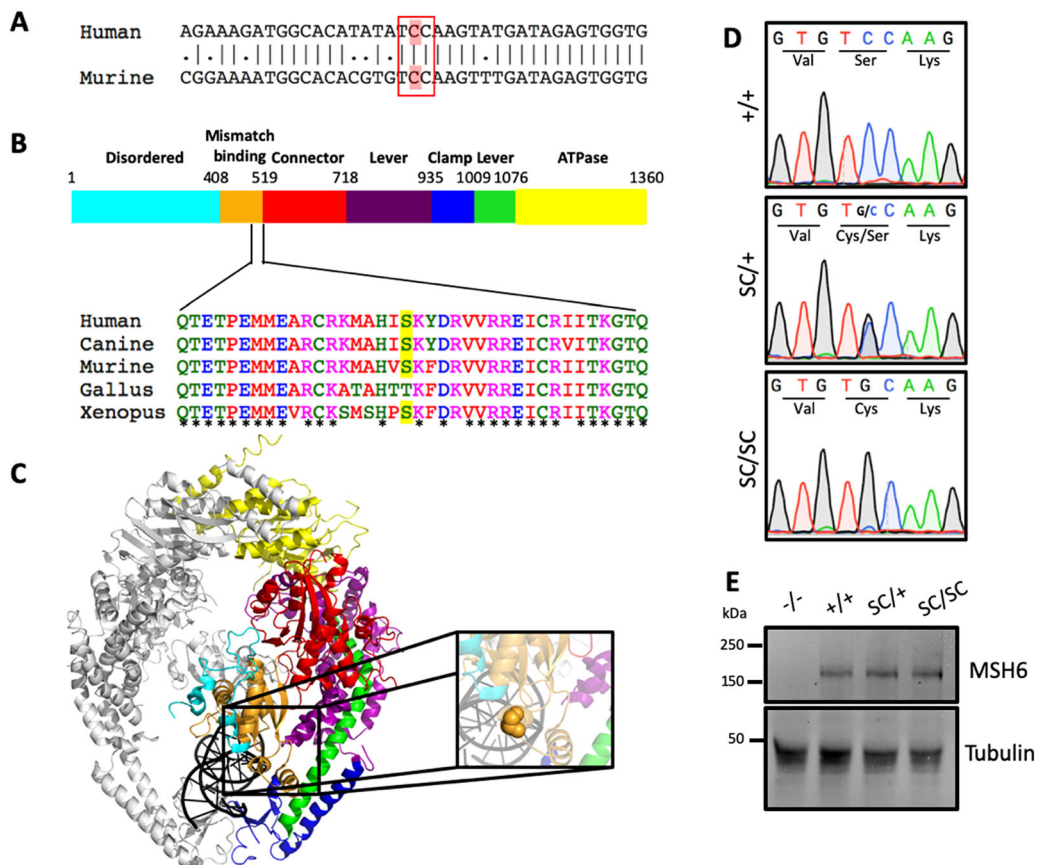

Figure 1. Creation of the *Msh6* variant mouse model. The codon (**A**) and amino acid (**B**) are conserved among different organisms. **C**, MSH2 is in gray, and MSH6 is colored as in **B**. Inset shows MSH6 S503 in gold (PDB: 2O8B). **D**, Chromatograms of the *Msh6* allele. The top is wildtype, the middle is heterozygous for the *Msh6*[S503C] allele, and the bottom is homozygous for the *Msh6*[S503C] allele. **E**, Mouse embryonic fibroblasts (MEFs) from the indicated genotypes show the MSH6 S503C protein is expressed when compared with MEFs nullizygous for *Msh6*.
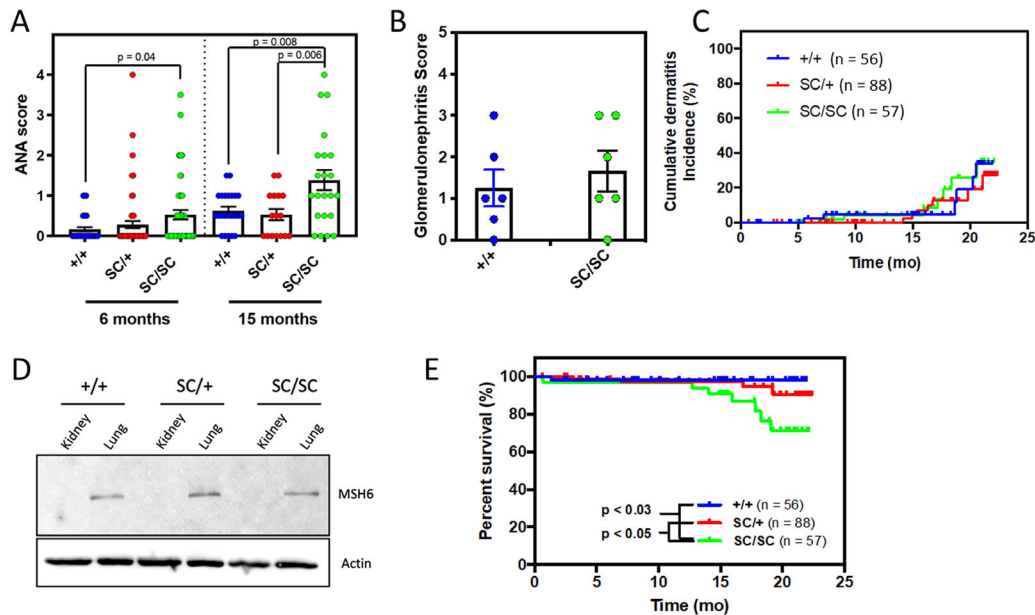
Figure 2. The *Msh6^S502C/S502C* mice develop high antinuclear antibodies (ANA) and reduced survival. **A**, Sera from mice at 6 and 15 months of age were tested for ANA on human epithelial (HEp-2) cells in 12-well slides (Diasorin). Each point represents an individual mouse, and a one-way analysis of variance test with Tukey's post hoc tests was used to compare the genotypes. Bars indicate means ± SEM. **B**, Glomerulonephritis was scored as described in Methods. **C**, Kaplan–Meier curves of cumulative incidence of dermatitis. **D**, Western blot shows no difference between the genotypes in MSH6 protein expression in 12-month-old mice. **E**, Kaplan–Meier survival curves of *Msh6^+/+*, *Msh6^S502C/+*, and *Msh6^S502C/S502C* mice. Comparisons were evaluated via log-rank tests. All statistical analyses were performed using GraphPad Prism.

Tumors were observed in 5 of 27 (18.5%), 5 of 32 (15.5%) and 4 of 28 (14%) wildtype, heterozygous, and homozygous mice, respectively (Supplementary Table 4). None of the *Msh6^S502C/S502C* mice developed intestinal tumors, whereas intestinal tumors were observed in one wildtype and one heterozygous mouse at the time of necropsy. This is in contrast to the observation of tumors in the majority of *Msh6*-deficient and *Msh6^T1217D* mutant mice (22,23). Overall, these data indicate that mice homozygous for the human SLE-associated *MSH6* GV are developing autoimmunity and have a shortened lifespan.

Lung disease in the *Msh6^S502C/S502C* mice. It is estimated that 80% to 98% of patients with SLE develop pulmonary disease, such as pleuritis, alveolar hemorrhaging, and alveolar pulmonary infiltrates (24). Given that MSH6 expression is high in lungs and not kidneys (Figure 2E), we wanted to determine if there were any pulmonary manifestations in the *Msh6^S502C/S502C* mice. Interestingly, we found that there is an increase in cleaved caspase-3-positive cells, a marker for apoptosis, at 7 to 12 months for both the *Msh6^S502C/+* and *Msh6^S502C/S502C* mice compared with wildtype mice (Figure 3A and B). This increase in apoptosis could elicit recruitment of phagocytes for the removal of these dead cells (25). Indeed, the *Msh6^S502C/S502C* mice have a significant increase in the immune infiltration of CD68+ cells, a marker for monocytes and macrophages, in the alveolar space of their lungs at 7 to 12 months that persists to 13 to 18 months

and is still high at 19 to 24 months, though it is not significant at this later time point (Figure 3C and D). Macrophages exhibit plasticity and alter their phenotypes in response to the tissue microenvironment (26). Classically activated M1 macrophages are considered to exhibit a proinflammatory phenotype generally developed to clear pathogens; these macrophages express high levels of iNOS. Alternatively activated M2 macrophages are suggested to possess an anti-inflammatory phenotype and participate in wound healing. M2 macrophages generally express high levels of ARG1. Therefore, the iNOS and ARG1 markers are thought to represent pro- and anti-inflammatory macrophages, respectively, in tissue. Intriguingly, we observed a significant increase in iNOS+ cells in *Msh6^S502C/S502C* versus wildtype mice at 13 to 18 months of age and a potential increase in iNOS+ cells at 19 to 24 months in the mutant mice that was not significant. We also observe an increase of ARG1+ cells at both 13 to 18 and 19 to 24 months of age in *Msh6^S502C/S502C* mice compared with wildtype mice (Supplementary Figure 2). Of note, the cells marked with iNOS are likely proinflammatory macrophages and are present at higher levels in the lungs of *Msh6^S502C/S502C* versus wildtype mice than anti-inflammatory ARG1+ cells at 13 to 18 months of age, but ARG1+ macrophages appear to be more highly enriched than iNOS-marked cells in the lungs of mutant mice at 19 to 24 months. This pattern suggests that the lung tissue environment differs in the mutant mice as a function of age and perhaps other factors and that at 19 to 24 months of age, the environment may be of an anti-inflammatory nature.
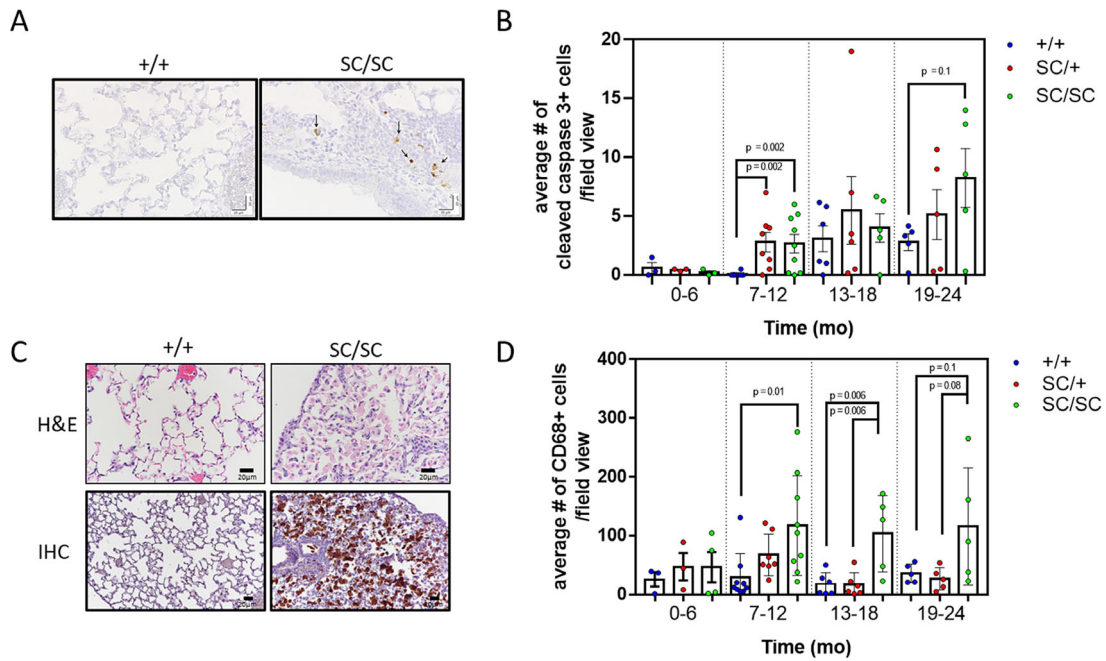
Figure 3. There are abnormalities in the lungs of $Msh6^{S502C/S502C}$ mice. **A,** There is increased cleaved caspase 3 staining in the lungs of the $Msh6^{S502C/S502C}$ mice as quantified in **B. C,** Hematoxylin and eosin (H&E) slides show infiltration of cells into the alveolar space. These cells stained as CD68+ cells via immunohistochemistry, and on quantification (**D**), the lungs from the $Msh6^{S502C/S502C}$ mice showed an increase in infiltration of CD68+ cells at 7-12 and 13-18 months of age when compared with $Msh6^{+/+}$ lungs. Black bars in **A** and **C** are 20 μm. Statistical analysis of **B** and **D** was performed by one-way analysis of variances with Tukey's multiple comparisons test.
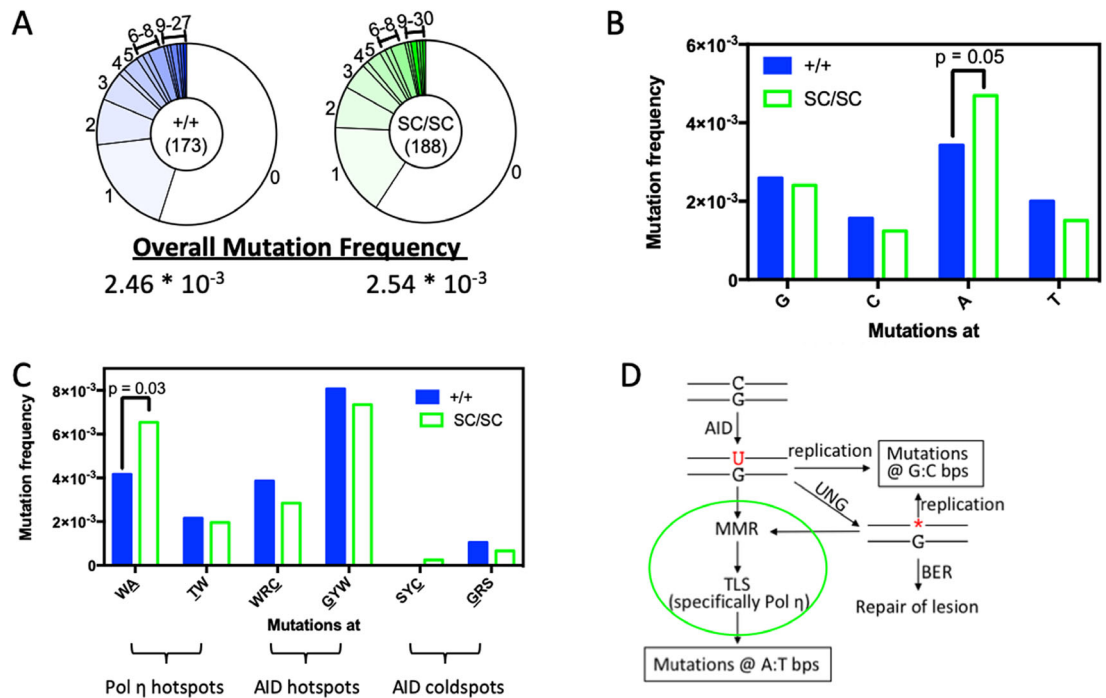

Figure 4. Somatic hypermutation (SHM) spectra in the $Msh6^{S502C/S502C}$ mice are altered. **A,** Overall mutation frequencies between wildtype and the $Msh6^{S502C/S502C}$ mice at the intronic region downstream of the $V_H J558$-rearranged J_{H4} gene. The number in the inner circle indicates the number of clones analyzed. The pie chart shows the percentage of clones that contained the indicated number of mutations. **B,** Mutation frequencies on the nontranscribed strand at each base and hot or cold spots (**C**), where underlined bases are the mutated sites (W = A/T; R = A/G; Y = C/T; S = G/C). For **A** and **B**, data are from three mice per genotype and P values are determined from χ^2 analyses as with Yate's correction. **D,** Model of SHM (58), in which the green circle indicates the proposed steps that the $Msh6^{S502C}$ allele influences during SHM.
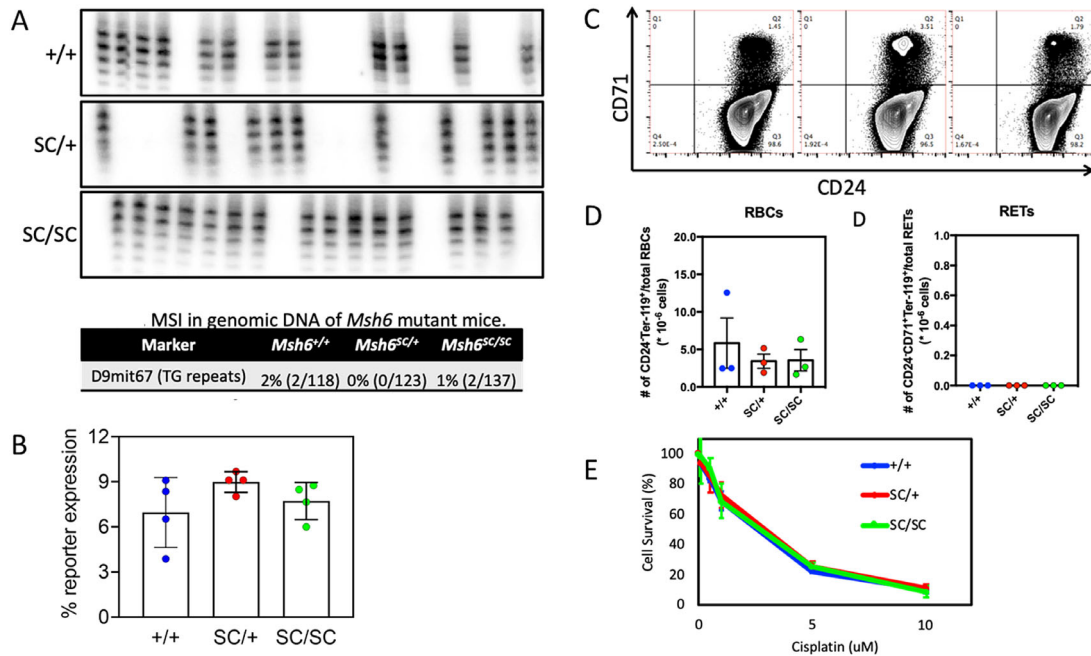
Figure 5. The *Msh6*[S502C] variant does not affect mutagenesis or canonical mismatch repair (MMR). **A**, Phosphor images of amplified D9mit67 microsatellites. The percentage of microsatellites that are unstable is shown in the table. **B**, Canonical MMR was determined by transfecting the cells with a reporter plasmid containing a DNA mismatch. **C**, The *Pig-a* assay was used as an indirect method to determine the mutation frequency at the *Pig-a* gene from mouse blood. The loss of the CD24 marker is indicative of a *Pig-a* mutation and is quantified in **D**. **E**, Cisplatin is a DNA crosslinker that has been shown to induce caspase-dependent cell death in MMR-proficient cells. Survival of cells was determined after 24-hour treatment with the indicated doses of cisplatin.

Furthermore, the 7- to 12-month-old *Msh6*[S502C/S502C] mice exhibited a significant increase in CD3+ cells, a marker for T cells, in the lungs (Supplementary Figure 3). These data show that the *Msh6*[S502C/S502C] mice develop pulmonary disease, as do patients with SLE.

The *Msh6*[S502C] allele alters SHM but does not affect CSR. MMR has known roles in SHM and CSR, processes important for antibody diversification. Interestingly, these processes contribute to ANA production and pathogenicity (27,28). Given that the *Msh6*[S502C/S502C] mice develop high levels of ANA, we asked if there was aberrant antibody diversification in mice expressing the *Msh6* variant.

The process of SHM is important for ~10^6-fold increase in mutations at immunoglobulin loci (29), allowing for the variable (V) region of antibodies in B cells to be diversified for the recognition of different antigens. SHM is initiated by transcription-coupled deamination of cytosine to uracil by activation-induced cytidine deaminase (AID), which creates the U:G mismatches. These mismatches are subsequently processed through BER, MMR, and/or translesion synthesis, resulting in mutations at G:C base pairs (bps) or A:T bps or complete repair of the site (30). To determine if the *Msh6*[S502C] allele influences SHM, the intronic region downstream of the V_HJ558-rearranged J_{H4} immunoglobulin gene was analyzed for the presence of somatic mutations in activated B cells

as described in Methods. As shown in Figure 4A, the overall mutation frequency was similar between wildtype (+/+) and the *Msh6*[S502C/S502C] (SC/SC) mice. However, the mutational spectra differed between the wildtype and mutant mice. Specifically, the frequency of mutations at adenine was greater in the *Msh6*[S502C/S502C] mice than in wildtype siblings, approaching statistical significance (Figure 4B). Importantly, there was a significantly increased mutation frequency at the DNA polymerase η (pol η) hotspot, WA, in the *Msh6*[S502C/S502C] mice when compared with the wildtype mice (Figure 4C), suggesting that pol η was influencing somatic mutations at the V_HJ558-rearranged J_{H4} Ig gene in the B cells of the *Msh6*[S502C/S502C] B cells (Figure 4D).

MMR is also important for CSR. Specifically, the processing of AID at immunoglobulin switch regions by both MMR and BER promote double-strand breaks. These breaks can then recombine the variable heavy-chain segment to a different constant heavy gene via nonhomologous end joining, whereby antibodies are class switched from immunoglobulin M (IgM) to immunoglobulin G (IgG), immunoglobulin A, and immunoglobulin E isotypes (30). To determine if the *Msh6*[S502C] variant alters CSR, we monitored the isotype switching of B cells from IgM to IgG1, IgG2a, IgG2b, and IgG3. In this in vitro assay, we isolated splenic B cells and induced isotype switching via treatment with different cytokines and lipopolysaccharide as described in Methods. As shown in Supplementary Figure 4, the *Msh6*[S502C/S502C] B cells do not show altered

isotype switching compared with wildtype B cells. Collectively, our data indicate that the *Msh6* variant alters SHM but not CSR.

The *MSH6^S502C^* mice and cells derived from them do not exhibit an increased mutation frequency or altered canonical MMR. Given that the S502C variant of our SLE model is in the MSH6 mismatch repair protein, we asked whether the mice exhibited a mutator phenotype. MutSα recognizes both small insertion and deletion loops and DNA mismatches. At dinucleotide repeat sites, if MutSα is not functioning properly, there is expansion or contraction of the sequence leading to microsatellite instability (MSI) (23). Analysis of mouse ear cells did not show any difference in MSI between the different genotypes (Figure 5A). We also employed the *Pig-a* assay, which measures the mutation frequency at the X-linked *Pig-a* gene by monitoring the loss of the glycosylphosphatidylinositol-anchored protein CD24 on red blood cells and reticulocytes. We determined via flow cytometry that the mutation frequency of *Pig-a* was similar among the genotypes (Figure 5C and D). To directly characterize the repair of DNA mismatches, we employed a previously established fluorescence-based host cell reactivation assay (31) and showed that the ability of mouse embryonic fibroblasts isolated from our mice to repair a DNA mismatch (G:G) is similar between the *Msh6^S502C^* and wildtype mice (Figure 5B). Overall, these data are in agreement with studies that used a yeast-based or human cell-free extract system (32,33), showing that this protein variant is able to function in canonical mismatch repair.

DISCUSSION

In this study, we provide evidence that the *MSH6^S502C^* GV is enriched in individuals with SLE. We then demonstrate that mice carrying the *Msh6^S502C^* allele develop several SLE-associated pathologies. Our results suggest that the *MSH6^S502C^* GV is linked to the development of lupus.

Currently, there is no cure for SLE. Therefore, it is important to understand the etiology of the disease to be able to discover new targets for treatment. Advancements in the field indicate that DNA repair plays a role in development of SLE given that lymphoid cells from patients with SLE are sensitive to DNA-damaging agents and are defective in the repair of DNA lesions (6–8). Additionally, specific DNA repair GVs in *OGG1*, *NEIL3*, *XRCC1*, and *POLB* have been linked to predisposition to SLE (5,30). In this study, we discovered that a GV in the mismatch repair pathway, the *MSH6^S502C^* GV, has an observed OR that is quite high for a complex disease such as SLE (34), suggesting that it is enriched in individuals with SLE.

MMR is important for antibody diversification given that human deficiencies in MMR proteins—MSH2 (35), MSH6 (36), or PMS2 (37)—are associated with impaired antibody maturation. Furthermore, studies in mice have shown that noncanonical MMR is important for the generation of mutations at A:T bps

during SHM, and CSR is diminished in mice that carry a deletion of the *MSH6* or *MSH2* gene (38,39). As shown in Figure 3D, AID deaminates cytosine to uracil, resulting in a U:G mispair. If uracil DNA glycosylase (UDG) binds to the U:G mispair, mutations during SHM are generally observed at G:C bps. However, if MSH2/6 bind to the U:G mispair, mutations are observed at A:T bps. Importantly, the absence of MSH6 and/or MSH2 leads to a 75% to 90% decrease in mutations at A:T bps (40).

In this study, we show that in mice harboring the *MSH6^S502C^* allele, SHM is altered. Specifically, we observe a significant increase in mutations at the WA motifs. Interestingly, the frequency of mutations at the opposite site, TW, is similar between *Msh6^S502C/S502C^* and *Msh6^+/+^* B cells. This indicates that there is a strand bias and that the presence of the *Msh6* variant leads to mutations at A:T bps on the nontranscribed strand (NTS) as opposed to the transcribed strand during SHM. It was previously shown that this preference for mutations induced on the NTS is most likely dependent on pol η because this strand bias is absent in mice that are deficient in pol η (41) (Figure 4D). Given that the MutSα complex stimulates pol η activity (42), the *MSH6^S503C^* variant may potentially stimulate pol η activity during SHM. An alternative hypothesis, though not necessarily mutually exclusive, is that given that UDG and MSH2/6 may compete for U:G mispairs, the MSH6 variant–MSH2 complex may outcompete UDG for binding to the U:G mispair, thus allowing for the increase in mutations at A and WA sites. The amino acid change that occurs in the MSH6 variant is in the DNA binding domain, and this change may increase the binding affinity of the MSH2/6 complex to U:G mispairs, resulting in the pol η mutation signature. If this is the case, we would also anticipate that canonical MMR will be unaffected or possibly even more accurate. Indeed, we do not observe increased mutagenesis in mice expressing the *MSH6^S503C^* variant, suggesting that canonical MMR does not differ between cells expressing wildtype or the variant *Msh6* (Figure 5). Additionally, a previous study indicated that MSH6 may act as a scaffold for AID to initiate deamination of cytosine (43), which can potentially alter the mutation frequency during SHM. However, we do not believe that the *MSH6^S503C^* variant is affecting AID recruitment to immunoglobulin genes because the overall mutation frequency between wildtype and *Msh6^S502C/S502C^* B cells is similar (Figure 4A). Finally, CSR in *Msh6^−/−^* B cells is decreased for switching to IgG1 and IgG3 by up to 40% and 70%, respectively, when compared with wildtype B cells (44). In our study, we do not observe this decrease in switching, suggesting that the *Msh6^S502C^* allele does not impact CSR.

An altered SHM repertoire can potentially result in increased levels of ANAs (45). Interestingly, a lupus-prone mouse model has been shown to harbor mutated AGC and AGT serine codons to arginine codons 66% of the time in autoreactive antibody V genes. Additionally, the AGC and AGT codons occur 1.5 times more frequently in V genes than other mouse genes (a similar trend is seen in humans), thus providing many sites that can lead

to arginine substitutions (46). Given that the *MSH6* GV B cells have a skewed SHM spectrum at A and W<u>A</u> sites (Figure 4), this could potentially result in AGC and AGT serine codons frequently mutating to the arginine codons CGC and CGT, respectively, when compared with wildtype cells. Arginine substitutions in the variable region of immunoglobulins are reactive to DNA (47). These ANAs can then bind to the nuclear antigens released by apoptotic cells and lead to inflammation, such as in the lungs of the *Msh6*[S502C/S502C] mice.

Apoptotic cells throughout the body release antigens, including DNA, as evidenced by circulating DNA in the blood and also as part of microparticles (for a review see ref. 48). The antibodies that arise in the *Msh6*[S502C/S502C] mice during SHM could bind to this DNA, potentially leading to immune complex formation. Because we do not observe immune complexes in the kidneys of our mice, it is unlikely that significant levels of circulating immune complexes are formed. Importantly, recognition of mismatches or oxidative damage in lung cells by the MSH2/6 heterodimer could eventually result in cell death as a result of futile cycling and/or activation of ATM/ATR (49). Given that the MSH2/6 subpathway of SHM seems to be preferentially used, it is possible that MSH6 also acts in dominant manner to bind to mismatches or oxidative damage in lungs, thereby increasing cell death in lungs compared to what is present in the lungs of wildtype mice. The presence of monocytes, along with increased levels of caspase 3 in the lungs of the *Msh6*[S502C/S502C] mice, indicates that there are dying cells in the lungs. These cells could release DNA antigens that are ultimately recognized by ANA generated as a result of the development of the skewed immune repertoire driven by the MSH6 variant protein during SHM, potentially promoting inflammation in the lung.

We discovered that MSH6 is highly expressed in the lungs and barely present in the kidneys of mice (Figure 2D). The minimal expression of MSH6 in kidneys may be the reason we did not observe kidney disease in the *Msh6*[S502C/S502C] mice (Figure 2B). Additionally, *Msh6* mRNA has been shown to be minimal in mouse skin tissue (50), which may explain why dermatitis incidence is not increased in the *Msh6*[S502C/S502C] mice (Figure 2C). In contrast, we do observe aberrant lung pathology in the mutant mice, an organ in which the expression of the MSH6 protein is high. The lungs are continually exposed to high levels of oxygen, which makes the cells susceptible to reactive oxygen species (ROS) (51). Interestingly, MutSα can recognize mispairs that involve DNA lesions, such as those produced by ROS (eg, 8-oxoguanine opposite adenine) (52). If the MSH6 variant does not process the lesion correctly, it can potentially lead to apoptosis by futile cycling, a process in which MMR stimulates DNA excision and resynthesis, which eventually creates strand breaks that ultimately induce cell death. Apoptosis can also occur via recognition of DNA damage by MMR that subsequently leads to a signal transduction cascade that activates the cell cycle checkpoint and eventually apoptosis (53). Given that expression of MSH6 is

high in the lung, the MSH6 variant can potentially recognize the abundant oxidative lesions that are in the lung cells and eventually lead to increased pulmonary apoptosis through either an apoptotic signaling cascade or futile cycling. Indeed, we detected an increase in apoptotic cells in the lungs of the *Msh6*[S502C/S502C] and *Msh6*[S502C/+] mice compared with wildtype mice (Figure 3A and B). The influx of apoptotic debris must be removed by phagocytic cells (54). As shown in Figure 3, there is concomitant increase in both apoptotic and CD68+ cells, which suggests that the monocytes and macrophages are in the lungs to remove the apoptotic cells in the *Msh6*[S502C/S502C] mice. The sustained levels of CD68+ cells and the increase in cells that express the pro- and anti-inflammatory macrophage markers (Supplementary Figure 2) in the lungs of *Msh6*[S502C/S502C] mice indicate that the immunologic microenvironment in lung alveoli is compromised. Interestingly, these pro- and anti-inflammatory macrophages have been both been implicated in SLE pathogenesis (55,56). However, we do note that oxidative lesions accumulate in aging mice (57), which should lead to the accumulation of apoptotic cells at older time points. Further investigations will have to be done to understand why an increase in apoptotic cells in the older *Msh6*[S502C/S502C] mice is not observed.

In summary, our results suggest that the *MSH6*[S503C] human genetic variant is linked to the development of SLE.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising is critically for important intellectual content, and all authors approved the final version to be published. Dr. Sweasy had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study conception and design. Meas, Sweasy, Criswell.
Acquisition of data. Meas, Nitiham, Taylor, Maher, Clairmont, Carufe, Kashgarian, Nottoli, Nagel, Gaffney.
Analysis and interpretation of data. Meas, Sweasy, Criswell.

REFERENCES

1. Thibault Flesher DL, Sun X, Behrens TW, Graham RR, Criswell LA. Recent advances in the genetics of systemic lupus erythematosus. Expert Rev Clin Immunol 2010;6:461–79.

2. Crispín JC, Hedrich CM, Tsokos GC. Gene-function studies in systemic lupus erythematosus. Nat Rev Rheumatol 2013;9:476–84.

3. Wang YF, Zhang Y, Lin Z, Zhang H, Wang TY, Cao Y, et al. Identification of 38 novel loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups. Nat Commun 2021;12:772.

4. Mohan C, Putterman C. Genetics and pathogenesis of systemic lupus erythematosus and lupus nephritis. Nat Rev Nephrol 2015;11: 329–41.

5. Senejani AG, Liu Y, Kidane D, Maher SE, Zeiss CJ, Park HJ, et al. Mutation of POLB causes lupus in mice. Cell Rep 2014;6:1–8.

6. Bashir S, Harris G, Denman MA, Blake DR, Winyard PG. Oxidative DNA damage and cellular sensitivity to oxidative stress in human autoimmune diseases. Ann Rheum Dis 1993;52:659–66.

7. Davies RC, Pettijohn K, Fike F, Wang J, Nahas SA, Tunuguntla R, et al. Defective DNA double-strand break repair in pediatric systemic lupus erythematosus. Arthritis Rheum 2012;64:568–78.

8. McCurdy D, Tai LQ, Frias S, Wang Z. Delayed repair of DNA damage by ionizing radiation in cells from patients with juvenile systemic lupus erythematosus and rheumatoid arthritis. Radiat Res 1997;147:48–54.

9. Modrich P. Mechanisms in E. coli and human mismatch repair (Nobel lecture). Angew Chem Int Ed Engl 2016;55:8490–501.

10. Genschel J, Modrich P. Mechanism of 5′-directed excision in human mismatch repair. Mol Cell 2003;12:1077–86.

11. Goellner EM, Smith CE, Campbell CS, Hombauer H, Desai A, Putnam CD, et al. PCNA and Msh2-Msh6 activate an Mlh1-Pms1 endonuclease pathway required for exo1-independent mismatch repair. Mol Cell 2014;55:291–304.

12. Kadyrov FA, Dzantiev L, Constantin N, Modrich P. Endonucleolytic function of MutLα in human mismatch repair. Cell 2006;126:297–308.

13. Longley MJ, Pierce AJ, Modrich P. DNA polymerase delta is required for human mismatch repair in vitro. J Biol Chem 1997;272:10917–21.

14. Bowen N, Kolodner RD. Reconstitution of saccharomyces cerevisiae DNA polymerase e-dependent mismatch repair with purified proteins. Proc Natl Acad Sci U S A 2017;114:3607–12.

15. Hegan DC, Narayanan L, Jirik FR, Edelmann W, Liskay RM, Glazer PM. Differing patterns of genetic instability in mice deficient in the mismatch repair genes Pms2, Mlh1, Msh2, Msh3 and Msh6. Carcinogenesis 2006;27:2402–8.

16. Zan H, Zhang J, Ardeshna S, Xu Z, Park SR, Casali P. Lupus-prone MRL/fas^lpr/lpr mice display increased AID expression and extensive DNA lesions, comprising deletions and insertions, in the immunoglobulin locus: concurrent upregulation of somatic hypermutation and class switch DNA recombination. Autoimmunity 2009;42:89–103.

17. Diogo D, Bastarache L, Liao KP, Graham RR, Fulton RS, Greenberg JD, et al. TYK2 protein-coding variants protect against rheumatoid arthritis and autoimmunity, with no evidence of major pleiotropic effects on non-autoimmune complex traits. PLoS One 2015;10:e0122271.

18. Webb R, Kelly JA, Somers EC, Hughes T, Kaufman KM, Sanchez E, et al. Early disease onset is predicted by a higher genetic risk for lupus and is associated with a more severe phenotype in lupus patients. Ann Rheum Dis 2011;70:151–6.

19. Barnetson RA, Cartwright N, van Vliet A, Haq N, Drew K, Farrington S, et al. Classification of ambiguous mutations in DNA mismatch repair genes identified in a population-based study of colorectal cancer. Hum Mutat 2008;29:367–74.

20. Tejon G, Hidalgo Y, Rosa Bono M, Rosemblatt M. A spontaneous mouse model of lupus: physiology and therapy. In: Lionaki S, editor. Lupus - new advances and challenges. London: IntechOpen; 2019. URL: https://www.intechopen.com/books/lupus-new-advances-and-challenges/a-spontaneous-mouse-model-of-lupus-physiology-and-therapy.

21. Li Z, Pearlman AH, Hsieh P. DNA mismatch repair and the DNA damage response. DNA Repair (Amst) 2016;38:94–101.

22. Edelmann W, Yang K, Umar A, Heyer J, Lau K, Fan K, et al. Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. Cell 1997;91:467–77.

23. Yang G, Scherer SJ, Shell SS, Yang K, Kim M, Lipkin M, et al. Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility. Cancer Cell 2004;6:139–50.

24. Quadrelli SA, Alvarez C, Arce SC, Paz L, Sarano J, Sobrino EM, et al. Pulmonary involvement of systemic lupus erythematosus: analysis of 90 necropsies. Lupus 2009;18:1053–60.

25. Mahajan A, Herrmann M, Muñoz LE. Clearance deficiency and cell death pathways: a model for the pathogenesis of SLE. Front Immunol 2016;7:35.

26. Funes SC, Rios M, Escobar-Vera J, Kalergis AM. Implications of macrophage polarization in autoimmunity. Immunology 2018;154:186–95.

27. Brink R. The imperfect control of self-reactive germinal center B cells. Curr Opin Immunol 2014;28:97–101.

28. Liu S, Cerutti A, Casali P, Crow MK. Ongoing immunoglobulin class switch DNA recombination in lupus B cells: analysis of switch regulatory regions. Autoimmunity 2004;37:431–43.

29. Saribasak H, Gearhart PJ. Does DNA repair occur during somatic hypermutation? [Review article]. Semin Immunol 2012;24:287–92.

30. Meas R, Burak MJ, Sweasy JB. DNA repair and systemic lupus erythematosus. DNA Repair (Amst) 2017;56:174–82.

31. Nagel ZD, Margulies CM, Chaim IA, McRee SK, Mazzucato P, Ahmad A, et al. Multiplexed DNA repair assays for multiple lesions and multiple doses via transcription inhibition and transcriptional mutagenesis. Proc Natl Acad Sci U S A 2014;111:e1823–32.

32. Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, de Wind N. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. Hum Mutat 2012;33:488–94.

33. Martinez SL, Kolodner RD. Functional analysis of human mismatch repair gene mutations identifies weak alleles and polymorphisms capable of polygenic interactions. Proc Natl Acad Sci U S A 2010;107:5070–5.

34. Ramos PS, Brown EE, Kimberly RP, Langefeld CD. Genetic factors predisposing to systemic lupus erythematosus and lupus nephritis. Semin Nephrol 2010;30:164–76.

35. Whiteside D, McLeod R, Graham G, Steckley JL, Booth K, Somerville MJ, et al. A homozygous germ-line mutation in the human MSH2 gene predisposes to hematological malignancy and multiple café-au-lait spots. Cancer Res 2002;62:359–62.

36. Gardès P, Forveille M, Alyanakian MA, Aucouturier P, Ilencikova D, Leroux D, et al. Human MSH6 deficiency is associated with impaired antibody maturation. J Immunol 2012;188:2023–9.

37. Péron S, Metin A, Gardès P, Alyanakian MA, Sheridan E, Kratz CP, et al. Human PMS2 deficiency is associated with impaired immunoglobulin class switch recombination. J Exp Med 2008;205:2465–72.

38. Martomo SA, Yang WW, Gearhart PJ. A role for Msh6 but not Msh3 in somatic hypermutation and class switch recombination. J Exp Med 2004;200:61–8.

39. Ehrenstein MR, Neuberger MS. Deficiency in Msh2 affects the efficiency and local sequence specificity of immunoglobulin class-switch recombination: parallels with somatic hypermutation. EMBO J 1999;18:3484–90.

40. Zanotti KJ, Gearhart PJ. Antibody diversification caused by disrupted mismatch repair and promiscuous DNA polymerases. DNA Repair (Amst) 2016;38:110–16.

41. Martomo SA, Yang WW, Wersto RP, Ohkumo T, Kondo Y, Yokoi M, et al. Different mutation signatures in DNA polymerase η- and MSH6-deficient mice suggest separate roles in antibody diversification. Proc Natl Acad Sci U S A 2005;102:8656–61.

42. Wilson TM, Vaisman A, Martomo SA, Sullivan P, Lan L, Hanaoka F, et al. MSH2-MSH6 stimulates DNA polymerase eta, suggesting a role for A:T mutations in antibody genes. J Exp Med 2005;201:637–45.

43. Li Z, Zhao C, Iglesias-Ussel MD, Polonskaya Z, Zhuang M, Yang G, et al. The mismatch repair protein Msh6 influences the in vivo AID targeting to the Ig locus. Immunity 2006;24:393–403.

44. Li Z, Scherer SJ, Ronai D, Iglesias-Ussel MD, Peled JU, Bardwell PD, et al. Examination of Msh6- and Msh3-deficient mice in class

switching reveals overlapping and distinct roles of MutS homologues in antibody diversification. J Exp Med 2004;200:47–59.

45. Bashford-Rogers RJ, Smith KG, Thomas DC. Antibody repertoire analysis in polygenic autoimmune diseases. Immunology 2018;155: 3–17.

46. Guo W, Smith D, Aviszus K, Detanico T, Heiser RA, Wysocki LJ. Somatic hypermutation as a generator of antinuclear antibodies in a murine model of systemic autoimmunity. J Exp Med 2010;207: 2225–37.

47. Li Z, Schettino EW, Padlan EA, Ikematsu H, Casali P. Structure-function analysis of a lupus anti-DNA autoantibody: central role of the heavy chain complementarity-determining region 3 Arg in binding of double- and single-stranded DNA. Eur J Immunol 2000;30: 2015–26.

48. Pisetsky DS. Evolving story of autoantibodies in systemic lupus erythematosus. J Autoimmun 2020;110:102356.

49. Li GM. Mechanisms and functions of DNA mismatch repair. Cell Res 2008;18:85–98.

50. Joost S, Annusver K, Jacob T, Sun X, Dalessandri T, Sivan U, et al. The molecular anatomy of mouse skin during hair growth and rest. Cell Stem Cell 2020;26:441–57.

51. Park HS, Kim SR, Lee YC. Impact of oxidative stress on lung diseases. Respirology 2009;14:27–38.

52. Ni TT, Marsischky GT, Kolodner RD. MSH2 and MSH6 are required for removal of adenine misincorporated opposite 8-oxo-guanine in S. cerevisiae. Mol Cell 1999;4:439–44.

53. Bellacosa A. Functional interactions and signaling properties of mammalian DNA mismatch repair proteins. Cell Death Differ 2001;8: 1076–92.

54. Gordon S, Plüddemann A. Macrophage clearance of apoptotic cells: a critical assessment. Front Immunol 2018;9:127.

55. Triantafyllopoulou A, Franzke CW, Seshan SV, Perino G, Kalliolias GD, Ramanujam M, et al. Proliferative lesions and metalloproteinase activity in murine lupus nephritis mediated by type I interferons and macrophages. Proc Natl Acad Sci U S A 2010;107:3012–7.

56. Orme J, Mohan C. Macrophage subpopulations in systemic lupus erythematosus. Discov Med 2012;13:151–8.

57. Hamilton ML, van Remmen H, Drake JA, Yang H, Guo ZM, Kewitt K, et al. Does oxidative damage to DNA increase with age? [Original article]. Proc Natl Acad Sci U S A 2001;98:10469–74.

58. Van Es JH, Gmelig Meyling FH, van de Akker WR, Aanstoot H, Derksen RH, Logtenberg T. Somatic mutations in the variable regions of a human IgG anti-double-stranded DNA autoantibody suggest a role for antigen in the induction of systemic lupus erythematosus. J Exp Med 1991;173:461–70.