Every year, the number of non-MHC genes with high contribution to SLE susceptibility has increased. However, the association between the copy number status and SLE susceptibility was analyzed in 334 SLE patients and 338 controls. CCL3L3-null status was significantly associated with SLE susceptibility (OR > 18, P < 0.0001), which remained significant by Bonferroni’s correction (corrected P = 0.0007). However, the significant association between C4B low-copy status and SLE susceptibility (OR = 1.6051, P = 0.0331) became non-significant by Bonferroni’s correction (corrected P = 0.3938). Except for these results, no other significant association between SLE susceptibility and copy number status in other genes was observed. The CCL3L3-null status may be a significant factor for SLE susceptibility.
the highest scoring loci in the non-MHC locus is \( \text{TNFAIP3} \) (geneID: 7128), which is associated with various autoimmune diseases including SLE\(^{17–19} \). Although the association of CNVs in \( \text{TNFAIP3} \) and \( \text{TNIP1} \) with SLE susceptibility has not been reported, associations with RA susceptibility have been reported\(^{20} \). Therefore, the contribution of CNs in \( \text{TNFAIP3} \) and \( \text{TNIP1} \) in SLE susceptibility was also evaluated in the present study. To evaluate the CNs, the modified real competitive polymerase chain reaction (mrcPCR) method was employed.

**Results**

**Establishment of mrcPCR assays for CN determination.** The mrcPCR method, which measures CN by estimating the signals from the amplified gene products of interest relative to those from the spiked internal reference sequences, was shown to be accurate and simple for determining CNs\(^{21,22} \). To establish mrcPCR assays for \( \text{TNFAIP3} \), \( \text{TNIP1} \), \( \text{IL12B} \), \( \text{TBX21} \), \( \text{TLR7} \), \( \text{C4A} \), and \( \text{C4B} \) in the present study, modified bases were introduced to produce competitors as shown in Fig. 1.

Some inconsistencies in disease association studies of the \( \text{CCL3L} \) family clusters may be related to assays which provide heterogeneous results due to their designs based on incomplete information on the \( \text{CCL3L} \) family cluster genes\(^{23} \). To avoid these possible errors, specific PCR primers and extension primers for \( \text{CCL3L1} \) and \( \text{CCL3L3} \) were designed for the mrcPCR assay in the present study (Fig. 1B) using \( \text{IGF1} \) as a control gene.

To confirm the specific amplification of \( \text{CCL3L1} \) and \( \text{CCL3L3} \) by the mrcPCR assay, Sanger sequencing of the PCR products in the mrcPCR assay PCR product (red box).
mrcPCR assay was performed and confirmed that only the sequences from CCL3L1 and CCL3L3 were specifically amplified (Fig. 1C). This suggests that the relative CN deduced from the products in our mrcPCR assay was not confounded by other homologous sequences such as CCL3 (geneID: 6348) and CCL3L2 (geneID: 390788).

**Determination of CNs by mrcPCR.** The mrcPCR assay was performed on control and SLE samples to determine the CNs of TNFAIP3, TNIP1, IL12B, TBX21, C4A/C4B, CCL3L1, and CCL3L3. To optimize the assays, the relative primer concentrations and competitors were empirically determined, and the final mrcPCR assay information is shown in Supplementary Tables S1–S3. Representative mrcPCR results are shown in Fig. 2. The relative CNs were determined using the relative peak ratios from the mrcPCR results as previously reported21,22. After the median peak ratios values for each gene from the mrcPCR were obtained, the raw peak ratio data were divided by half of the median value, and the standardized CN (sCN) values were employed for the comparison. In genes such as TNFAIP3, TNIP1, IL12B, and TLR7, CNV was absent or very low (Fig. 3). When the CNs between the controls and SLE patients were compared by Wilcoxon rank-sum tests, most genes were significant (TNFAIP3, P < 0.0001; IL12B, P < 0.0001; TBX21, P = 0.0043; and TLR7, P < 0.0001 for both male and female), except for TNIP1 (P = 0.7034). However, most of the differences seemed to be related to experimental variations, as the CNs for the controls or SLE cases did not show any separate, distinct groups but showed continuous values (Fig. 3). When the CNs were compared by the Chi-squared or Fisher’s exact tests after the sCNs were transformed into digitized CNs (dCNs) as described in the Methods section, no significant CN difference between the controls and SLE cases was observed for the TNFAIP3, TNIP1, IL12B, TBX21, and TLR7 genes (Table 1), suggesting that the direct comparison of CNs may lead to false-positive results. Although we found several distinct CNV cases in IL12B (Fig. 3C,E) in our dataset, the cases with CNVs were quite limited. In a male SLE patient with two copies of TLR7 in the X chromosome, the sex was confirmed to be male by a short tandem repeat marker test, suggesting the presence of TLR7 CN variants also in the Korean population.

In genes such as C4A, C4B, CCL3L1, and CCL3L3, the CNVs were relatively high (Fig. 4). In a comparison of the median CNs in those genes by the Wilcoxon rank-sum test, only C4A showed a significant difference compared to the controls.
The significant association of CNV in C4A with SLE may be related to experimental variations because the significance was lost when the dCNs were compared. The dCNs were compared, C4B (P = 0.0379) and CCL3L3 (P = 0.0002) CNs were significantly associated with SLE susceptibility (Table 1). Especially, the dCNs for low C4B (OR = 1.6051, P = 0.0331) and CCL3L3-null status (OR > 18.5355, P < 0.0001) were significantly associated with SLE (Table 1). After Bonferroni’s correction, the significant association between C4B low-copy status and SLE susceptibility became non-significant (corrected P = 0.2414).

Figure 3. Difference in CNs between the controls and SLE patients in genes showing relatively low or no CNV. (A) TNFAIP3 (P < 0.0001). (B) TNIP1 (P = 0.7035). (C) IL12B (P < 0.0001). (D) TBX21 (P = 0.0043). (E) TLR7 (P < 0.0001 for both male and female). Although statistical significance was observed in several genes by the Wilcoxon rank-sum test, CNVs were not evident except in IL12B and TLR7 in a few cases.

Table 1. Correlation between gene copy number and SLE. 1 dCN, digitized copy number: the standardized copy number (sCN) in which the raw copy number data were divided by half of their median value was digitized depending upon the sCN range as described in “Materials and methods”. 2 P-value was estimated by the Chi-squared test (†) or Fisher’s exact test (‡). Significant differences are marked by boldface. P-values adjusted for multiple testing with Bonferroni’s correction method (11 tests) are shown in parentheses. * No statistical test was performed because there was no case with a copy number of 1 for the TLR7 gene in either the SLE patients or the controls. 3 TLR7 (M), TLR7 data for males. 4 TLR7 (F), TLR7 data for females. 5 C4B and 6 CCL3L3 were analyzed for the contributions of a copy status of 1 (for C4B) or 0 (for CCL3L3) to SLE susceptibility. Odds Ratios (OR) were 1.6051 for 5 low C4B CN status, and > 18.5355 for 6 CCL3L3-null status in SLE patients.
Therefore, only CCL3L3-null status, which was still significant by Bonferroni’s correction (corrected \( P = 0.0007 \)), was considered a significant CNV for SLE in the present study.

We tested if the dCNs for C4A, C4B, CCL3L1, and CCL3L3 were in Hardy–Weinberg equilibrium by the method described previously. The \( P \)-values were significant in the controls for C4B (\( P = 2.0 \times 10^{-13} \)), suggesting that the dCN for C4B in the controls may be skewed and that the difference in C4B CN between the controls and SLE cases may not be valid. The \( P \)-values for the other cases were not significant (controls for CCL3L3, \( P = 0.8595 \); SLE cases for CCL3L3, \( P = 1.0 \); controls for CCL3L1, \( P = 1.0 \); SLE cases for C4A, \( P = 0.9481 \); SLE cases for C4B, \( P = 0.8447 \)).

In the analysis of associations between CCL3L3-null status and clinical variables, CCL3L3-null status showed a significant association with non-scarring alopecia (\( P = 0.0434 \)), whereas CCL3L1-null status did not (\( P = 0.2048 \)). Except for non-scarring alopecia, CCL3L3-null status did not show any significant association with any of the other clinical variables (Table S4).

Confirmation of CCL3L3-null copy status in SLE patients by PCR sequencing and digital droplet PCR. In our mrcPCR results, CCL3L3-null copy status showed the most significant correlation with SLE susceptibility. To confirm the CCL3L3-null copy status, first, we sequenced the PCR products produced in the mrcPCR assay in the CCL3L1-null and CCL3L3-null cases and found that the PCR products in the null cases did not contain the null-gene-specific base (Supplementary Figure S1), suggesting the specificity of the mrcPCR assay for CCL3L1 and CCL3L3. Then, we established a digital droplet (dd)PCR assay to confirm the CCL3L3-null copy status in our mrcPCR results. We observed both higher and lower-level signals from CCL3L1 or CCL3L3 using the probe for CCL3L1 (Fig. 5A,B), and the higher and lower level signals were switched when the probe for CCL3L3 was employed (Fig. 5C,D), suggesting that the lower-level signals for each probe were related to cross-reactions. To evaluate CCL3L1 and CCL3L3 CNs by the ddPCR assay, the results from the CCL3L1 probes were employed as the signals from CCL3L1 and CCL3L3 were well-separated. In the comparison of the results between mrcPCR and ddPCR, the correlations were significant (\( P < 0.0001 \) for both CCL3L1 and CCL3L3, Fig. 5E,F), confirming that the quantitative CN results from mrcPCR were comparable to those from ddPCR. All five randomly selected CCL3L3-null cases from the mrcPCR assay were also negative in the ddPCR assays (five cases were shown in Fig. 5A,C), confirming again that the CCL3L3-null status estimated by mrcPCR was comparable to that from ddPCR.

Discussion

Although CNVs in many immunity-related genes have implications in SLE susceptibility, the contribution of CNVs to SLE susceptibility awaits more investigation. The present study evaluated the CNs for immunity-related non-MHC genes such as TNFAIP3, TNIP1, IL12B, Tbx21 (T-bet), TLR7, CCL3L1, and CCL3L3, and MHC genes such as C4A and C4B using mrcPCR assays in 327 SLE patients and 338 controls. Among the CNVs for nine immunity-related genes tested in the present study, a significant correlation between CCL3L3-null status (\( P < 0.0001 \)) and SLE was found.
CNVs in various non-MHC immunity-related genes including \( TNFAIP3 \)\(^{20,25}\), \( TNIP1 \)\(^{20}\), \( IL12B \)\(^{3}\), \( TBX21 \)\(^{3}\), and \( TLR7 \)\(^{4,10,11}\) have been reported to have an association with autoimmune diseases. Also, the correlation of CNVs in \( IL12B \), \( TBX21 \), and \( TLR7 \) with human SLE susceptibility has been reported, and the CNs over 2 were 11.3% for \( IL12B \), 8.5% for \( TBX21 \), and 21.6% for \( TLR7 \) in SLE patients in the previous reports\(^{26,27}\). However, when dCNs were analyzed in the present study, few CNVs were found in \( IL12B \), \( TBX21 \), and \( TLR7 \) in the SLE patients and controls, which fact may be related to ethnic differences. For \( TNFAIP3 \) and \( TNIP1 \), it has already been reported that their CNVs are rare (0.2 and 0.4%, respectively)\(^{20}\), which result was reproduced in the present study. Given the few variants in those genes in the present study, therefore, we could draw any conclusions on the significance of CNVs in \( TNFAIP3 \), \( TNIP1 \), \( IL12B \), \( TBX21 \), and \( TLR7 \) to the SLE susceptibility of a Korean population.

\( C4A \) and \( C4B \) showed high CNV in the present study, and one copy of \( C4B \) (OR = 1.6051, \( P = 0.0331 \)) was significantly associated with SLE, which is partially consistent with a previous report that suggested that one copy of \( C4A \) (OR = 1.613, \( P = 0.022 \)) was a risk factor for SLE susceptibility\(^{7}\). However, we could not find \( C4A \)-null cases, which was suggested as a strong risk factor for SLE (OR = 5.267, \( P = 0.001 \)) in the previous study\(^{7}\). Again, this might be related to ethnic differences. In addition, Bonferroni’s correction of the association of low \( C4B \) copy status lost significance (\( P = 0.3938 \)) in the present study. In further analysis of Hardy–Weinberg Equilibrium (HWE), the CN distribution of \( C4B \) in the control group was not in HWE, suggesting that the comparison between the controls and cases may not be valid for the present study cases. Therefore, the correlation between \( C4A \) or \( C4B \)-null status and SLE susceptibility was not confirmed in the present study, probably due to the absence of \( C4A \) or \( C4B \)-null cases in our study cohort, which might be related to ethnic differences.

Figure 5. Correlation between mrcPCR and ddPCR results for \( CCL3L1 \) and \( CCL3L3 \). (A) ddPCR results for \( CCL3L1 \) and \( CCL3L3 \) using a \( CCL3L1 \)-specific probe in a \( CCL3L1 \)-null case, five \( CCL3L3 \)-null cases, and a case with both alleles. In addition to specific signals from \( CCL3L1 \) (marked as \( CCL3L1 \)), weak signals from the \( CCL3L3 \) sequence (marked as \( CCL3L3 \)) were also detected. (B) ddPCR results for the control \( GAPDH \) gene with a probe labeled by fluorescein phosphoramidite (FAM) in assays using a \( CCL3L1 \)-specific probe. (C) ddPCR results for \( CCL3L1 \) and \( CCL3L3 \) using a \( CCL3L3 \)-specific probe in the same cases as in (A). In addition to specific signals from \( CCL3L3 \) (marked as \( CCL3L3 \)), weak signals from the \( CCL3L1 \) sequence (marked as \( CCL3L1 \)) were also detected. (D) ddPCR results for the control \( GAPDH \) with a probe labeled by Hexachloro-Fluorescein in assays for the \( CCL3L3 \)-specific probe in (C). (E) Linear correlation between mrcPCR and ddPCR results for \( CCL3L1 \) (\( R^2 = 0.9925, P < 0.0001, N = 16 \)). (F) Linear correlation between mrcPCR and ddPCR results for \( CCL3L3 \) (\( R^2 = 0.9912, P < 0.0001, N = 16 \)). For (E) and (F), three cases for each dCN (0–4) by mrcPCR assay were analyzed. Figures (A) to (D) for ddPCR results were obtained from QuantaSoft analysis software version 1.74 (BioRad Laboratories).
CCL3 (geneID: 6348), CCL3L1 (geneID: 6349), CCL3L2 (geneID: 390788), and CCL3L3 (geneID: 414062) are CCL3-related genes located closely on chromosome 17q12. CCL3-related genes share over 95% sequence identity at both the genomic and amino acid levels, and they encode macrophage inflammatory protein (MIP)-1α, which is secreted from epithelial cells, lymphocytes, platelets, and macrophages.26–28. MIP-1α is a chemokine acting as a pro-inflammatory cytokine on immune cells including CD8+ T cells and dendritic cells via the CCR5 receptor.29

In a previous report, SLE patients had a trend to have higher concentrations of MIP-1α and higher serum levels of MIP-1α was associated with discoid lupus.31. Also, serum MIP-1α level was higher in patients with active renal disease than those without.32. Lower CCL3L1 CN was first reported to be associated with enhanced HIV/acquired immunodeficiency syndrome (AIDS) susceptibility33, and with the durability of immune recovery during anti-HIV-1 therapy.34. The influence of higher CCL3L1 CN in autoimmune diseases such as SLE, rheumatoid arthritis, and Kawasaki disease has been reported. However, the non-specificity of the assays used for determining the CNs in those studies due to high sequence similarity among CCL3-related genes has been raised, and the authors argued the necessity of new methodologies to specifically measure highly homologous CCL3-related genes. In addition, a report showed that the rounded CCL3L1-CNs were not in HWE, raising issues on the more careful interpretation of CN data.

The present study tested the association between CCL3L1/CCL3L3 CNs and SLE susceptibility using an mrcPCR assay, which was designed for the determination of CNs specific for CCL3L1 and CCL3L3, and was not confounded by each other or the other CCL3L-related genes that may have confounded the estimation of CNs for CCL3L1 or CCL3L3 as indicated previously.36. The specificity of our mrcPCR assay for CCL3L1 and CCL3L3 was confirmed by Sanger sequencing in the CCL3L1-null and CCL3L3-null cases. In addition, the distribution of CCL3L1 and CCL3L3 CNs by mrcPCR did not deviate from HWE in the present study, suggesting that our mrcPCR assay could be a useful tool for the validation of previous associations with susceptibility to HIV/AIDS or autoimmune diseases including RA and Kawasaki disease.37. The significance of CCL3L3-null status in SLE (OR > 17, P < 0.0001) was maintained after Bonferroni’s correction (P = 0.0008). CCL3L3-null status by mrcPCR was confirmed by the ddPCR assay, which was developed in the present study, along with the linearity of the results obtained from the mrcPCR and ddPCR assays. Therefore, our results suggest that CCL3L3-null status may be a significant factor for SLE susceptibility in the Korean population.

CCL3-related genes encode MIP-1α which is a ligand for the CCR5 receptor. CCR5 is also the co-receptor used by the HIV-1 virus for cell entry. Therefore, MIP-1α and the HIV-1 virus compete for the CCR5 on lymphocytes. CCL3 and CCL3L1 encode protein products that differ in 3 amino acids,39 but their inhibitory effects on viral replication are tenfold different, suggesting the different roles of the CCL3-related genes in disease susceptibility. However, CCL3L1 and CCL3L3 have three identical exons and encode identical proteins, and their protein products have been posited as having the same functions. Therefore, the distinction between CCL3L1 and CCL3L3 in their CN estimation was unimportant so far. The present study showed a significant association of CCL3L3-null but not CCL3L1-null status with SLE susceptibility, suggesting that CCL3L1 and CCL3L3 may have distinctive roles, and raising the question of the necessity for separate evaluation of their CNs for disease susceptibility risk assessment. Although CCL3L1 and CCL3L3 encode identical protein, they might have a differential role in SLE susceptibility; pertinent hypotheses, however, are not yet available. Cells from CCL3L1-null and CCL3L3-null cases may yield more insights into their possibly differential expression or roles. The present study’s significant association of CCL3L3-null status but not of CCL3L1-null status with non-scarring alopecia may also suggest their differential roles. Meanwhile, we must seek confirmation.

This study has limitations. First, the participants were ethnically limited to Korean patients, which could reduce the generalizability of our results. Secondly, this study analyzed the association between CCL3L3-null status and clinical manifestations of SLE based on a relatively small sample size. Further studies on various ethnic backgrounds with a larger number of SLE cases will be necessary in order to more fully explicate the significance of CCL3L3-null status.

Conclusion

CCL3L3-null status may be a significant factor for SLE susceptibility in the Korean population.

Materials and methods

Patients and controls. SLE patients were recruited from a rheumatology outpatient clinic in Seoul National University Hospital. All of them had been diagnosed and followed by certified rheumatologists and met the 2019 European League Against Rheumatism and the American College of Rheumatology classification criteria.41 The control group samples were age- and sex- matched healthy participants of Korean national health screening program. The use of samples and clinical information was approved by the Institutional Review Boards of Seoul National University Hospital and National Cancer Center and informed consent was obtained from all the participants. All methods were performed in accordance with the relevant guidelines and regulations.

DNA isolation. DNA from blood cells was isolated using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA, USA) and TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) for DNA solubilization. The purified DNA stock was maintained at − 80 °C, and diluted DNA (10 ng/μL) made from the stock using distilled water (Gibco, Carlsbad, CA, USA) was stored at −20 °C until use.

Cloning of competitor DNA sequences for modified real competitive PCR. To obtain the competitor sequences, the sequences for each gene were amplified with the primer pairs in Supplementary Table S1 except for CCL3L1 or CCL3L3, which employed the following primers: CAA GGT GTT TGG CAG GCG TTT AAG and CTC TGC ACC ACG TGA GTC CAT GTT GGT. After purification of the amplified products and clon-
ing into the pGEM-T Easy Vector (Promega, Madison, WI, USA), a Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA, USA) was employed to introduce artificial base changes into the competitor sequences (Fig. 1A and Supplementary Figure S1). The previously reported IGF1 competitor3, where two bases were changed (Fig. 1B), was employed. The cloned competitors were digested with the restriction enzyme SalI to reduce non-specific amplification due to the closed circular plasmid structure. The competitors were diluted and aliquoted.

**Establishment of mrcPCR assay for the determination of CNs.** PCR amplification, the purification of amplified products, primer extension reactions in the mrcPCR assay were performed as previously reported21. The relative peak heights for single-base-extended products were analyzed using a GeneMapper software ver. 5.0 (Thermo Fisher Scientific), and the relative CNs from the peak heights were analyzed as previously reported21,22. The PCR primers for mrcPCR are shown in Supplementary Table S1. For the simultaneous amplification of the genomic sequence and competitor sequence, the diluted competitor(s) was spiked into genomic DNA, and the PCR primers for the control gene, IGF1, were added together with the PCR primers for the gene of interest. The amount of PCR primers and the spiked competitors were determined empirically, and the amount per reaction employed in the present study is shown in Supplementary Table S2. The primers employed for the extension reaction are shown in Supplementary Table S3, along with information on the amount per extension reaction used in the present study.

The raw relative CN data for each gene were divided by half of their median value, and the resulting standardized CN (sCN) was employed for comparison by the Mann–Whitney U test. The sCNs were converted to digitized CNs (dCN) as follows: 0 ≤ sCN < 0.5, 0; 0.5 ≤ sCN < 1.5, 1; 1.5 ≤ sCN < 2.5, 2; 2.5 ≤ sCN < 3.5, 3; 3.5 ≤ sCN < 4.5, 4; 4.5 ≤ sCN < 5.5, 5; and 5.5 ≤ sCN < 6.5, 6.

**Droplet digital PCR for CCL3L1 and CCL3L3 CNs.** ddPCR was carried out according to the manufacturer's protocol (QX100; BioRad Laboratories, Hercules, CA, USA). The reaction mixture was prepared according to the protocol for 2 × ddPCR Supermix (BioRad Laboratories) with 20 × primers and probes (final concentrations of 900 and 250 nM, respectively), and 25 ng of template DNA. In the reaction, the PCR amplification primers for CCL3L1 and CCL3L3 were the same ones employed for the mrcPCR assay (Supplementary Table S1). The detection probes for CCL3L1 and CCL3L3 were 5′-Hexachloro-Fluorescein (HEX)- GTC TTT TTT TGC GGC CTC AGA AGC-BHQ1-3′ and 5′-FAM- GTC TTT TTT TGT GCC CTC AGA AGC-BHQ1-3′, respectively. GAPDH was employed as a reference gene with the following primers and probe: 5′-TGC CTT CCT GCC TCT TGT CT-3′ (forward), 5′-AAC GAA GGG GTC ATT GAT GG-3′ (reverse) for amplification primers and 5′-HEX- TCA CCA GGG CTG CTG TTA AC-BHQ1-3′ (probe employed for CCL3L1-specific probe) or 5′-HEX-TCA CCA GGG CTG CTT TTA AC-BHQ1-3′ (probe employed for CCL3L3-specific probe) for the probe. Each reaction mixture was loaded into a sample well of an eight-channel disposable droplet generator cartridge (BioRad Laboratories). The emulsified samples were generated from a droplet generator (QX100; BioRad Laboratories) and then transferred into a 96-well plate. After heat-sealing with foil seal, the emulsified samples underwent a 2-step thermal cycling protocol in a T-100 Touch Thermal Cycler (BioRad Laboratories) as follows: 95 °C for 10 min, 40 cycles of 95 °C for 30 s and 55 °C for 60 s (ramp rate set to 2 °C per second), and 98 °C for 10 min. The 96-well droplet PCR plates were loaded into a droplet reader (BioRad Laboratories), which automatically read the droplets from each well of the plate. Analysis of the ddPCR data was performed with QuantaSoft analysis software version 1.74 (BioRad Laboratories).

**Statistical analyses.** Age- and sex-matched SLE patients (N = 368) and controls (N = 375) were accrued. However, the cases with less than 100 ng of purified genomic DNA were excluded (N = 5 for SLE patients), and those showing failure in the mrcPCR assays for any of the seven genes (N = 36 for SLE patients, and N = 37 for controls) were excluded. Therefore, the results from 327 SLE patients and 338 controls were finally analyzed. The differences in CNs between the SLE patients and controls were evaluated by the Wilcoxon rank-sum test for continuous variables, and the Chi-squared test or Fisher’s exact test for categorical variables. The correlation of the CNs measured by ddPCR and mrcPCR was analyzed by linear regression. HWE of the CNs was evaluated by a previously reported method24, which estimated the expected frequencies using an estimation maximization approach and calculated the Pearson Chi-squared statistics for HWE from the expected and observed frequencies. All statistical analyses were performed using R software version 4.0.4 (R Core Team (2021). R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/).

Received: 30 April 2021; Accepted: 30 August 2021
Published online: 27 September 2021

**References**

1. Ghodke-Puranik, Y. & Niewold, T. B. Immunogenetics of systemic lupus erythematosus: A comprehensive review. J. Autoimmun. 64, 125–136. https://doi.org/10.1016/j.jaut.2015.08.004 (2015).
2. Deng, Y. & Tsao, B. P. Advances in lupus genetics and epigenetics. Curr. Opin. Rheumatol. 26, 482–492. https://doi.org/10.1097/ BOR.0000000000000886 (2014).
3. Yu, B. et al. Copy number variations of Interleukin-12B and T-bet are associated with systemic lupus erythematosus. Rheumatology (Oxford) 50, 1201–1205. https://doi.org/10.1093/rheumatology/keq439 (2011).
4. Garcia-Ortiz, H. et al. Association of TLR7 copy number variation with susceptibility to childhood-onset systemic lupus erythematosus in Mexican population. Ann. Rheum. Dis. 69, 1861–1865. https://doi.org/10.1136/ard.2009.124313 (2010).
5. Mantani, M. et al. CCL3L1 gene-containing segmental duplications and polymorphisms in CCR5 affect risk of systemic lupus erythematosus. *Ann. Rheum. Dis.* 67, 1076–1083. https://doi.org/10.1136/ard.2007.070848 (2008).
6. Pereira, K. M. et al. Low C4, C4A and C4B gene copy numbers are stronger risk factors for juvenile-onset than for adult-onset systemic lupus erythematosus. *Rheumatology* (Oxford) 55, 869–873. https://doi.org/10.1093/rheumatology/kev436 (2016).
7. Yang, Y. et al. Gene copy-number variation and associated polymorphisms of complement component C4 in human systemic lupus erythematosus (SLE): Low copy number is a risk factor for and high copy number is a protective factor against SLE susceptibility in African Americans. *Am. J. Hum. Genet.* 80, 1037–1054. https://doi.org/10.1086/518237 (2007).
8. Mosmann, T. R. & Coffman, R. L. TH1 and TH2 cells: Different patterns of lymphokine secretion lead to different functional properties. *Ann. Rev. Immunol.* 7, 145–173. https://doi.org/10.1146/annurev.immunol.070408.163038 (1989).
9. Wakeland, E. K., Liu, G., Graham, R. R. & Behrens, T. W. Delineating the genetic basis of systemic lupus erythematosus.
10. Andrews, B. S.
11. Adrianto, I.
12. Hauptmann, G., Grosshans, E. & Heid, E. Lupus erythematosus syndrome and complete deficiency of the fourth component of complement. *Boll Ist Sieroter Milan* 53(suppl), 228 (1974).
13. McKinney, C. et al. Evidence for an influence of chemokine ligand 3-like 1 (CCL3L1) gene copy number on susceptibility to rheumatoid arthritis. *Ann. Rheum. Dis.* 67, 409–413. https://doi.org/10.1136/ard.2007.070528 (2008).
14. Shamlov, R. & Aneskiевич, B. J. TNIP1 in autoimmune diseases: Regulation of toll-like receptor signaling. *J. Immunol. Res.* 2018, 941269. https://doi.org/10.1155/2018/941269 (2018).
15. Adrianto, I. et al. Association of two independent functional risk haplotypes in TNIP1 with systemic lupus erythematosus. *Arthritis Rheumatol.* 64, 3695–3705. https://doi.org/10.1002/art.34642 (2012).
16. Allanore, Y. et al. Genome-wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis. *PLoS Genet.* 7, e1002091. https://doi.org/10.1371/journal.pgen.1002091 (2011).
17. Flesher, D. L., Sun, X., Behrens, T. W., Graham, R. R. & Criswell, L. A. Recent advances in the genetics of systemic lupus erythematosus. *Expert Rev. Clin. Immunol.* 6, 461–479. https://doi.org/10.1586/09740396.6.4.461 (2010).
18. Mosunse, S. L. et al. Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. *Nat. Genet.* 40, 1062–1064. https://doi.org/10.1038/ng.202 (2008).
19. Graham, R. R. et al. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nat. Genet.* 40, 1059–1061. https://doi.org/10.1038/ng.200 (2008).
20. Uddin, M., Sturge, M., Rahman, P. & Woods, M. O. Autosomal-wide copy number variation association analysis for rheumatoid arthritis using the WTC/Cc high-density SNP genotype data. *J. Rheumatol.* 38, 797–801. https://doi.org/10.3899/jrheum.100758 (2011).
21. Kim, H. K. et al. Simple and versatile molecular method of copy-number measurement using cloned competitors. *PLoS One* 8, e69414. https://doi.org/10.1371/journal.pone.0069414 (2013).
22. Kim, Y. H. et al. FAK-cop gene gain is a predictive marker for sensitivity to FAK inhibition in breast cancer. *Cancers* https://doi.org/10.3390/cancers11091288 (2019).
23. Shrestha, S., Nyaku, M. & Edberg, J. C. Variations in CCL3L gene cluster sequence and non-specific gene copy numbers. *Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat. Genet.* 39, 1059–1061. https://doi.org/10.1038/ng.200 (2008).
24. Fritz, R. S.
25. Musone, S. L. et al. Sequencing of TNFAIP3 and association of variants with multiple autoimmune diseases. *Genes Immun.* 12, 176–182. https://doi.org/10.1038/gene.2010.64 (2011).
26. Yang, S. K., Eckmann, L., Panja, A. & Kagnoff, M. F. Differential and regulated expression of C-X-C, C-C, and C-chemokines by human intestinal epithelial cells. *Gastroenterology* 113, 1214–1223. https://doi.org/10.1053/gast.1997.v113.pm1932256 (1997).
27. Aihara, S. K. et al. CCL3L1-CCL5 genotype influences durability of immune recovery during antiretroviral therapy of HIV-1-infected individuals. *Nat. Med.* 14, 413–420. https://doi.org/10.1038/nm1741 (2008).
28. Ayabe, Y. et al. Genetic variations in the receptor-ligand pair CCR5 and CCL3L1 are important determinants of susceptibility to Kawasaki disease. *J. Infect. Dis.* 192, 344–349. https://doi.org/10.1086/403953 (2005).
29. Field, S. E. et al. Experimental aspects of copy number variant assays at CCL3L1. *Nat. Med.* 15, 1115–1117. https://doi.org/10.1038/nm1099-1115 (2009).
30. Berger, E. A., Murphy, P. M. & Farber, J. M. Chemokine receptors as HIV-1 coreceptors: Roles in viral entry, tropism, and disease. *Annu. Rev. Immunol.* 17, 657–700. https://doi.org/10.1146/annurev.immunol.17.1.657 (1999).
31. Lusso, P. HIV and the chemokine system: 10 years later. *EMBO J.* 25, 447–456. https://doi.org/10.1093/emboj/d609476 (2006).
32. Colobran, R., Pedrosa, E., Carretero-Iglesia, L. & Juan, M. Copy number variation in chemokine superfamily: The complex scene of CCL3L-CCL4L genes in health and disease. *Clin. Exp. Immunol.* 162, 41–52. https://doi.org/10.1111/j.1365-2491.2010.04224.x (2010).
33. Aguero, S. et al. The LD78beta isoform of MIP-1alpha is the most potent CC-chemokine in inhibiting CCR5-dependent human immunodeficiency virus type 1 replication in human macrophages. *J. Virol.* 75, 4402–4406. https://doi.org/10.1128/jvi.75.9.4402-4406.2001 (2001).
34. Aringer, M. et al. 2019 European League Against Rheumatism/American College of Rheumatology Classification Criteria for Systemic Lupus Erythematosus. *Arthritis Rheumatol.* 71, 1400–1412. https://doi.org/10.1002/art.40930 (2019).
Acknowledgements
The single-base extension reaction for mrcPCR analysis was supported by Dr. Jung-Ah Hwang at the Core Center of the National Cancer Center (NCC), Korea. The statistical analysis was supported by Joongyub Lee, MD. PhD at the Department of Preventive Medicine Seoul National University College of Medicine, Seoul, Korea.

Author contributions
Conceptualization, Y.-W.S., and K.-M.H.; methodology, Y.-H.K., and H.W.S.; validation, E.-K.K., and Y.-H.W.; data acquisition and formal analysis, H.-Y.K., and E.E.L.; statistical analysis, D.-e.L.; data curation, Y.-H.K., and E.E.L.; writing—original draft preparation, Y.-H.K., E.E.L., and K.-M.H.; writing—review and editing, Y.-W.S.; supervision, K.-M.H., and Y.-W.S.; and funding acquisition, K.-M.H., Y.-H.K., and Y.-W.S.

Funding
This study was funded by grants from the National Cancer Center, Korea (1910150 to K-M.H.), a grant from the National Research Foundation of Korea (NRF-2018R1D1A1B07048952 to Y-H.K), grants from the Ministry of Science, ICT, and Future Planning (NRF-2020M3E5E2037430 to Y-W.S.) and a grant from Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, Republic of Korea (HI13C1754).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-98531-6.

Correspondence and requests for materials should be addressed to K.-M.H. or Y.-W.S.

Reprints and permissions information is available at www.nature.com/reprints.

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

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

© The Author(s) 2021