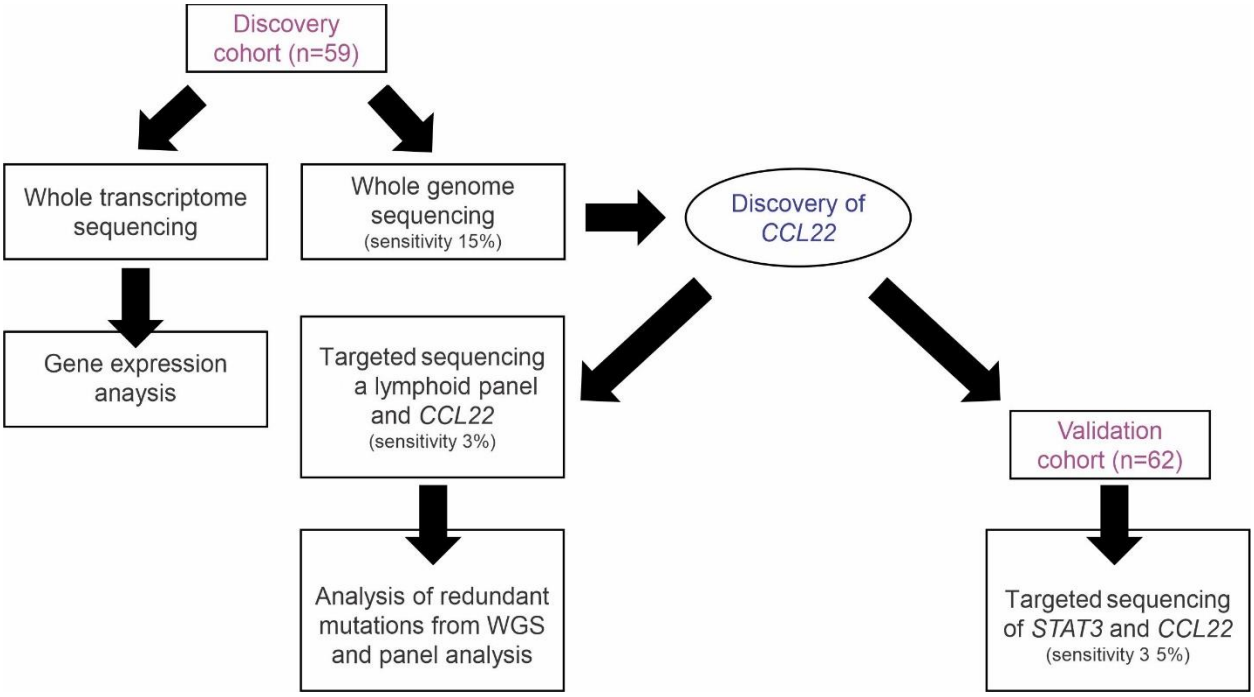
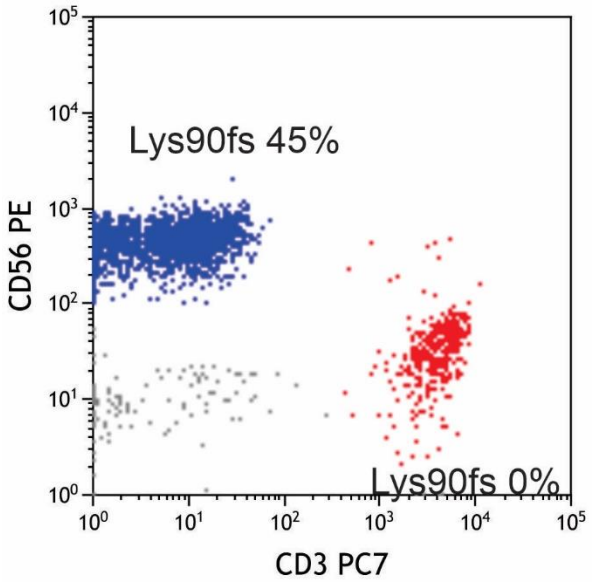
CCL22 mutations drive natural killer cell lymphoproliferative disease by deregulating microenvironmental crosstalk

Constance Baer*, Shunsuke Kimura*, Mitra S. Rana, Andrew B. Kleist, Tim Flerlage, David J. Feith, Peter Chockley, Wencke Walter, Manja Meggendorfer, Thomas L. Olson, HeeJin Cheon, Kristine C. Olson, Aakrosh Ratan, Martha-Lena Mueller, James M. Foran, Laura J. Janke, Chunxu Qu, Shaina N. Porter, Shondra M. Pruett-Miller, Ravi C. Kalathur, Claudia Haferlach, Wolfgang Kern, Elisabeth Paietta, Paul G. Thomas, M. Madan Babu, Thomas P. Loughran Jr., Ilaria Iacobucci, Torsten Haferlach and Charles G Mullighan

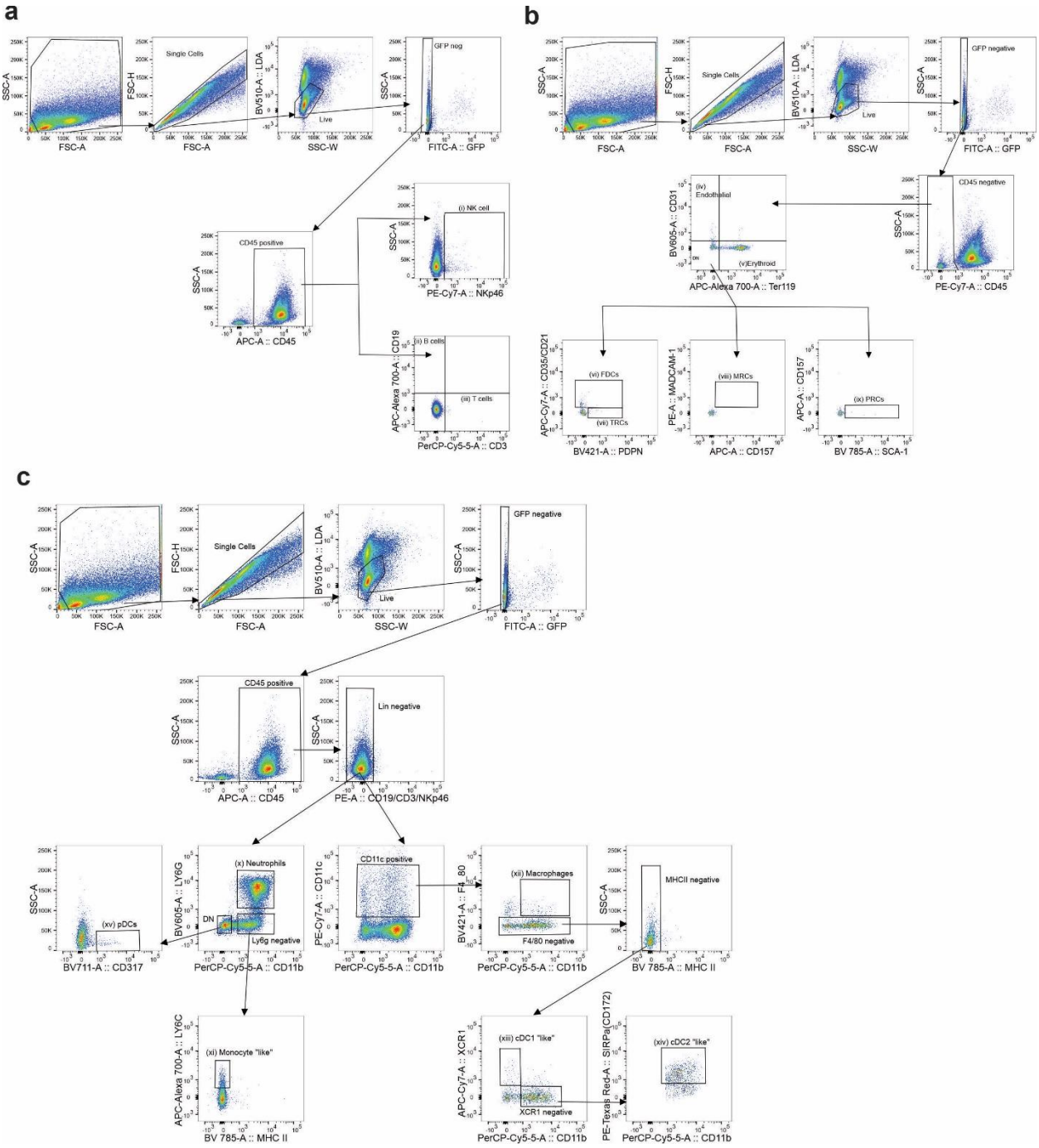
Supplementary Figures



Supplementary Figure 1. Schema of study design with integrated genomic analysis in the discovery (n = 59) and validation (n = 62) cohorts. Sensitivity is the lower limit of variant allele frequency to filter the candidate calls from whole genome sequencing and targeted sequencing.



Supplementary Figure 2. Flow-sorted CD56+ NK and CD3+ T cells. The *CCL22* mutation (Lys90fs) was detected only from the CD56+ NK cell population but not from CD3+ T cells. The same finding was observed for the other Pro79Arg *CCL22* mutation; the valiant allele frequency (VAF) was 49% in the CD56+ NK cell population and 0% in the CD3+ T cell population.



Supplementary Figure 3. Example of gating strategy described in Figure 5a,b. Three panels were used for detecting infiltrated mouse immune and stromal cell populations in spleens of engrafted NSG-Tg(Hu-IL15) mice transplanted with NK-92 cells expressing wild type (n=4, biologically independent animal samples in a single experiment) or mutant CCL22 (Pro79Arg, n=5), or GFP-expressing lentiviral vector (empty vector, n=7). **a,** The panel for (i) NK (NKp46+), (ii) B (CD19+/CD3-), and (iii) T cell (CD19-/CD3-) populations. **b,** The panel for (iv) endothelial (CD45-/CD31+/Ter119-), (v) erythroid cells (CD45-/CD31-/Ter119+), (vi) follicular dendritic

cells (FDC; CD45-/CD31-/Ter119-/CD35+ or CD21+), (vii) T cell zone reticular cell (TRC; CD45-/CD31-/Ter119-/CD35-/CD21-), (viii) marginal zone reticular cells (MRC; CD45-/CD31-/Ter119-/MADCAM-1+/CD157+), and (ix) perivascular reticular cells (PRC; CD45-/CD31-/Ter119-/CD157-/SCA-1+) populations. **c,** The panel for (x) neutrophils (Lin-/LY6G+), (xi) monocyte (Lin-/Ly6G-/Ly6C+/MHCII-), (xii) macrophage (Lin-/CD11c+/CD11b+/F4-80+), (xiii) conventional DC1 (cDC1; Lin-/CD11c+/F4-80-/MHCII-/XCR1+), cDC2 (Lin-/CD11c+/F4-80-/MHCII-/XCR1-/CD172+), and (xv) plasmacytoid DC (pDC; Lin-/LY6G-/CD11b-/CD317+) populations.

Supplementary note

Supplementary methods

Whole genome, whole transcriptome and targeted sequencing

For the CLPD-NK discovery cohort, we performed TruSeq DNA PCR-Free whole-genome sequencing library preparation according to manufacturer's instructions (llumina, ILMN, San Diego, CA) on the automated NGS Star liquid handling platform (Hamilton, Bonaduz, Switzerland) followed by 2x150 bp paired-end sequencing on the HiSeqX or NovaSeq6000 (ILMN). An average coverage of >100x was achieved.

For whole transcriptome analysis, the TruSeq Total Stranded RNA kit was used, starting with 250 ng of total RNA, to generate RNA libraries following the manufacturer's recommendations (ILMN). 2x100bp paired-end reads were sequenced on the NovaSeq 6000 with a median of 50 mio. reads per sample (ILMN).

Targeted sequencing of the CLPD-NK was performed for *STAT3* and *CCL22* cohort by PCR amplicons generation with Illumina adapters (primers: ggctgagacatacaggacaga-acttgtaaactgaggcccaga, atcctctggtcaccatccttc-aggagtctgaggtccagtaga, tgattacgtccgttaccgtct-tgaactgagccatctcttcca, aatgctaagctcccgagggtg-cagctataatggcagggaggt). Sequencing libraries were prepared using a custom rhAmpSeq design (IDT, Integrated DNA Technologies Inc. Coralville, IA) according to the manufacturer's recommended high throughput method, with 50 ng of DNA or PCR amplicons generation with Illumina adapters and sequenced on the MiSeq. The CLPD-NK discovery cohort was further analyzed with a lymphatic gene panel (*ARID1A, ATM, BCL2, BIRC3, BRAF, BTK, CARD11, CCND1, CD79A, CD79B, CREBBP, CXCR4, DDX3X, DIS3, DNMT3A, EP300, EZH2, FAM46C, FAS, FAT4, FBXW7, GPR98, ID3, IKBKB, IL2RG, JAK1, JAK3, KLHL6,*

KMT2D, KRAS, MAP2K1, MAPK1, MEF2B, MYBBP1A, MYD88, NFKBIE, NOTCH1, NOTCH2, NRAS, PHF6, POT1, RPS15, RUNX1, SF3B1, STAT3, STAT5B, TBL1XR1, TCF3, TET2, TLR2, TNFAIP3, TNFRSF14, TP53, TRAF3, UBR5, WHSC1, XPO1, ZMYM3). Library prep was done by Nextera DNA Flex Library Prep (ILMN) followed by enrichment with the ITD xGen system (IDT). A sensitivity limit of 3% was achieved by sequencing on the NovaSeq6000 (ILMN).

Lentiviral cloning and transduction

A gateway-compatible entry clone containing the *CCL22* and *CCR4* cDNA was obtained from Genecopoeia (CCL22, GC-Z9059; CCR4, GC-A0663). The *CCL22* mutations were generated by site directed mutagenesis using the Q5 Site-Directed Mutagenesis Kit (New England BioLabs, #E0554S) and primers (Supplementary table 14). Wild-type or mutated cDNAs were then transferred into a Gateway-compatible CL20-MSCV-IRES-GFP (for *CCL22*) or CL20-MSCV-IRES-ametrine (for *CCR4*) lentiviral vector using the LR Clonase II enzyme mix (Life Technologies, #11791100). Constructs were verified by Sanger sequencing. To produce lentivirus, HEK 293T (293T) cells were cotransfected with pEcopac helper packaging plasmid and CL20-MSCV-IRES-GFP/ametrine as an empty vector or containing the wild-type or mutant gene of interest by using FuGENE HD Transfection Reagent (Promega, E2311). Viral supernatants were harvested 48 hours and 72 hours post transfection and filtered through a 0.45 μm filter (Millipore, #SE1M003M00). Mouse hematopoietic Ba/F3 cells or human NK NK-92 cells were infected on RetroNectin (Takara, #T100B)-coated plates preloaded with viral supernatants and cultured in the presence of cytokines for 48 hours. Transduced GFP and/or ametrine-positive cells were collected by fluorescence-activated cell sorting.

Recombinant chemokine preparation, cAMP and β-arrestin assay

CCL22 wild-type and mutant chemokines, fused to thioredoxin with a hexa-histidine metal affinity tag and a human rhinovirus 3C protease site, were expressed in BL21 Star (DE3) cells (Invitrogen, #C601003). The insoluble chemokines were solubilized in 50mM TrisHCl, pH8.0, 6M GdnHCl buffer and refolded using a dialysis method with a redox cysteine:cystine mix at 10:1 ratio. Following refolding, the thioredoxin fusion along with the hexa-histidine tag was removed by treatment with PreScission protease. The chemokines were further purified by cation-exchange chromatography on a MonoS 5/50 GL (Cytiva, #17516901) column in an AKTA Pure FPLC system (Cytiva). Purity of the chemokines were assessed by SDS-PAGE. Quality of the chemokines were assessed by nanoDSF using Tycho NT.6 in the presence and absence of dithiothreitol (DTT) to ascertain disulfide bond formation. Purified chemokines were dialyzed into PBS buffer and frozen in aliquots at –80°C for later use.

Recombinant chemokine activities were measured by the GloSensor cAMP assay. 293T cells stably expressing CCR4 receptor were transfected with the pGloSensor 22F (Promega, #E2301) plasmid. A day later, cells were resuspended in HBSS containing 0.5mg/mL D-luciferin (GoldBio, #LUCK-100), plated on 96-well microplates, and treated with 10 µM Forskolin followed by serially diluted concentrations of the chemokines ranging from ~ 1.3 µM to ~0.03 pM. Luminescence from the plates were read on Spectramax iD5 (Molecular Devices) on a kinetic mode. The decrease in normalized luminescence signal was plotted as a function of the chemokine concentration and fit with an in-built function in OriginPro Graphing and Analysis software (Origin Lab) to extract the EC_{50} value. β-arrestin recruitment was measured by the PathHunter® eXpress CCR4 CHO-K1 β-Arrestin GPCR Assay (DiscoverX, 93-0193E2CP0M) according to the manufacturer's protocol. PathHunter® eXpress β-Arrestin GPCR-CCR4 cells were treated with

recombinant wild type and mutant CCL22 for three hours. Luminescence from the plates were read on Synergy HT (Bio Tek) and analyzed by Prism (GraphPad, version 8).

Immunofluorescence assay

293T cells expressing wild type CCR4 were seeded overnight on Millicell EZ slides (Millipore, #PEZGS0416) and transiently transfected with the vector containing wild type or mutant *CCL22* (Lys90Argfs, Pro79Arg, Leu45Arg) by using FuGENE HD (Promega, #E2311). After 48 hours incubation, cells were fixed for 5 min at room temperature with 4% paraformaldehyde, and permeabilized in 0.1%/PBS Triton-X 100 (Sigma-Aldrich, #X100) for 10 min (Permeabililized condition) or incubated with PBS for 10 min (Non-Permeabilized condition). Non-specific sites were blocked by incubating cells for 60 min in Donkey serum (Sigma-Aldrich, #D9663) 1x/PBS. Cells were stained with Alexa Fluor 647-conjugated anti-human CD194 (CCR4) antibody (BD, #561992) for 1 h at 37°C. Slides were washed three times with PBS and incubated with ProLong Diamond Antifade Mountant with DAPI (Thermo Fisher Scientific, #P36962) for nucleic acid staining and mounting. Fluorescent images were acquired on a Lecia SP8 AOBS (Leica Microsystems) confocal microscope using a 63x/1.4 NA objective with a 4x zoom. Fluorescent probes were excited with either a 405 (DAPI) laser or a white light laser (Leica Microsystems) tuned to 653 nm (Alexa 647). Emission was collected on a HyD detector and pseudo-colored to either blue or red, respectively. Post-acquisition, image deconvolution was performed using Leica Lightning software with adaptive settings. Images were exported to tiffs using the LASX software (Leica Microsystems).

Gene editing

NKL cells were modified by CRISPR-Cas9 genome editing to correct the mutation and express wild type *CCL22*. Briefly, approximately 400,000 NKL cells were subjected to a transient

transfection via nucleofection (Lonza, 4D-Nucleofector™ X-unit) with the program CA-137 in a small (20ul) cuvette using solution P3 according to the manufacturer's recommended protocol and with precomplexed ribonuclear proteins (RNPs). RNPs included the following components: 100 pmol of chemically modified sgRNA (Synthego) and 33 pmol of Cas9 protein produced at the St. Jude Protein Production Core. Additionally, 3ug of ssODN donor (IDT; AltR™ modifications); and 200ng of pMaxGFP (Lonza) were included in the nucleofection. Five days post-nucleofection, NKL transfected cells (GFP+) were single cell sorted by FACS in 96-well plates containing filtered 50:50 conditioned media (mixture of 50% fresh, complete media and 50% media harvested from cultured cells) and clonally expanded. The clones harboring the desired genomic modification were identified by targeted deep sequencing. Briefly, DNA was extracted from clonal cell pellets and amplified with primers overlapping the region containing the modification and containing partial Illumina adapter overhangs[1]. The generated amplicons were indexed and pooled with other amplicons for different genomic loci to generate sequence diversity. Additionally, 10% PhiX Sequencing Control V3 (Illumina) was added. Paired 2 X 250bp reads were generated on an Miseq Sequencer System (Illumina). Samples were demultiplexed by the index sequences, fastq files were generated, and sequencing analysis was performed using CRIS.py.[2] The authentication of the final clones was performed by the PowerPlex® Fusion System (Promega) at the Hartwell Center (St. Jude Children's Research Hospital). Clones were tested for mycoplasma by the MycoAlertTMPlus Mycoplasma Detection Kit (Lonza) and resulted negative. Supplementary table 14 lists the editing construct sequences and screening primers that were used.

Cytokine array

NKL (CCL22-Arg44_Leu45insSer) and gene edited NKL (CCL22 wild type) cell lines at a concentration of 2.5 x 10^5 cells/mL were seeded in 24 well plate and incubated in a medium

containing Minimum Essential Medium Eagle - alpha modification supplemented with 12.5% heat-inactivated FBS, 12.5% heat-inactivated horse serum, and 100 U/mL penicillin-streptomycin in the presence of 500 U/ml recombinant human IL-2 at 37°C for 72 hours. Supernatants were subsequently examined for cytokine profiling with Proteome Profiler Human XL Cytokine Array Kit (R&D, #ARY022B) according to the manufacture's instruction. Odyssey® CLx Imaging System (LI-COR) and IRDye 800CW Streptavidin (LI-COR, Catalog #926-32230) were used for the imaging.

Supplementary table information

Supplementary Table 1.

Patients' data in discovery and validation cohort.

Supplementary Table 2.

Summary of clinical features of CLPD-NK cases with hotspot *CCL22* mutations (at residues Leu45, Pro79).

Supplementary Table 3.

Z-score of 59 cases with CLPD-NK and 12 normal NK cells from peripheral blood and bone marrow (GSE133383). This tables describes normalized gene expression of each sample.

Supplementary Table 4.

Differentially expressed genes analyzed by DESeq2 between *CCL22* mutations vs others, *STAT3* mutations vs others, and non-*CCL22/STAT3* mutations (Others) vs others.

Supplementary Table 5.

Differentially expressed genes analyzed by DESeq2 between sorted BM/PB $CD56^{bright}$ and $CD56^{dim}$ normal NK cells from the public database (GSE133383).

Supplementary Table 6.

Pathway analysis using differentially expressed genes in CLPD-NK cases with *CCL22* mutations, *STAT3* mutations, and Other (non-*CCL22/STAT3* mutations), compared with normal NK cell.

Supplementary Table 7.

Gene set enrichment analysis of 59 cases of primary CLPD-NK patients with *CCL22* mutations v.s. others (c2all).

Supplementary Table 8.

The results of intracellular phosphorylation of pERK in BaF3-CCR4 cells after treatment of CCL22 wild type/mutant proteins detected by isoplexis.

Supplementary Table 9.

Differentially expressed genes of engrafted NK-92 cells in NSG-Tg(Hu-IL15) mice analyzed by DESeq2 between CCL22 mutant v.s. wild type/empty vector transduced NK-92 cells.

Supplementary Table 10.

Gene set enrichment analysis of engrafted NK-92 cells in NSG-Tg(Hu-IL15) mice between mutant v.s. wild type/empty vector transduced NK-92 cells..

Supplementary Table 11.

Gene sets generated by differentially expressed genes for GSEA.

Supplementary Table 12.

CCL22-dependent chemokine secretion. This tables describes the results from chemokine array performed to compare chemokine secretion between gene edited *CCL22* wild type and mutant NKL cell line.

Supplementary Table 13.

Antibody list used in the manuscript.

Supplementary Table 14.

Primer sequences for cloning and gene editing.

Supplementary references

1 Sentmanat, M. F., Peters, S. T., Florian, C. P., Connelly, J. P. & Pruett-Miller, S. M. A Survey of Validation Strategies for CRISPR-Cas9 Editing. Sci Rep 8, 888 (2018).

2 Connelly, J. P. & Pruett-Miller, S. M. CRIS.py: A Versatile and High-throughput Analysis Program for CRISPR-based Genome Editing. Sci Rep 9, 4194 (2019).