Induced pluripotent stem cells from CINCA syndrome patients as a model for dissecting somatic mosaicism and drug discovery

Takayuki Tanaka,[1] Kazutoshi Takahashi,[1] Mayu Yamane,[1] Shota Tomida,[1] Saori Nakamura,[1] Koichi Oshima,[1] Akira Niwa,[1] Ryuta Nishikomori,[2] Naotomo Kambe,[3] Hideki Hara,[4] Masao Mitsuyama,[4] Nobuhiro Morone,[5] John E. Heuser,[5] Takuya Yamamoto,[1] Akira Watanabe,[1] Aiko Sato-Otsubo,[6] Seishi Ogawa,[6] Isao Asaka,[1] Toshio Heike,[2] Shinya Yamanaka,[1,5,7,8] Tatsutoshi Nakahata,[1,2] and Megumu K. Saito[1]

[1]Center for iPS Cell Research and Application and [2]Department of Pediatrics, Kyoto University, Kyoto, Japan; [3]Department of Dermatology, Chiba University Graduate School of Medicine, Chiba, Japan; [4]Department of Microbiology and [5]Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto, Japan; [6]Cancer Genomics Project, University of Tokyo, Tokyo, Japan; [7]Yamanaka iPS Cell Special Project, Japan Science and Technology Agency, Kawaguchi, Japan; and [8]Gladstone Institute of Cardiovascular Disease, San Francisco, CA

Chronic infantile neurologic cutaneous and articular (CINCA) syndrome is an IL-1–driven autoinflammatory disorder caused mainly by *NLRP3* mutations. The pathogenesis of CINCA syndrome patients who carry *NLRP3* mutations as somatic mosaicism has not been precisely described because of the difficulty in separating individual cells based on the presence or absence of the mutation. Here we report the generation of *NLRP3*-mutant and nonmutant-induced pluripotent stem cell (iPSC) lines from 2 CINCA syndrome patients with somatic mosaicism, and describe their differentiation into macrophages (iPS-MPs). We found that mutant cells are predominantly responsible for the pathogenesis in these mosaic patients because only mutant iPS-MPs showed the disease relevant phenotype of abnormal IL-1β secretion. We also confirmed that the existing anti-inflammatory compounds inhibited the abnormal IL-1β secretion, indicating that mutant iPS-MPs are applicable for drug screening for CINCA syndrome and other NLRP3-related inflammatory conditions. Our results illustrate that patient-derived iPSCs are useful for dissecting somatic mosaicism and that *NLRP3*-mutant iPSCs can provide a valuable platform for drug discovery for multiple NLRP3-related disorders. (*Blood.* 2012;120(6):1299-1308)

Introduction

Chronic infantile neurologic cutaneous and articular syndrome (CINCA syndrome; MIM #607715) is a dominantly inherited autoinflammatory disease characterized by systemic inflammation with an urticaria-like rash, neurologic manifestations, and arthropathy.[1] *NLRP3* mutation is the first and so far the only identified mutation that is responsible for CINCA syndrome.[2,3] *NLRP3* is expressed mainly in myelomonocytic lineage cells and chondrocytes[3] and acts as an intracellular sensor of danger signals from various cellular insults. In normal macrophages, a first stimulus, such as lipopolysaccharide (LPS), induces the synthesis of NLRP3 and the biologically inactive proIL-1β.[4] A second stimulus, such as ATP, enhances the assembly of a protein complex called the NLRP3-inflammasome.[5] The inflammasome contains caspase1, which executes the proteolytic maturation and secretion of IL-1β. Although normal monocytes/macrophages show no or limited IL-1β secretion in response to LPS stimulation alone, CINCA patients' cells exhibit robust IL-1β secretion because the mutant NLRP3-inflammasome is autoactivated without the need for any second stimulus.[6] It is therefore thought that the manifestations of CINCA syndrome are predominantly caused by the excessive secretion of the proinflammatory cytokine, IL-1β, and this concept is supported by the efficacy of an IL-1 receptor antagonist (IL-1Ra) for decreasing most of the symptoms.[7] However, because IL-1Ra treatment does not seem to ameliorate the characteristic arthropathy of cartilage overgrowth and joint contraction,[8] a more specific therapeutic approach that directly modulates the NLRP3-inflammasome is desired.

Although approximately half of CINCA patients carry heterozygous gain-of-function mutations of the *NLRP3* gene,[2,3] 30% to 40% of all patients have mutations in *NLRP3* in only a small number of somatic cells.[9,10] Because the population of mutant cells is relatively small (4.2%-35.8% in blood cells), it remains controversial whether the small fraction of *NLRP3*-mutated cells actually causes the strong autoinflammation observed in CINCA patients, or whether the *NLRP3* mutations found in mosaic patients are just a bystander, with all cells carrying an unknown mutation of another gene that causes the disease.[11]

Somatic mosaicism refers to the presence of more than 1 genetically distinct cell population in a single person, and has been identified in patients with various diseases.[12,13] The relevance of somatic mosaicism to the onset of diseases has been suggested mainly through sequence-based approaches. However, direct evidence that a cell population with a distinct genetic property shows disease-specific characteristics is lacking because it has been impossible to separately extract individual live cells from affected tissues to assess their biologic characteristics. Regarding hematopoietic disorders in which mutant cells show decreased expression of a certain protein, genetic heterogeneity caused by somatic mutations was detected by flow cytometry after intracellular staining,[14-16] but sorting out alive mutant and nonmutant cells for evaluating biologic property has been impossible.

Submitted March 27, 2012; accepted May 29, 2012. Prepublished online as *Blood* First Edition paper, June 21, 2012; DOI 10.1182/blood-2012-03-417881.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Induced pluripotent stem cells (iPSCs) are pluripotent cell lines directly reprogrammed from somatic cells.[17] Patient-derived iPSCs can provide somatic cells, which cannot be directly obtained from patients, and this discovery has led to the development of a new field of disease modeling (reviewed by Grskovic et al[18]). In addition, iPSC technology has another interesting characteristic that each iPSC clone originates from a single cell,[19] which may make it possible to obtain genetically different iPSC clones from a person.

In this study, we established mutant and nonmutant iPSC lines from the same patients by deriving iPSCs from patients carrying a mutation of an autosomal gene as somatic mosaicism. By analyzing the disease-relevant characteristic of IL-1β secretion, we demonstrated that mutant macrophages are mainly responsible for the disease phenotype in the mosaic patients. Moreover, using a robust differentiation protocol to generate macrophages and purifying them by their surface marker expression, we showed that drug candidates inhibit the IL-1β secretion from mutant macrophages. Our data prove the usefulness of iPSC technology both for dissecting somatic mosaicism and as a platform for drug discovery of multiple NLRP3-related inflammatory diseases.

Methods

Human iPSC generation

We obtained skin biopsy specimens from 2 independent patients (patient 1, CIRA188Ai; and patient 2, CIRA086Ai). This study was approved by Ethics Committee of Kyoto University, and informed consent was obtained from both the patients and their guardians in accordance with the Declaration of Helsinki. We expanded the fibroblasts in DMEM (Nacalai Tesque) containing 10% FBS (Invitrogen) and 0.5% penicillin and streptomycin (Invitrogen). Generation of iPS cells was performed as described previously.[17] In brief, we introduced *OCT3/4*, *SOX2*, *KLF4*, and *c-MYC* using ecotropic retroviral transduction into fibroblasts expressing the mouse *Slc7a1* gene. Six days after transduction, the cells were harvested and replated onto mitotically inactivated SNL feeder cells. The next day, we replaced the medium with Primate ES cell medium (ReproCELL) supplemented with 4 ng/mL bFGF (Wako). Three weeks after this period, individual colonies were isolated and expanded. Cell culture was performed under 37°C, with 5% CO_2 and 21% O_2 unless otherwise stated. Cells were examined using Olympus CKX41 inverted microscope equipped with Nikon Digital Sight DS-L2 camera. A UPlan FLN 4×/0.13 objective (Nikon) was used for image acquisition.

Genetic analysis

Genomic DNA from either fibroblasts or iPSCs was isolated. The PCR product of exon 3 of *NLRP3* was sequenced directly or after subcloning with a TOPO TA cloning kit (Invitrogen), using an ABI 3100 sequencer (Applied Biosystems). For pyrosequencing, the PCR product of exon 3 of *NLRP3* was analyzed by PyroMarkQ96ID (QIAGEN).

RNA isolation and quantitative PCR for *NANOG* and the transgene

Total RNA was purified with the Trizol reagent (Invitrogen) and treated with a Turbo DNA-free kit (Ambion) to remove genomic DNA contamination. A total of 1 μg of total RNA was used for a reverse transcription reaction with ReverTraAce-α (Toyobo) and the dT_{20} primer, according to the manufacturer's instructions. Quantitative PCR was performed on the 7900HT Fast Real-Time PCR System (Applied Biosystems) with SYBR Premix ExTaqII (Takara). The primer sequences are described in supplemental Table 4 (available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).

Southern blotting

Genomic DNA (5 μg) was digested with BglII and ScaI overnight. The digested DNA fragments were separated on 1% agarose gels and were transferred to a nylon membrane (GE Healthcare). The membrane was incubated with a digoxigenin (DIG)–labeled human *cMYC* DNA probe in DIG Easy Hyb buffer (Roche Diagnostics) at 42°C overnight with constant agitation. After washing, an alkaline phosphatase-conjugated anti-DIG antibody (1:10 000; Roche Diagnostics) was added to a membrane. Signals were obtained using CDP-star (Roche Diagnostics) and detected by an LAS4000 imaging system.

Teratoma formation

Approximately 2×10^6 cells were injected subcutaneously into the dorsal flank of immunocompromised NOD/scid/γc^{null} mice (Central Institute for Experimental Animals). Masses were excised 8 to 10 weeks after injection and fixed with PBS containing 4% paraformaldehyde. Paraffin-embedded tissues were sliced and stained with hematoxylin and eosin. Slides were examined using BIOREVO BZ-9000 (KEYENCE). A PlanApo 20×/0.75 objective (Nikon) and BZ-II Viewer software (KEYENCE) were used for image acquisition.

In vitro differentiation into macrophages

Undifferentiated human embryonic stem cell (ESC) and iPSC lines were cultured on mitotically inactivated SNL feeder cells with Primate ES cell medium supplemented with 4 ng/mL bFGF. During the differentiation of the cells into macrophages, cells were cultured under 37°C, with 5% CO_2 and 5% O_2. On day 0, the iPSCs were plated at a ratio of 1:15 onto a mitotically inactivated OP9 feeder layer on 100-mm cell culture plates in α-MEM (Invitrogen) containing 10% FBS and 1% Antibiotic-Antimycotic (Invitrogen) supplemented with 50 ng/mL VEGFα (R&D Systems). On day 5, the medium was changed. On day 10, the differentiating iPSCs were collected by trypsinization, and Tra-1-85$^+$ CD34$^+$ and KDR$^+$ hematopoietic progenitors were sorted on a FACSAria II instrument (BD Biosciences). The progenitors were plated at 2×10^4 cells on another mitotically inactivated OP9 feeder layer on 100-mm cell culture plates or at 3×10^3 cells/well in 6-well cell culture plates in α-MEM containing 10% FBS and 1% Antibiotic-Antimycotic supplemented with 50 ng/mL IL-3, 50 ng/mL stem cell factor, 10 ng/mL thrombopoietin, 50 ng/mL Flt-3 ligand, and 50 ng/mL M-CSF (all R&D Systems). On day 18, the medium was changed. On day 26, differentiating cells were collected with Accumax (Innovative Cell Technologies), and CD14$^+$ iPSC-derived macrophages were purified on an autoMACSpro instrument (Miltenyi Biotec).

Peripheral blood mononuclear cells (PBs) were obtained from healthy volunteers, and CD14$^+$ monocytes were purified on the autoMACSpro instrument. For macrophage differentiation, 5×10^5 monocytes were plated in 6-well cell culture plates in RPMI 1640 (Sigma-Aldrich) containing 10% FBS and 1% Antibiotic-Antimycotic supplemented with 50 ng/mL M-CSF. On day 5, the adherent cells were collected with Accumax, and CD14$^+$ blood-derived macrophages (B-MPs) were purified on the autoMACSpro instrument. May-Giemsa–stained slides were examined using BIOREVO BZ-9000. A PlanApo 40×/0.95 objective (Nikon) and BZ-II Viewer software were used for image acquisition.

FACS analysis

Hematopoietic marker expression was evaluated on a MACSQuant Analyzer (Miltenyi Biotec). Primary antibodies Tra-1-85-FITC (R&D Systems), CD34-PE (Beckman Coulter), KDR–AlexaFluor-647 (BioLegend), CD45-PE (BD Biosciences PharMingen), and CD14-APC (Beckman Coulter) were used.

Immunocytochemistry

For immunocytochemistry, cells were fixed with PBS containing 4% paraformaldehyde for 5 minutes, permeabilized in PBS containing 0.1% Tween 20 for 5 minutes, and blocked in PBS containing 3% BSA for 10 minutes, all at room temperature. The primary antibody was for CD68 (1:50; Santa Cruz Biotechnology), and the secondary antibody was Cy3-conjugated

AffiniPure Donkey Anti–Mouse IgG (1:100; Jackson ImmunoResearch Laboratories). Nuclei were stained with 1 μg/mL Hoechst 33342 (Invitrogen). Cells were examined using BIOREVO BZ-9000. A Plan Fluor DL 10×/0.30 Ph1 objective (Nikon) and BZ-II Viewer software were used for image acquisition.

Electron microscopy

The 5×10^4 macrophages in 20 μL suspension were placed on the poly-L-lysine treated, carbon-coated sapphire disks (3 mm in diameter) and incubated for 30 minutes at 37°C with 5% CO_2. The cell-adsorbed disk was then subjected to chemical fixation with 2.5% glutaraldehyde in NaHCa buffer (100mM NaCl, 30mM HEPES, 2mM $CaCl_2$, adjusted at pH 7.4 with NaOH). These specimens were postfixed with 1% osmium and 1.5% $K_4Fe(CN)_6$ in 0.1M PBS buffer, washed, dehydrated with a series of ethanol, and embedded in Epoxy resin (TAAB EPON812). After the polymerization at 70°C, the ultra-sections (70 nm) obtained by Ultramicrotome (Leica FC6) were mounted in EM grids, stained with uranyl acetate/lead citrate, and then observed by conventional TEM (JEOL JEM1400).

PCR and microarray analysis of macrophages

Total RNA was column-purified with the RNeasy kit (QIAGEN) and treated with RNase-free DNase (QIAGEN). A total of 20 ng of total RNA was reverse transcribed into cDNA using random primers and the Sensiscript RT Kit (QIAGEN). Quantitative PCR was performed on a StepOne Plus Real-Time PCR System (Applied Biosystems) with TaqMan Gene Expression Master Mix (Applied Biosystems). The primer sequences are described in supplemental Table 4. For the microarray analysis, RNA probes were hybridized to SurePrint G3 Human GE 8 × 60K Microarrays (Agilent Technologies) according to the manufacturer's protocols. Microarrays were scanned, and the data were analyzed using the GeneSpring GX Version 11 software program (Agilent Technologies). The complete dataset from this analysis is available at the NCBI Gene Expression Omnibus using accession no. GSE38626.

LM infection

Listeria monocytogenes EGD (LM) were grown in brain heart infusion broth (Eieken Chemical), washed, suspended in PBS supplemented with 10% glycerol, and stored in aliquots at −80°C. Macrophages were seeded into an 8-well chamber slide at 2×10^5 cells/well in RPMI containing 10% FBS and then infected with bacteria at a multiplicity of infection of 10 for 60 minutes at 37°C. Cells were cultured for further 1 or 5 hours in the presence of 5 μg/mL gentamicin. The cells were fixed in 4% paraformaldehyde and incubated with PBS containing 10% Blocking One (Nacalai Tesque) and 0.1% saponin. F-actin and nuclei were visualized by staining with Alexa-488-phalloidin (Invitrogen) and 4′,6-diamidino-2-phenylindole (Dojindo), respectively. The bacteria were stained by treatment with a goat anti-*Listeria* polyclonal antibody (Kirkegaard & Perry Laboratories) and then with the Alexa 546 anti–goat IgG antibody (Invitrogen). Slides were examined using BIOREVO BZ-9000. A PlanApo_VC 100×H/1.40 objective (Nikon) and BZ-II Viewer software were used for image acquisition, and BZ-II Analyzer (KEYENCE) was used for image processing. Immunofluorescence was evaluated with the IN Cell Analyzer 2000, and samples were analyzed with the IN Cell Developer Toolbox Version 1.8 software program (GE Healthcare).

Cytokine secretion from macrophages

Purified iPS-MPs or B-MPs were seeded at the indicated counts per well or 5×10^4 cells/well unless otherwise stated in 96-well cell culture plates in RPMI 1640 containing 10% FBS and 1% Antibiotic-Antimycotic. Cells were cultured for 2 hours in the presence or absence of inhibitors. The plates were centrifuged at 300*g* for 10 minutes; then the medium was changed. Cells were cultured for 4 hours in the presence of LPS or recombinant human IL-1β. LPS concentration was 1 μg/mL unless otherwise stated. After the 30 minute or 1-hour culture after the addition of 1mM ATP (Sigma-Aldrich), we collected the supernatants and cell lysates. As second signal stimulants, we also used 500 μg/mL silica crystals (U.S. silica) for 1 hour, or 100 μg/mL monosodium urate crystals (Sigma-Aldrich) for 3 hours. For the supernatant transfer experiments, we harvested the supernatant from the wells of mutant or wild-type iPS-MPs, which were stimulated with LPS for 4 hours. After centrifugation, we transferred the supernatants to the wells of other iPS-MPs and cultured them for another 4 hours. The cytokine concentration of the supernatants was determined using a Th1/Th2 11plex FlowCytomix Kit (Bender MedSystems) following the manufacturer's instructions. Reagents were purchased as follows: CA074Me (Calbiochem), IL-1Ra (R&D Systems), oxidized ATP (oATP; Sigma-Aldrich), pyridoxalphosphate-6-azophenyl-2′,4′-disulphonic acid (PPADS; Sigma-Aldrich), cycloheximide (Sigma-Aldrich), MG132 (Calbiochem), Bay11-7082 (Sigma-Aldrich), and Ac-YVAD-CHO (Calbiochem).

LDH secretion assay

The lactate dehydrogenase (LDH) concentration of the supernatants of iPS-MPs after a 4-hour culture with LPS was determined with an LDH Cytotoxicity Detection kit (Takara) following the manufacturer's instructions.

Statistical analysis

The data were processed using the SPSS Statistics Version 18 software package. The values are reported as the mean ± SEM. Comparisons between groups were performed using the unpaired Student *t* test. $P < .05$ was considered statistically significant.

Results

Establishment and characterization of iPSCs

Dermal fibroblasts were obtained from 2 male CINCA patients who had mutations of *NLRP3* as somatic mosaicism. Both patients had nonsynonymous point mutations in the *NLRP3* coding region. The fibroblasts from patients 1 and 2 contained 34% and 9.8% mutant cells, respectively (Figure 1A; supplemental Figure 1A). These fibroblasts were reprogrammed to iPSCs after transduction with retroviral vectors encoding *OCT3/4*, *SOX2*, *KLF4*, and cMYC.[17] Twelve of the 28 isolated clones from patient 1, and 3 of 30 clones from patient 2 had a heterozygous mutation of the *NLRP3* gene, whereas the rest of the clones were wild-type (Figure 1A; supplemental Figure 1B-C). The frequency of mutants was comparable among blood cells,[9,20] fibroblasts, and iPSCs (Table 1). We randomly selected 3 mutant (M1-M3) and 3 wild-type clones (W1-W3) from patient 1 and 3 mutant (m1-m3) and 3 wild-type clones (w1-w3) from patient 2 for the propagation and subsequent analyses.

All iPSC clones showed a characteristic human ESC-like morphology (Figure 1B), the reactivation of endogenous pluripotency genes (*OCT3/4*, *SOX2*, *NANOG*; Figure 1C-D; supplemental Figure 1D) and the demethylation of the *OCT3/4* promoter regions (supplemental Figure 1E). Transgene expression was rarely detected (Figure 1D; supplemental Figure 1D), and the retroviral integration patterns were confirmed by a Southern blot analysis (Figure 1E; supplemental Figure 1F). All of the iPSC clones maintained a normal karyotype (data not shown). There were neither proviral integration nor copy number changes observed in any of the genes that might affect the function of the NLRP3 inflammasome (supplemental Tables 1 and 2). Genetic identity was proven by a short tandem repeat analysis (supplemental Table 3), and the pluripotency of the iPSC clones was confirmed by the presence of cell derivatives of all 3 germ layers by teratoma formation after injection of undifferentiated iPSCs into immunocompromised NOD/scid/γc^{null} mice (Figure 1F; supplemental Figure 1G).

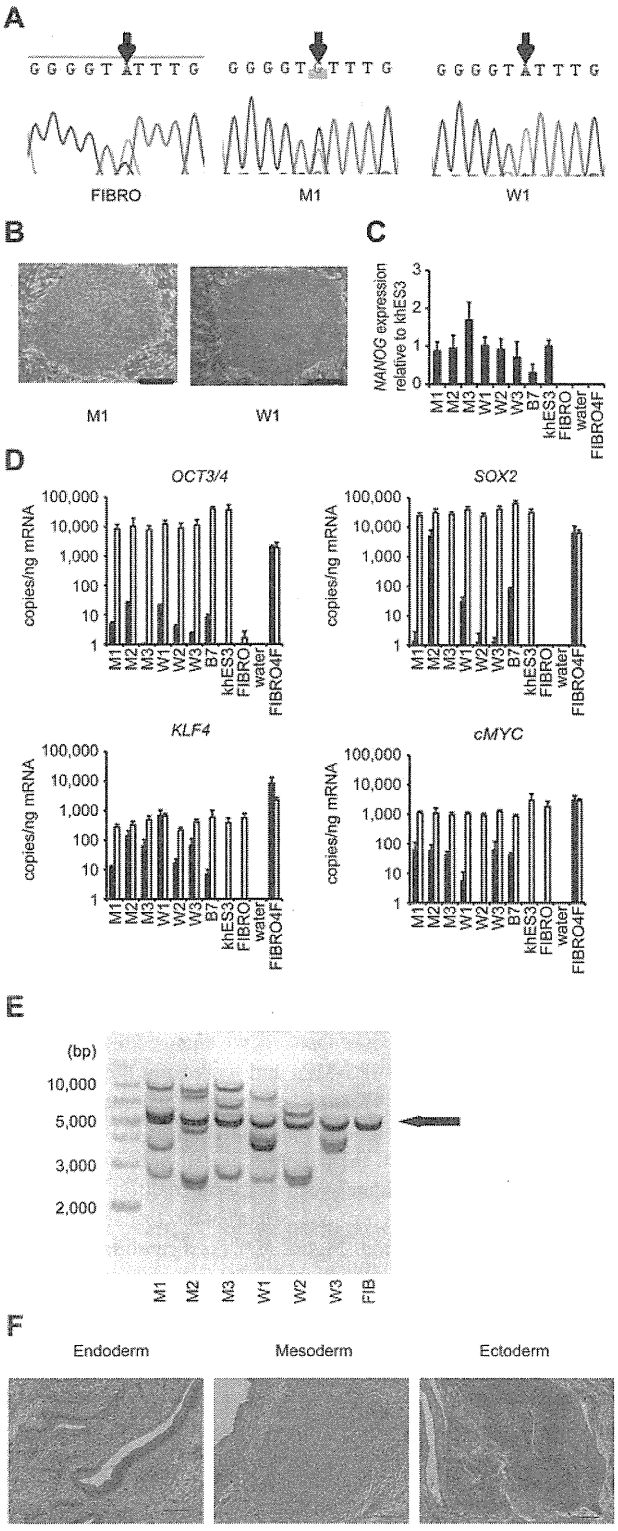

Figure 1. Establishment and characterization of iPSCs. (A) Sequencing of the *NLRP3* 1709 A > G mutation (Y570C) in fibroblasts (FIBRO), mutant iPSCs (M1), and wild-type iPSCs (W1) in patient 1. (B) The morphology of the mutant and wild-type iPSCs. (C) *NANOG* expression in CINCA iPSCs, control iPSCs (B7), control ESCs (khES3), fibroblasts (FIBRO), and fibroblasts transduced with 4 factors (FIBRO4F) normalized to *GAPDH*. n = 3. (D) A quantitative RT-PCR assay for the expression of *OCT3/4*, *SOX2*, *KLF4*, and *cMYC* in iPSCs. One primer set detects only the transgene (in black), and the other primer set detects both the transgene and endogenous gene (in white). n = 3. (E) Retroviral transgene integration analyses. Southern blot analyses were performed with DIG-labeled DNA probes against *c-MYC*. The parental fibroblasts carried a band in common with all of the iPSC lines (arrow). (F) A teratoma derived from a mutant iPSC clone, M1. Scale bars represent 100 μm. Data are mean ± SEM.

Differentiation and characterization of iPSC-derived macrophages

To compare the most prominent features of the disease, we differentiated the patient-derived iPSCs into the monocyte/macrophage lineage using a murine stromal cell line, OP9.[21] After culturing the iPSCs on an OP9 feeder layer for 10 days, we collected KDR$^+$ CD34$^+$ hemangioblasts (Figure 2A). All of the iPSC clones, whether they carried an *NLRP3* mutation or not, differentiated into KDR$^+$ CD34$^+$ progenitors as efficiently as the control ESC or iPSC clones (Figure 2B; supplemental Figure 2A). Adherent CD68$^+$ macrophages emerged after culturing the KDR$^+$ CD34$^+$ cells on another OP9 feeder layer for 16 days (Figure 2C; supplemental Figure 2B). Approximately 80% of the differentiated cells expressed CD14, and magnetic-activated cell sorting increased the purity to almost 100% (Figure 2D). All of the clones we used efficiently produced comparable amounts of iPSC-derived macrophages (iPS-MPs; Figure 2E; supplemental Figure 2C). The iPS-MPs visualized by light and electron microscopy showed a typical morphology, with a high cytoplasm-to-nucleus ratio and cytoplasmic vacuoles (Figure 2F; supplemental Figure 2D). The iPS-MPs showed a global gene expression pattern closer to that of blood-derived macrophages than to the parental iPSC clone (supplemental Figure 2E-F). Both mutant and wild-type iPS-MPs phagocytosed bacteria to the same extent when we infected the cells with Gram-positive LM, an intracellular bacterium that escapes into the cytosol (Figure 2G-H). These data indicate that both the mutant and wild-type iPS-MPs derived from mosaic CINCA patients are indistinguishable based on their gene expression and their phagocytic function.

Elucidation of the pathogenesis of somatic mosaic CINCA syndrome

Monocytes derived from CINCA syndrome patients usually do not spontaneously secrete IL-1β and become active after LPS stimulation.[6] Monocytes or mononuclear cells from untreated CINCA syndrome patients, however, sometimes show an increased synthesis of proIL-1β[2] and secretion of mature IL-1β,[7] even in the absence of LPS stimulation, because they can be activated by persistent inflammation or by the purification procedure. As spontaneous activation complicates the functional analysis, we herein evaluated the IL-1β activation status both before and after the stimulation. We observed that the mRNA expression of *IL1B* was low in unstimulated iPS-MPs and increased to comparable levels in mutant and wild-type iPS-MPs in response to LPS stimulation (supplemental Figure 3A). Similarly, the mRNA level of *NLRP3* was relatively low before LPS stimulation (supplemental Figure 3A). Mature IL-1β was not detectable in the supernatant of the cell culture medium (data not shown). Collectively, these data indicate that the unstimulated iPS-MPs were in an "inactive" state before stimulation.

To identify which iPS-MP clones showed the specific features compatible to patients' monocytes, we evaluated their IL-1β secretion. Although LPS stimulation alone led to IL-1β secretion

Table 1. Mutation frequency among different cell types

Patient no.	Site of mutation	Frequency (%) of mutant cells: Whole blood*	Fibroblasts	iPSCs
1	1709A > G(Y570C)	33.3	34.3	42.9
2	919G > A(G307S)	8.5	9.8	10.0

*The frequency in whole blood was reported previously.[9,20]

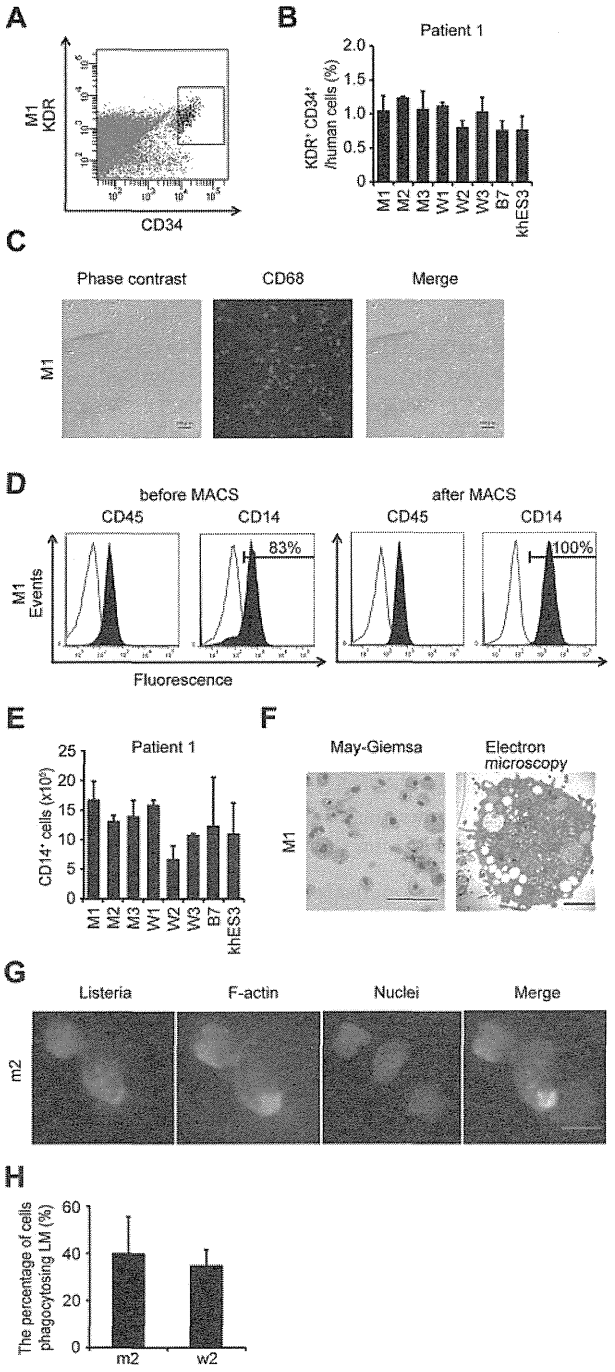

Figure 2. Differentiation and characterization of iPSCs-derived macrophages. (A) KDR$^+$ CD34$^+$ hematopoietic progenitors purified 10 days after differentiation. (B) The percentage of KDR$^+$ CD34$^+$ cells in Tra-1-85$^+$ human cells. n = 3. (C) CD68 immunostaining of macrophages. Scale bars represent 100 μm. (D) The histograms show antibody staining (in black) relative to the isotype-matched controls (in white) for a blood cell marker (CD45), and a macrophage marker (CD14), in cells before (left 2 panels) or after (right 2 panels) magnetic-activated cell sorting purification. (E) CD14$^+$ cell counts obtained from iPSCs plated on an OP9 feeder layer on one 100-mm dish. n = 3. (F) Representative morphology of iPS-MPs evaluated by May-Giemsa staining or transmission electron microscopy. Scale bars represent 100 μm and 2 μm, respectively. (G) The phagocytosis by iPS-MPs after LM infection. The cells were treated with anti-LM antibody, phalloidin, and 4′,6-diamidino-2-phenylindole. Scale bar represents 20 μm. (H) The percentage of iPS-MPs phagocytosing LM was calculated as the average of 9 fields of vision. Data are mean ± SEM.

from the mutant iPS-MPs, the addition of ATP was necessary to induce IL-1β secretion from wild-type iPS-MPs, as it was from either ESC-derived or blood-derived macrophages (Figure 3A). The IL-1β level from mutant iPS-MPs was significantly higher than that from wild-type macrophages, even in the presence of LPS plus ATP. Both groups of macrophages showed similar kinetics in their secretion of other cytokines, such as IL-6 or TNFα (Figure 3A). The results were similar in the iPS-MPs from patient 2 (Figure 3B). Although iPS-MPs showed a similar response at lower LPS concentrations (Figure 3C-D; supplemental Figure 3B-C), no IL-1β secretion was detectable from mutant iPSCs, wild-type iPSCs, or parental fibroblasts in response to stimulation with 1 μg/mL LPS (data not shown). These data demonstrate that the abnormal function of the iPS-MPs is predominantly determined by the *NLRP3* mutation, and not by some unknown genetic alteration(s) prevalent in all cells. We next investigated whether iPS-MPs show pyronecrosis: a pathogen-induced, cathepsin B-dependent, necrosis-like programmed cell death that is characteristically observed in *NLRP3*-mutant monocytes/macrophages.[22,23] When we compared LDH secretion as a marker of membrane rupture, we found that LPS stimulation evoked a significantly higher LDH secretion only from the mutant iPS-MPs, which was inhibited by the cathepsin B inhibitor, CA074Me (Figure 3E).

Despite the low percentage of mutant cells, the clinical manifestation of mosaic CINCA patients is similar to that of patients with a heterozygous mutation.[9,10] We hypothesized that an interaction between the mutant and wild-type macrophages leads to exacerbation of the inflammation. To test this hypothesis, we modeled a mosaic condition by coculturing mutant and wild-type cells. After stimulating mutant iPS-MPs with LPS in separate cultures or in cocultures with wild-type counterparts, we determined the IL-1β level in the supernatant. We found that the IL-1β secretion significantly increased after coculture (Figure 4A; supplemental Figure 4A). Although increasing the cell concentration raised the total amount of the IL-1β secretion from mutants, it did not accelerate the IL-1β secretion per cell from mutant iPS-MPs or enhance the secretion from wild-type macrophages (Figure 4B). To determine the ratio of mutant/wild-type cells at which the additional IL-1β secretion is most enhanced, we changed the ratio using a fixed number of mutant iPS-MPs and increasing the number of wild-type iPS-MPs. We observed a significant increase only at a percentage of 25% mutant macrophages (Figure 4C). Thus, we capitulated, at least in part, the patient's mosaic condition in vitro.

Next, we tried to elucidate whether the interaction is mediated by some humoral factor(s), but supernatant transfer did not facilitate the IL-1β secretion (Figure 4D). As a candidate that may mediate this interaction, we selected ATP because necrotic cells trigger NLRP3-inflammasome activation in part through ATP release.[24] We therefore investigated whether the necrosis-induced ATP secretion activates the wild-type iPS-MPs using ATP receptor antagonists, oxidized ATP (oATP) and PPADS. Although both antagonists markedly inhibited the IL-1β secretion after LPS plus ATP stimulation (supplemental Figure 4B), neither of them abrogated the additional IL-1β secretion in the mixed culture (Figure 4E; compare column 2 with column 3, and column 4 with column 5). The IL-1β secretion from mutant iPS-MPs may have decreased because of off-target effects of oATP.[25] Overall, although it remains to be elucidated how this effect is mediated, these results suggest that the interaction between mutant and wild-type macrophages may enhance IL-1β secretion in mosaic patients.

Validation for future applications for drug screening

An NLRP3-targeted therapeutic approach would be attractive because (1) the progressive arthropathy despite anti–IL-1 therapy indicates that the presence of additional proteins processed by the

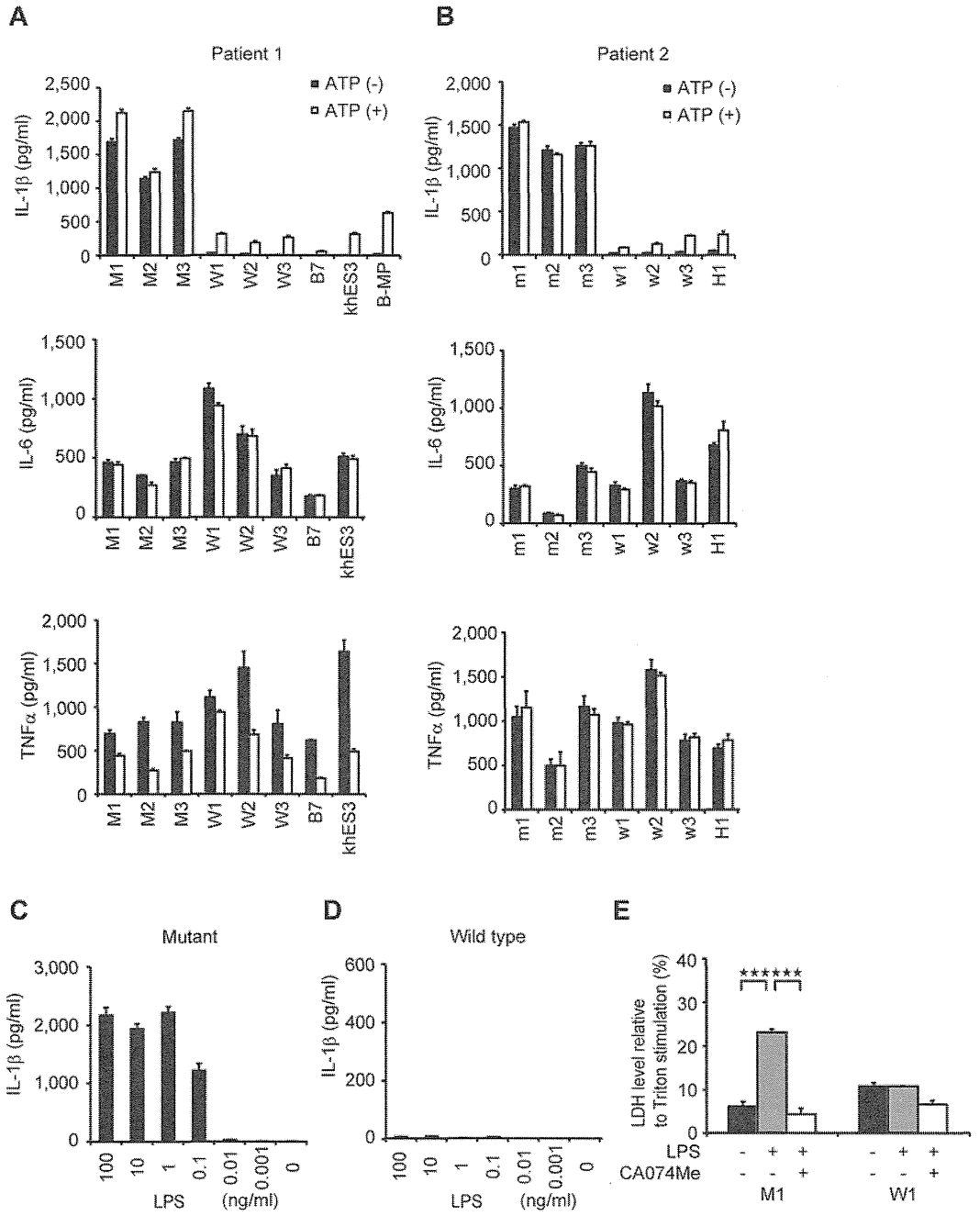

Figure 3. Elucidation of the pathogenesis of somatic mosaic CINCA syndrome. (A) Cytokine secretion from iPS-MPs derived from patient 1. After stimulating iPS-MPs by LPS with or without ATP, we determined the IL-1β (top panel), IL-6 (middle panel), or TNFα (bottom panel) level of the supernatant. n = 3. (B) Cytokine secretion from iPS-MPs derived from patient 2, determined as in panel A. (C) IL-1β secretion from mutant iPS-MPs in the presence of 10-fold dilutions of LPS from 100 ng/mL. n = 3. (D) IL-1β secretion from wild-type iPS-MPs, determined as in panel C. (E) LDH secretion from iPS-MPs stimulated with LPS in the presence or absence of the cathepsin B inhibitor, CA074Me. n = 3. Data are mean ± SEM. ***$P < .001$ (Student t test).

inflammasome is also involved in the pathogenesis of CINCA syndrome; (2) specific inhibition of the NLRP3-inflammasome can avoid unfavorable suppression of other IL-1β–processing pathways in response to various triggers; and (3) these drugs may be also effective for various other NLRP3-related chronic inflammatory conditions, such as Alzheimer disease, diabetes, severe gout, and atherosclerosis.[26-30] Because drug screening using NLRP3 autoactivated cells has not been described previously, we examined whether the iPS-MPs from CINCA patients can serve as a prototype for seeking drug candidates that directly modulate NLRP3-inflammasome activation.

When wild-type iPS-MPs were stimulated with LPS and ATP in the presence of various inhibitors, inhibitors known to modulate molecules upstream of the NLRP3-inflammasome (a protein synthesis inhibitor, cycloheximide, and an NF-κB inhibitor, MG132), downstream of the inflammasome (a caspase-1 inhibitor, Ac-YVAD-CHO), and both upstream of and the inflammasome itself[31] (Bay11-7082) successfully inhibited IL-1β secretion (Figure 5A). Although the precise mechanism is unknown, a cathepsin B inhibitor, CA074Me, also efficiently inhibited IL-1β secretion. As expected, upstream inhibitors inhibited the secretion of other cytokines, such as IL-6 and IL-8, but a downstream inhibitor,

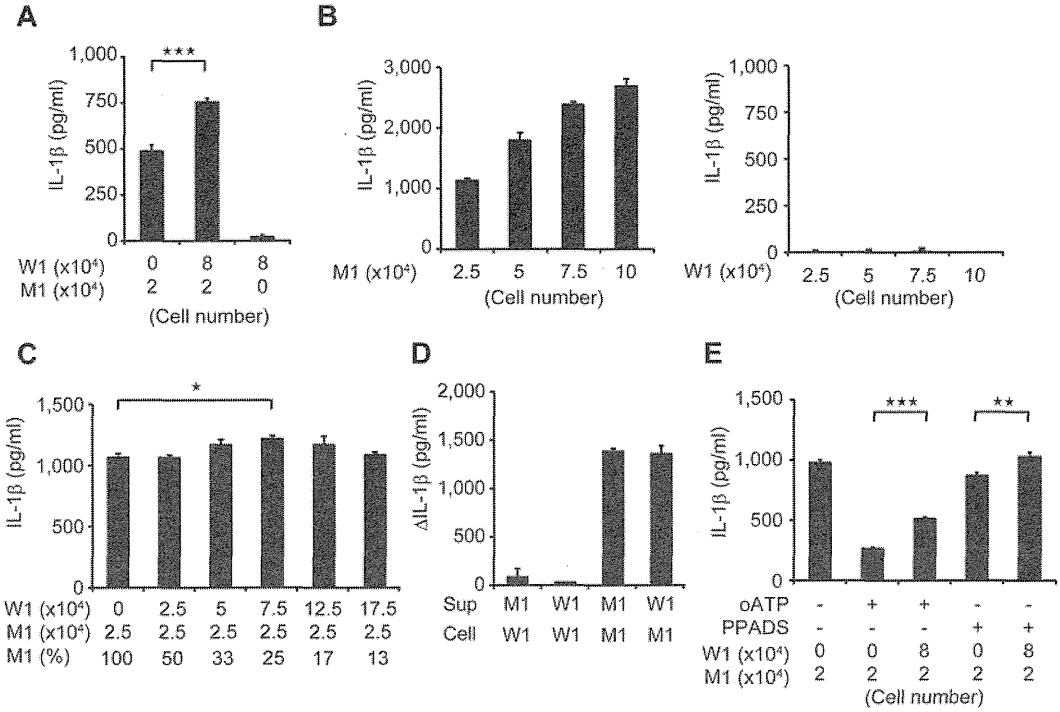

Figure 4. Remodeling mosaicism by coculturing mutant and wild-type iPS-MPs. (A) IL-1β secretion from cocultured iPS-MPs. We used 2 × 10^4 mutant iPS-MPs (M1) and 8 × 10^4 wild-type iPS-MPs (W1) as indicated. n = 6. (B) IL-1β secretion from various numbers of mutant (left panel) or wild-type (right panel) iPS-MPs. The iPS-MPs were seeded at the indicated numbers. n = 3. (C) IL-1β secretion from iPS-MPs that were cocultured at various ratios. The wild-type or mutant iPS-MPs were seeded at the numbers indicated in the first and second rows, respectively. The percentage of mutants is indicated in the third row; n = 3. (D) Increase of IL-1β levels during stimulation by the supernatant. The supernatant was harvested from the wells of the indicated iPS-MPs (Sup) and transferred to the wells of other iPS-MPs (Cell); n = 3. (E) IL-1β secretion from cocultured iPS-MPs in the presence of the ATP receptor antagonist, oATP (300μM) or PPADS (300μM). We used 2 × 10^4 mutant iPS-MPs (M1) and 8 × 10^4 wild-type iPS-MPs (W1) as indicated. n = 6. Data are mean ± SEM. ***$P < .001$ (Student t test). **$P < .01$ (Student t test). *$P < .05$ (Student t test).

Ac-YVAD-CHO, specifically affected IL-1β secretion (Figure 5A). Although CA074Me and Ac-YVAD-CHO inhibited IL-1β secretion regardless of the second signals that were present, PPADS, an inhibitor of extracellular ATP signaling, failed to inhibit IL-1β secretion by following exposure to other second signals, such as monosodium urate and silica crystals (Figure 5B), proving that wild-type iPS-MPs can be activated in a second signal-dependent manner. Therefore, the results of the wild-type iPS-MP-based compound screening depended on the choice of second signals, and such a screening makes it possible to extract candidate compounds that modulate specific second signaling pathways.

Next, we examined the response of mutant iPS-MPs to the inhibitors. In the absence of inhibitors, mutant iPS-MPs secreted a higher level of IL-1β, but treatment with inhibitors dose-dependently decreased IL-1β secretion to the comparable level produced by WT iPS-MPs (Figure 5C). We thus demonstrated the efficacy of these chemical compounds, even for excessive IL-1β production by constitutively hyperactivated inflammasomes. As expected, the mutant iPS-MPs did not respond to PPADS, confirming their autoactivation in a second signal-independent manner (Figure 5D). Therefore, because they can be activated independently from the type of second signals, mutant iPS-MP–based screening would enable the exclusion of compounds that inhibit IL-1β secretion depending on a specific type of second signal transduction. Overall, through using the IL-1β inhibition as the initial criteria and weeding out upstream inhibitors by measuring the levels of other cytokines, we can use *NLRP3*-mutant iPS-MPs to screen for drugs for CINCA syndrome and possibly for other NLRP3-related chronic inflammatory conditions.

Discussion

Since the first identification of a CINCA syndrome patient carrying *NLRP3* mutation as somatic mosaicism,[20] it has been controversial whether the small fraction of *NLRP3*-mutated cells actually causes the strong autoinflammation. It remained unanswered because of the difficulty to separately obtain live mutant and nonmutant blood cells. In this study, we reprogrammed fibroblasts from mosaic patients and obtained macrophages with different genotypes. By showing that only *NLRP3*-mutant iPS-MPs exhibit the distinct proinflammatory phenotype, we demonstrated that the *NLRP3*-mutant macrophages are mainly responsible for the pathogenesis of mosaic CINCA syndrome.

In this study, we established both *NLRP3*-mutant and nonmutant iPSC clones from the same person. One of the potential limitations of studies with patient-derived iPSCs is the difficulty in obtaining isogenic control counterparts, which do not carry the responsible mutations. One possible strategy to solve this problem is to correct the affected gene locus of patient-derived iPSC clones using novel techniques that facilitate homologous recombination.[32,33] As another solution, both affected and control iPSC clones can be obtained from patients of some X-linked hereditary diseases because each iPSC clone originated from somatic cells carrying either a mutated or nonmutated allele as an active X chromosome.[34-36] In the present study, we have retrieved both mutant and wild-type iPSC clones from patients with somatic autosomal mutations. These clones theoretically have the same genetic backgrounds, except for the *NLRP3* gene, and should serve as an ideal pair of mutant and control clones for disease research.

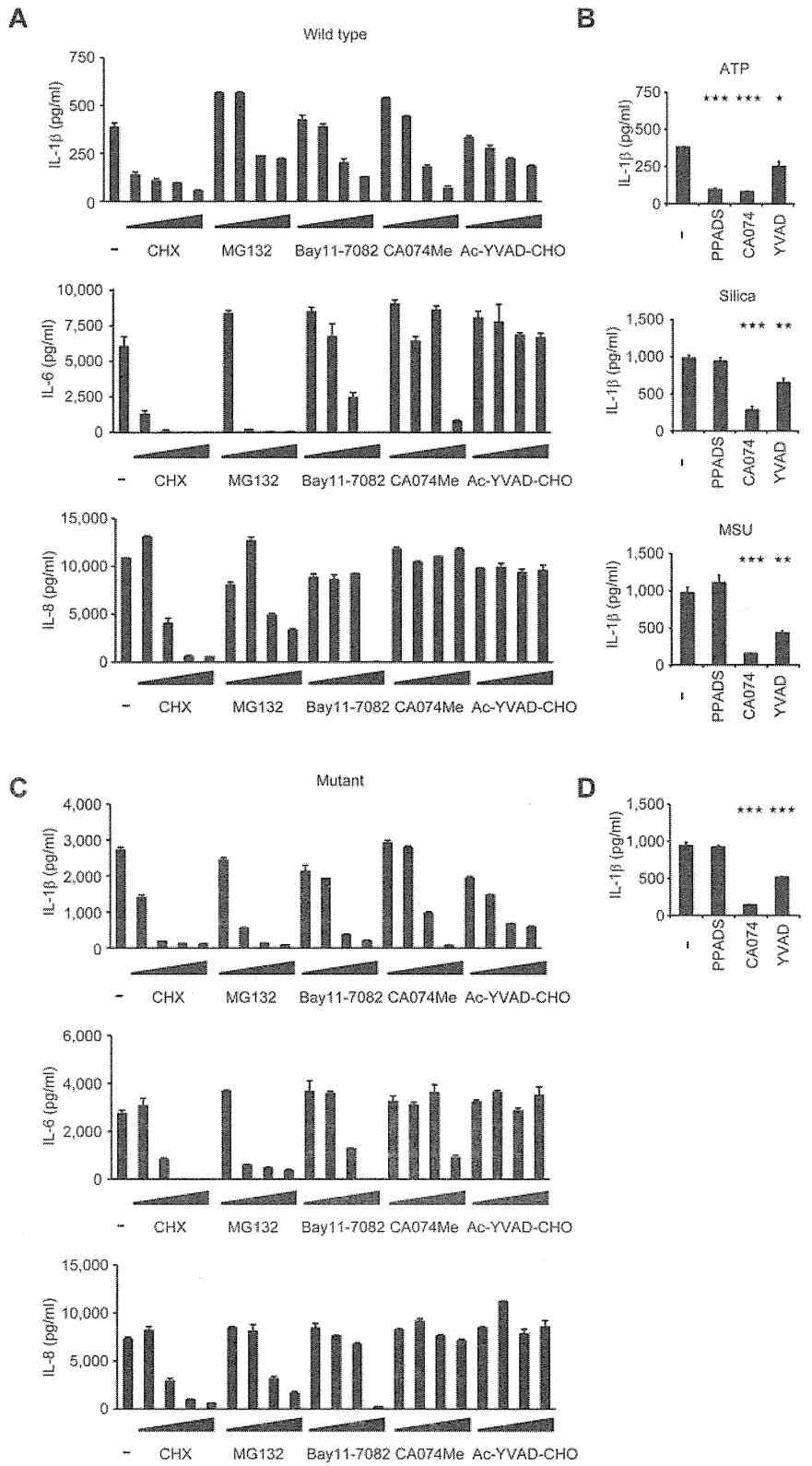

Figure 5. Validation of the cells for future applications for drug screening. (A) Inhibition of IL-1β (top panel), IL-6 (middle panel), or IL-8 (bottom panel) secretion from wild-type iPS-MPs by various inhibitors. The iPS-MPs were cultured for 2 hours in the presence of 100μM cycloheximide (CHX), 100μM MG132, 10μM Bay11-7082, 25μM CA074Me, 50μM Ac-YVAD-CHO, as well as 10-fold dilutions of each inhibitor, except CA074Me (which was diluted 5-fold), followed by LPS treatment plus ATP stimulation. n = 3. (B) The differential inhibition of IL-1β secretion from wild-type iPS-MPs by various inhibitors. In the presence of inhibitors, such as PPADS (300μM), CA074Me (25μM), or Ac-YVAD-CHO (50μM), LPS-primed wild-type iPS-MPs were stimulated with second signal triggers, such as ATP for 1 hour (top panel), silica crystals for 1 hour (middle panel), or monosodium urate crystals for 3 hours (bottom panel). n = 3. (C) Inhibition of IL-1β (top panel), IL-6 (middle panel), or IL-8 (bottom panel) secretion from mutant iPS-MPs by various inhibitors was evaluated as in panel A; n = 3. (D) Inhibition of IL-1β secretion from mutant iPS-MPs by various inhibitors. In the presence of inhibitors, such as PPADS (300μM), CA074Me (25μM), or Ac-YVAD-CHO (50μM), mutant iPS-MPs were stimulated with LPS for 4 hours. n = 3. Data are mean ± SEM. ***$P < .001$ (Student t test). **$P < .01$ (Student t test). *$P < .05$ (Student t test).

In addition to obtaining isogenic controls, iPSCs from patients with somatic autosomal mutations enable dissection and modeling of somatic mosaicism. Despite the fact that each person contains various minor somatic mutations,[37] the effects of mosaicism can often be overlooked because of the difficulty in assessing the possible biologic effects caused by the small cell populations carrying the genetic alterations. Here we dissected somatic mosaicism by obtaining the component cells with heterogeneous genetic identity separately and established an in vitro model to evaluate the interaction between these cells, although precise mechanism of interaction remains to be elucidated. As an approach to determining the disease-causing potential of a specific somatic mutation found in a person, iPSC technology provides advantages compared with ordinary methods, such as the use of transgenic cell lines. First,

iPSCs can be differentiated into the affected cell types or tissues, allowing direct functional assays to be performed that are associated with the pathology. Second, because the disease-causing potential of some mutations is dependent on the genetic backgrounds of the patients,[38] it may be better to obtain both mutant and wild-type clones from a single mosaic patient to more accurately assess the impact of the mutation(s).

Considering that a mutation of *NLRP3* in 10% of the cells is sufficient to cause a distinct disease phenotype, somatic mutations of various genes at an even rarer frequency may also affect the biologic characteristics of a person. Because the presence of the *NLRP3* mutation did not affect the efficacy of reprogramming to the iPSCs, we may be able to obtain both mutant and wild-type iPSC clones from CINCA syndrome patients who carry *NLRP3* mutant cells at a lower percentage. In some diseases, such as Fanconi anemia, however, mutant cells may be resistant to reprogramming.[39,40] Even though there are some possible limitations, establishing both mutant and wild-type iPSC clones is a promising approach to dissect the extent and role of somatic mosaicism.

We demonstrated that several inhibitors that are considered to be effective against CINCA syndrome actually attenuated the disease-relevant phenotype of iPSC-derived macrophages. Before a successful drug screening using iPSC-derived somatic cells can be developed, several limitations need to be overcome, such as the heterogeneity of differentiation and difficulties associated with purification.[18] In this report, we used an efficient and robust differentiation protocol and obtained plenty of macrophages free from the clonal variations.

In conclusion, we elucidated the pathologic roles of both mutant and wild-type cells in mosaic CINCA syndrome patients. After obtaining iPSC-derived macrophages in large quantity and with high purity, we showed they are applicable for drug screening. The iPSC-based approach may help to illuminate the pathogenesis of various diseases that are caused by somatic mosaicism, and facilitate drug discovery for the treatment of NLRP3-related inflammatory diseases.

Acknowledgments

The authors thank the CINCA syndrome patients who participated in this study; Y. Sasaki, Y. Jindai, A. Okada, M. Narita, A. Nagahashi, T. Ohkame, S. Nishimoto, Y. Inoue, and S. Arai for technical assistance; I. Kato for help with animal experiments; M. Nakagawa, K. Okita, Y. Yoshida, T. Aoi, and M. Yanagimachi for scientific comments; and R. Kato, E. Nishikawa, S. Takeshima, Y. Otsu, H. Hasaba, H. Watanabe, T. Ishii, H. Kurokawa, N. Takasu, and Y. Takao for administrative assistance.

This work was supported by the Ministry of Health, Labor and Welfare (N.M. and T.N.), the Ministry of Education, Culture, Sports, Science and Technology (MEXT; N.M. and T.N.), the Leading Project of MEXT (S.Y. and T.N.), the Promotion of Fundamental Studies in Health Sciences of National Institute of Biomedical Innovation (S.Y.), the Funding Program for World-Leading Innovative Research and Development on Science and Technology (FIRST Program) of Japan Society for the Promotion of Science (JSPS; T.N., and S.Y.), JSPS and MEXT (Grants-in-Aid for Scientific Research; S.Y.), JSPS (T.N., T.T., and M.K.S.), the Takeda Science Foundation, SENSHIN Medical Research Foundation, and Suzuken Memorial Foundation to (M.K.S.).

Authorship

Contribution: T.T. planned the project, established iPSCs, performed experimental work, analyzed data, and prepared the manuscript; K.T. planned the project, established iPSCs, and analyzed data; M.Y., S.T., and S.N. performed experimental work; K.O., A.N., and T.H. analyzed data; R.N. and N.K. planned the project; H.H. and M.M. performed *L monocytogenes* infection; N.M. and J.E.H. performed electron microscopy; T.Y. identified retroviral integration sites; A.W. performed bisulfite sequencing; A.S.-O. and S.O. analyzed CNV; I.A. established iPSCs; S.Y. and T.N. planned the project and analyzed data; M.K.S. planned the project, analyzed data, and prepared the manuscript; and all authors read and approved the manuscript.

Conflict-of-interest disclosure: S.Y. is a member without salary of the scientific advisory boards of iPierian, iPS Academia Japan, and Megakaryon Corporation. The remaining authors declare no competing financial interests.

Correspondence: Megumu K. Saito, Center for iPS Cell Research and Application, Kyoto University, Kyoto 606-8507, Japan; e-mail: msaito@cira.kyoto-u.ac.jp.

References

1. Prieur AM, Griscelli C, Lampert F, et al. A chronic, infantile, neurological, cutaneous and articular (CINCA) syndrome: a specific entity analysed in 30 patients. *Scand J Rheumatol Suppl.* 1987;66:57-68.
2. Aksentijevich I, Nowak M, Mallah M, et al. De novo CIAS1 mutations, cytokine activation, and evidence for genetic heterogeneity in patients with neonatal-onset multisystem inflammatory disease (NOMID): a new member of the expanding family of pyrin-associated autoinflammatory diseases. *Arthritis Rheum.* 2002;46(12):3340-3348.
3. Feldmann J, Prieur AM, Quartier P, et al. Chronic infantile neurological cutaneous and articular syndrome is caused by mutations in CIAS1, a gene highly expressed in polymorphonuclear cells and chondrocytes. *Am J Hum Genet.* 2002;71(1):198-203.
4. Bauernfeind FG, Horvath G, Stutz A, et al. Cutting edge: NF-kappaB activating pattern recognition and cytokine receptors license NLRP3 inflammasome activation by regulating NLRP3 expression. *J Immunol.* 2009;183(2):787-791.
5. Mariathasan S, Weiss DS, Newton K, et al. Cryopyrin activates the inflammasome in response to toxins and ATP. *Nature.* 2006;440(7081):228-232.
6. Gattorno M, Tassi S, Carta S, et al. Pattern of interleukin-1beta secretion in response to lipopolysaccharide and ATP before and after interleukin-1 blockade in patients with CIAS1 mutations. *Arthritis Rheum.* 2007;56(9):3138-3148.
7. Goldbach-Mansky R, Dailey NJ, Canna SW, et al. Neonatal-onset multisystem inflammatory disease responsive to interleukin-1beta inhibition. *N Engl J Med.* 2006;355(6):581-592.
8. Neven B, Marvillet I, Terrada C, et al. Long-term efficacy of the interleukin-1 receptor antagonist anakinra in ten patients with neonatal-onset multisystem inflammatory disease/chronic infantile neurologic, cutaneous, articular syndrome. *Arthritis Rheum.* 2010;62(1):258-267.
9. Saito M, Nishikomori R, Kambe N, et al. Disease-associated CIAS1 mutations induce monocyte death, revealing low-level mosaicism in mutation-negative cryopyrin-associated periodic syndrome patients. *Blood.* 2008;111(4):2132-2141.
10. Tanaka N, Izawa K, Saito MK, et al. High incidence of NLRP3 somatic mosaicism in patients with chronic infantile neurologic, cutaneous, articular syndrome: results of an International Multicenter Collaborative Study. *Arthritis Rheum.* 2011;63(11):3625-3632.
11. Masters SL, Simon A, Aksentijevich I, Kastner DL. Horror autoinflammaticus: the molecular pathophysiology of autoinflammatory disease. *Annu Rev Immunol.* 2009;27:621-668.
12. Youssoufian H, Pyeritz RE. Mechanisms and consequences of somatic mosaicism in humans. *Nat Rev Genet.* 2002;3(10):748-758.
13. Erickson RP. Somatic gene mutation and human disease other than cancer: an update. *Mutat Res.* 2010;705(2):96-106.
14. Ariga T, Kondoh T, Yamaguchi K, et al. Spontaneous in vivo reversion of an inherited mutation in the Wiskott-Aldrich syndrome. *J Immunol.* 2001;166(8):5245-5249.

15. Nishikomori R, Akutagawa H, Maruyama K, et al. X-linked ectodermal dysplasia and immunodeficiency caused by reversion mosaicism of NEMO reveals a critical role for NEMO in human T-cell development and/or survival. *Blood.* 2004; 103(12):4565-4572.
16. Lutskiy MI, Beardsley DS, Rosen FS, Remold-O'Donnell E. Mosaicism of NK cells in a patient with Wiskott-Aldrich syndrome. *Blood.* 2005;106(8):2815-2817.
17. Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell.* 2007;131(5): 861-872.
18. Grskovic M, Javaherian A, Strulovici B, Daley GQ. Induced pluripotent stem cells: opportunities for disease modelling and drug discovery. *Nat Rev Drug Discov.* 2011;10(12):915-929.
19. Hanna J, Markoulaki S, Schorderet P, et al. Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell.* 2008;133(2): 250-264.
20. Saito M, Fujisawa A, Nishikomori R, et al. Somatic mosaicism of CIAS1 in a patient with chronic infantile neurologic, cutaneous, articular syndrome. *Arthritis Rheum.* 2005;52(11):3579-3585.
21. Nakano T, Kodama H, Honjo T. Generation of lymphohematopoietic cells from embryonic stem cells in culture. *Science.* 1994;265(5175):1098-1101.
22. Fujisawa A, Kambe N, Saito M, et al. Disease-associated mutations in CIAS1 induce cathepsin B-dependent rapid cell death of human THP-1 monocytic cells. *Blood.* 2007;109(7):2903-2911.
23. Willingham SB, Bergstralh DT, O'Connor W, et al. Microbial pathogen-induced necrotic cell death mediated by the inflammasome components CIAS1/cryopyrin/NLRP3 and ASC. *Cell Host Microbe.* 2007;2(3):147-159.
24. Iyer SS, Pulskens WP, Sadler JJ, et al. Necrotic cells trigger a sterile inflammatory response through the Nlrp3 inflammasome. *Proc Natl Acad Sci U S A.* 2009;106(48):20388-20393.
25. Beigi RD, Kertesy SB, Aquilina G, Dubyak GR. Oxidized ATP (oATP) attenuates proinflammatory signaling via P2 receptor-independent mechanisms. *Br J Pharmacol.* 2003;140(3):507-519.
26. Martinon F, Petrilli V, Mayor A, Tardivel A, Tschopp J. Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature.* 2006; 440(7081):237-241.
27. Halle A, Hornung V, Petzold GC, et al. The NALP3 inflammasome is involved in the innate immune response to amyloid-beta. *Nat Immunol.* 2008;9(8):857-865.
28. Duewell P, Kono H, Rayner KJ, et al. NLRP3 inflammasomes are required for atherogenesis and activated by cholesterol crystals. *Nature.* 2010; 464(7293):1357-1361.
29. Masters SL, Dunne A, Subramanian SL, et al. Activation of the NLRP3 inflammasome by islet amyloid polypeptide provides a mechanism for enhanced IL-1beta in type 2 diabetes. *Nat Immunol.* 2010;11(10):897-904.
30. Vandanmagsar B, Youm YH, Ravussin A, et al. The NLRP3 inflammasome instigates obesity-induced inflammation and insulin resistance. *Nat Med.* 2011;17(2):179-188.
31. Juliana C, Fernandes-Alnemri T, Wu J, et al. Anti-inflammatory compounds parthenolide and Bay 11-7082 are direct inhibitors of the inflammasome. *J Biol Chem.* 2010;285(13):9792-9802.
32. Aizawa E, Hirabayashi Y, Iwanaga Y, et al. Efficient and accurate homologous recombination in hESCs and hiPSCs using helper-dependent adenoviral vectors. *Mol Ther.* 2012;20(2):424-431.
33. Soldner F, Laganiere J, Cheng AW, et al. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell.* 2011;146(2):318-331.
34. Cheung AY, Horvath LM, Grafodatskaya D, et al. Isolation of MECP2-null Rett syndrome patient hiPS cells and isogenic controls through X-chromosome inactivation. *Hum Mol Genet.* 2011;20(11):2103-2115.
35. Kim KY, Hysolli E, Park IH. Neuronal maturation defect in induced pluripotent stem cells from patients with Rett syndrome. *Proc Natl Acad Sci U S A.* 2011;108(34):14169-14174.
36. Pomp O, Dreesen O, Leong DF, et al. Unexpected X chromosome skewing during culture and reprogramming of human somatic cells can be alleviated by exogenous telomerase. *Cell Stem Cell.* 2011;9(2):156-165.
37. Gore A, Li Z, Fung HL, et al. Somatic coding mutations in human induced pluripotent stem cells. *Nature.* 2011;471(7336):63-67.
38. Crotti L, Lundquist AL, Insolia R, et al. KCNH2-K897T is a genetic modifier of latent congenital long-QT syndrome. *Circulation.* 2005;112(9): 1251-1258.
39. Raya A, Rodriguez-Piza I, Guenechea G, et al. Disease-corrected haematopoietic progenitors from Fanconi anemia induced pluripotent stem cells. *Nature.* 2009;460(7251):53-59.
40. Müller LU, Milsom MD, Harris CE, et al. Overcoming reprogramming resistance of Fanconi anemia cells. *Blood.* 2012;119(23):5449-5457.

DNA Research **19**, 143–152, (2012) doi:10.1093/dnares/dsr047
Advance Access Publication on 24 January 2012

Detection of Base Substitution-Type Somatic Mosaicism of the *NLRP3* Gene with >99.9% Statistical Confidence by Massively Parallel Sequencing

Kazushi Izawa[1,†], Atsushi Hijikata[2,†], Naoko Tanaka[1], Tomoki Kawai[1], Megumu K Saito[3], Raphaela Goldbach-Mansky[4], Ivona Aksentijevich[5], Takahiro Yasumi[1], Tatsutoshi Nakahata[3], Toshio Heike[1], Ryuta Nishikomori[1,*], and Osamu Ohara[2,6,*]

Department of Pediatrics, Kyoto University Graduate School of Medicine, 54 Shogoin Sakyo, Kyoto 606-8507, Japan[1]; Laboratory for Immunogenomics, RIKEN Research Center for Allergy and Immunology, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho Tarumi-ku, Yokohama, Kanagawa 230-0045, Japan[2]; Clinical Application Department, Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto, Japan[3]; Translational Autoinflammatory Disease Section NIH/NIAMS, Bethesda, MD, USA[4]; The National Human Genome Research Institute, Bethesda, MD, USA[5] and Department of Human Genome Research, Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan[6]

*To whom correspondence should be addressed. Tel. +81 75-751-3291 (R.N.); +81 438-52-3913/+81 45-503-9696 (O.O.). Fax. +81 75-752-2361 (R.N.); +81 438-52-3914/+81 45-503-9694 (O.O.).
Email: rnishiko@kuhp.kyoto-u.ac.jp (R.N.); ohara@kazusa.or.jp/oosamu@rcai.riken.jp (O.O.)

Edited by Mitsuo Oshimura
(Received 18 November 2011; accepted 26 December 2011)

Abstract

Chronic infantile neurological cutaneous and articular syndrome (CINCA), also known as neonatal-onset multisystem inflammatory disease (NOMID), is a dominantly inherited systemic autoinflammatory disease and is caused by a heterozygous germline gain-of-function mutation in the *NLRP3* gene. We recently found a high incidence of *NLRP3* somatic mosaicism in apparently mutation-negative CINCA/NOMID patients using subcloning and subsequent capillary DNA sequencing. It is important to rapidly diagnose somatic *NLRP3* mosaicism to ensure proper treatment. However, this approach requires large investments of time, cost, and labour that prevent routine genetic diagnosis of low-level somatic *NLRP3* mosaicism. We developed a routine pipeline to detect even a low-level allele of *NLRP3* with statistical significance using massively parallel DNA sequencing. To address the critical concern of discriminating a low-level allele from sequencing errors, we first constructed error rate maps of 14 polymerase chain reaction products covering the entire coding *NLRP3* exons on a Roche 454 GS-FLX sequencer from 50 control samples without mosaicism. Based on these results, we formulated a statistical confidence value for each sequence variation in each strand to discriminate sequencing errors from real genetic variation even in a low-level allele, and thereby detected base substitutions at an allele frequency as low as 1% with 99.9% or higher confidence.

Key words: next generation sequencing; mosaicism; DNA diagnosis; chronic infantile neurological cutaneous and articular syndrome

1. Introduction

Chronic infantile neurological cutaneous and articular syndrome (CINCA; MIM #607115), also known as neonatal-onset multisystem inflammatory disease (NOMID), is a dominantly inherited autoinflammatory disease that is characterized by neonatal onset and a triad of symptoms, including an urticarial-like skin rash, neurological manifestations, and arthritis/arthropathy.[1–3] Patients often experience

† These authors contributed equally to this work.

recurrent fever and systemic inflammation. CINCA/NOMID is the most severe clinical phenotype in the spectrum of cryopyrin-associated periodic syndromes (CAPS), which also include two less severe but phenotypically similar syndromes, familial cold autoinflammatory syndrome (FCAS; MIM #120100), and Muckle–Wells syndrome (MWS; MIM #191900). CAPS are caused by mutations in the *NLRP3* gene, which is a member of the Nod-like receptor (NLR) family of the innate immune system.[4–6]

Approximately 60% of CINCA/NOMID patients carry heterozygous germline missense mutations in the *NLRP3* coding region (mutation-positive patients).[7] More than 80 different disease-causing mutations have been reported to date.[8] However, the remaining clinically diagnosed CINCA/NOMID patients (~40%) show no heterozygous germline *NLRP3* mutation based on conventional DNA sequencing-based genetic analyses (mutation-negative patients). In a previous international collaborative study, we found that there was a high incidence of somatic *NLRP3* mosaicism in mutation-negative CINCA/NOMID patients worldwide.[9] The level of mosaicism ranges from 4.2 to 35.8% (median = 10.2%). Rapidly diagnosing somatic *NLRP3* mosaicism is important to ensure proper treatment. However, the conventional approach used to identify somatic mosaicism of the *NLRP3* gene is time and labour intensive due to the subcloning of the *NLRP3* exon polymerase chain reaction (PCR) products, hereafter designated as amplicons, followed by capillary DNA sequencing of more than 100 subclones for each patient. Thus, this approach is not suitable to routinely diagnose somatic mosaicism of the *NLRP3* gene and additional labour and time will be required to reliably identify somatic mosaicism that occurs at a lower rate. The aim of the present study was to establish a new method that can be used to reliably diagnose somatic mosaicism using the *NLRP3* gene as a model. Massively parallel DNA sequencing (MPS) technology is an obvious method of choice to identify somatic mosaicism, and this approach has been already reported by other groups.[10–12] However, a well-known caveat of MPS is the high rate of sequencing errors, which cannot be disregarded when identifying low-level somatic mosaicism. To our knowledge, there have been no reports of a reliable method to discriminate MPS sequencing errors from somatic mosaicism with statistical confidence.

In this study, we first analysed the patterns of sequencing errors in *NLRP3* coding exons at a single-residue resolution by MPS using a Roche 454 GS-FLX sequencer and then constructed an error rate map for each base position in the *NLRP3* exons. Based on the error rate map, we could formulate a discrimination pipeline of somatic mosaicism from sequencing errors and thereby detect new somatic mosaicism in mutation-negative CINCA/NOMID patients, whose somatic mutations were subsequently confirmed by subcloning and Sanger sequencing. This approach can also be generally used to identify low-level somatic mosaicism in other genes.

2. Patients and methods

2.1. *Patients and DNA materials*

Patients were clinically diagnosed with CAPS by their referring physicians and the *NLRP3* gene was examined using the conventional Sanger sequencing method. DNA samples were obtained from Japanese *NLRP3* somatic mosaic patients ($n = 5$) who have been previously described,[9,13] CAPS patients ($n = 5$) with heterozygous *NLRP3* mutations, and healthy donors ($n = 50$). Genomic DNA samples from mutation-negative CINCA/NOMID patients ($n = 10$) were obtained from the National Institute of Health, Bethesda, USA. To generate DNA samples with no mosaicism, we constructed a set of subcloned plasmids containing each exon and its flanking intronic regions in the *NLRP3* gene from healthy donor genomic DNA using a Topo TA cloning kit (Invitrogen, San Diego, CA, USA). The cloned plasmids containing each exon and the flanking regions were validated by Sanger sequencing. Written informed consent was obtained from all the patients and their families. The study was approved by the ethical committees of Kyoto University and Kazusa DNA Research Institute and was conducted in accordance with the Helsinki Declaration.

2.2. *MPS of* NLRP3 *gene amplicons*

Genomic DNA samples were extracted from whole blood or peripheral blood mononuclear cells as previously described. We used a two-step PCR assay and pooled sample libraries for MPS. To cover the entire *NLRP3* coding exonic regions and flanking intronic regions, 14 amplicons were designed to be as long as an average read length for a 454 GS-FLX sequencer (up to 450 bases) and then amplified from each genomic DNA sample (Fig. 1A). The sequences of the PCR primers that were used to generate these 14 amplicons are provided in Supplementary Table S1. The upper and lower amplicon-specific primer sequences were flanked by common 15-base adapter sequences (TGTAAAACGACGGCC and GGAAACAGCTATGAC for the upper and lower primers, respectively) at the 5′ end in order to fuse the primer-binding sequence for MPS in the second-step PCR. The first PCR amplifications were performed in 50-μl reactions using 30 ng of genomic DNA, 1× PrimeSTAR GXL buffer, 0.2 mM of each dNTP,

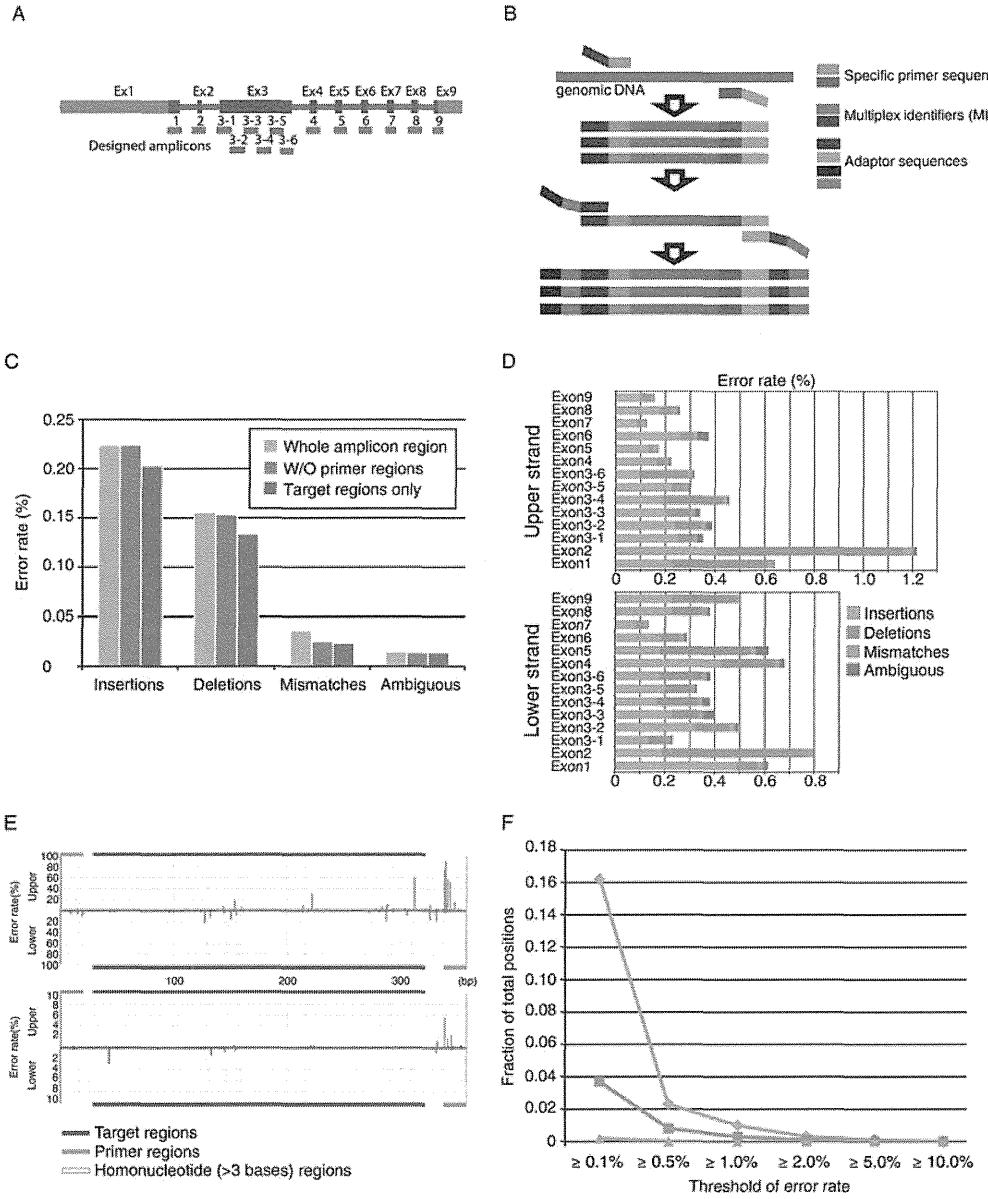

Figure 1. The amplicon analysis for *NLRP3* exons and its error rate. (A) Exon–intron structure of the *NLPR3* gene. Thick and thin rectangles depict exons and introns, respectively. Blue thick rectangles indicate the CDS region. The 14 designed amplicons (red) for nine exons are shown under the exon–intron structure. (B) Amplicon design schema. (C) Error rate for each error category in the region of entire amplicon (pale blue), that without designed primer regions (light blue), and the target regions (CDS + flanking intron; dark blue), respectively. (D) Strand-wise error rate for each amplicon. (E) Error rates along the amplicon sequence of exon 1 in each strand for insertions and deletions in the upper panel and mismatches and ambiguous base calls in the lower panel. The orange and blue lines depict the primer and target regions, respectively. The yellow shaded area depicts the homonucleotide ($n > 3$) region. The colour representation for the bars is the same as (D). (F) Co-occurrence error rate in both strands. The fraction of positions where a certain error occurred with the error rate for insertions, deletions, and mismatches. The colour representation is the same as in (D) and (E).

12.5 pmol of each forward and reverse primer, and 1.25 U of PrimeSTAR GXL DNA polymerase (Takara Bio, Shiga, Japan). The thermal cycling profile consisted of an initial denaturation step at 98°C for 1 min, followed by 28–32 cycles of 10 s denaturation at 98°C, 15 s of annealing at 60°C, and a 30 s extension at 68°C. The lengths of the PCR products ranged from 291 to 421 bp. The second PCR amplifications were performed using primers with adapter sequences at the 3′ end and Multiplex Identifier (MID) sequences at the 5′ end (Fig. 1B), which was used as a tag for each sample. The PCR reactions were performed in 50-μl volumes using 0.5 μl of the first PCR products, 1× PrimeSTAR GXL buffer, 0.2 mM of each dNTP, 12.5 pmol of each forward and reverse primer, and 1.25 U of PrimeSTAR GXL

DNA polymerase to attach the anchor sequences for MPS. The thermal cycling profile consisted of an initial denaturation step at 98°C for 20 s, followed by 20 cycles of 10 s denaturation at 98°C, 15 s of annealing at 60°C, and a 40 s extension at 68°C.

After confirming the amount and integrity of the PCR products by agarose gel electrophoresis, we mixed virtually equal amounts of the respective PCR amplicons that were generated using the same genomic DNA and applied the samples to a 454 Genome Sequencer (GS)-FLX system (Roche Diagnostics Corp., USA). All amplicons were amplified by emPCR and sequenced together in a multiplex fashion. MPS on this platform was performed as instructed by Roche. The sequencing reads from each of the pooled libraries were identified by their MID tags.

2.3. *Sequence data analysis*

The sequence read data were generated using GS RunProcessor ver.2.5.3 with default settings. Reads were sorted according to the MID tag sequences and were mapped to the reference amplicon sequences using the BLAT program[14] with the '-fine' option. In order to identify positions where the bases in a read differed from those in the reference sequence, each read was aligned to its reference sequence with the dpAlign module in the BioPerl package (http://www.bioperl.org/). The 454 pyrosequencing-related errors were categorized as insertions, deletions, mismatches, or ambiguous base calls. When aligning sequences, insertions/deletions are allocated based on the sequence context and strand orientation. To eliminate alignment artefacts due to insertion/deletion positions, the lower strand reads were converted to the reverse complement sequence, i.e. keeping the same strandness as the upper strand reads, when aligned with the reference sequence. A sequence error was defined as discordance in an equivalent position between the reference and control (from the 49 healthy individuals and a cloned plasmid vector). The error rate for a specified category was defined as the number of errors divided by the total number of bases in a read. The error rates of a base position on each strand were calculated from 50 control samples.

2.4. *Confirmation of somatic mosaicism of the* NLRP3 *gene by subcloning and subsequent capillary DNA sequencing*

To confirm the somatic mutational frequency that was identified based on the 454 sequencing data, we subcloned the PCR products and performed capillary DNA sequencing as previously described.[9] A Topo TA cloning kit (Invitrogen, San Diego, CA, USA) was used to subclone each of the 14 amplicons.

2.5. *Functional analysis*

To determine whether the identified *NLRP3* mutants are disease-causing, we assessed both ASC [apoptosis-associated speck-like protein containing a caspase recruitment domain; PYCARD, an approved symbol from the HUGO Gene Nomenclature Committee (HGNC) database]-dependent NF-κB activation in HEK293FT cells and transfection-induced cell death in THP-1 cells, a human monocytic cell line, as previously described.[9,13,15] cDNAs encoding carboxy-terminal green fluorescent protein (GFP)-tagged NLRP3 and its mutants were subcloned into pcDNA5/TO (Invitrogen). Before being introduced into THP-1 cells (10^6) using a Cell Line Nucleofector Kit V (Amaxa Biosystems, Cologne, Germany), phorbol myristate acetate (10 ng/ml) was added to enhance transient expression of *NLRP3* gene with minimizing spontaneous cell death.[15] Four hours after the introduction of plasmids (0.5 μg), cell death of GFP-positive THP-1 cells was measured by flow cytometry.

Expression plasmids for NLRP3 and ASC in the pEF-BOS vector background have been previously described.[13] HEK293FT cells (10^5) were transfected using TransIT-293 Transfection Reagent (Milus Bio, Madison, WI, USA) with an NF-κB reporter construct (pNF-κB-luc; 20 ng; BD Biosciences Clontech, Palo Alto, CA, USA), an internal control construct (pRL-TK; 5 ng; Toyo Ink, Tokyo, Japan), and wild-type or mutant NLRP3 expression plasmid (20 ng) in the presence or absence of ASC expression plasmid (20 ng). The amounts of total plasmid DNA used for transfection experiments were kept constant by adding pEF-BOS vector DNA. Twenty-four hours later, the transfected cells were harvested and subjected to dual luciferase assay by which the ability of each construct to induce NF-κB activation was assessed as previously described.[9]

3. Results

3.1. *Construction of base- and strand-specific error rate maps of* NLRP3 *exons from the MPS data of 50 control samples*

Errors in sequence reads generated by a Roche 454 GS-FLX sequencer are not randomly distributed along the sequence and depend on various factors.[16] Although this is a well-known characteristic of 454 sequencing, the occurrence pattern of these errors has not been explored in detail simply because these sequencing errors are considered noise that can be filtered out in most cases. However, it is highly critical to understand the occurrence pattern of sequencing errors on the MPS platform because low-level somatic mosaicism might appear at a rate close to that of sequencing errors. To address this, we collected

~1 million sequence reads using the 454 GS-FLX sequencer for 14 amplicons of *NLRP3* exons from 50 control samples that were thought to be free from somatic mosaicism, and ~94% of those reads were mapped to one of the reference *NLRP3* exon sequences. The number of sequencing depths for each amplicon of each sample on each strand was between 65 and 2139 (mean = 565.3, Supplementary Table S2). We found that the average error rate for each mutation category (insertion, deletion, mismatch, and ambiguous base calls) at each base position on each strand of the amplicons in the control samples was 0.22, 0.16, 0.036, and 0.014%, respectively (Fig. 1C). These values were consistent with those reported in a recent study on the error rates with 454 sequencing data.[16] The sequencing error in the 454 GS-FLX system tends to occur at the beginning and end of the reads,[11,16] and we confirmed this trend in our amplicon sequencing data (Supplementary Fig. S1). Moreover, after removing the end regions of the read sequences, we found that the error rates of the target regions for each category were 0.20, 0.134, 0.023, and 0.014%, respectively (Fig. 1C and Supplementary Table S3). When generating the amplicon sequences for the *NLRP3* exons, the target sequence (CDS region and flanking intron in 10-bp length) was designed to be 300−400 bp and not adjacent to primer sequences in order to obtain relatively low sequencing error rates (Fig. 1C). However, when the base- and strand-specific error rates of the respective amplicons were compared, we noticed that there were large variations in the error rate among amplicons in a strand-specific manner (Fig. 1D). We further examined the occurrence pattern of sequencing errors, as shown in Fig. 1E; the average sequencing error rates at each base in the 50 control amplicons are shown in a bar graph, where the bars in the upper or lower direction show the sequence error rates at the base position on the upper or lower strand of the amplicons, respectively. As evident in Fig. 1E, the error rates at most residues were low (<1%) with some hotspots for each type of error. Most of the insertion/deletion errors preferentially occurred at a homonucleotide region (yellow regions in Fig. 1E) as previously described,[17] but it was not always the case for all of homonucleotide regions. We could not find any tight relationship between other sequence patterns and the error rate. In addition, there was almost no position where sequencing errors occurred at a similar rate on both strands. This is more clearly shown in Fig. 1F, which indicates the numbers of positions with sequence variations in both strands that were higher than the threshold along the horizontal axis. These results indicate that the sequence errors can be discriminated from real genetic alterations when the sequence is read in both directions. However, it is important to keep in mind that PCR errors are not distinct from real genetic alterations. We did not observe any base substitution at a rate higher than 1% in our experiments (Fig. 1F), and the overall PCR error rate under MPS conditions was lower than 1% as long as a high-fidelity DNA polymerase was used to generate the amplicons.

Because Gilles *et al.*[16] recently reported that the occurrence of sequencing errors using the Roche 454 GS-FLX DNA sequencer depends on various factors, we first examined variations in the sequencing error rates of *NLRP3* exons among samples in the same run. For each mutation category, we found a similar trend in the error distribution rate in the amplicon sequences among the control samples (Supplementary Figs S2−S4). We confirmed that, for almost all residues, the error rate distributions among the 50 control samples fitted a Poisson distribution (data not shown). We next examined the run-to-run variation of the sequencing error rate for *NLRP3* exons. For this purpose, we performed an additional MPS run with seven amplicons (exons 3, 4, and 6) that were newly prepared and compared the number and positions of the sequencing errors between two independent sequencing runs. Out of 1993 base positions in the target regions, there was a low occurrence rate of mismatch errors in both runs and this seemed to fit a Poisson distribution. However, insertion/deletion errors (>1% error rate) were observed at ~100 base positions (<5% in the target regions) in each run, and only a half of these errors were shared between both runs (Table 1).

Table 1. Run-to-run variations in the error occurrence (>1% frequency)

Error category	Upper strand			Lower strand			All[a]
	First run	Second run	Overlap	First run	Second run	Overlap	
Insertions	63	73	42	76	96	52	10
Deletions	36	44	24	29	65	20	2
Mismatches	0	0	0	3	0	0	0
Ambiguous base calls	6	8	6	12	10	10	0

[a]The number of positions where the error rates in each category were commonly >1% for both strands in two independent runs.

This indicated that the occurrence of insertion/deletion errors was considerably affected by the run conditions (probably due to variations in the absolute signal strengths of pyrosequencing). Thus, as previously reported, the detection of insertion/deletion mutations by MPS on the 454 GS-FLX system was quite error-prone at least at a limited number of residues. However, the results also implied that false-positive mosaic mutations could be avoided by considering the sequencing data for both strands because these run-dependent insertion/deletion errors occur only in a single strand. Taken together, we conclude that the obtained sequence error map is stable and sufficiently robust to discriminate substitution sequencing errors from low-level mosaicism.

3.2. *Discrimination formula for detection of somatic mosaicism with statistical confidence*

We next examined known SNPs, known heterozygous mutations and somatic mosaic mutations of CAPS patients using MPS. All of these variations appeared on both strands at the expected allele frequencies as shown in Fig. 2, again indicating that filtering the strand-specific sequence variations is unlikely to eliminate real genetic variations.

Based on the experimentally observed sequencing errors with the 454 GS-FLX system described above, we established a discrimination formula to detect low-level somatic mosaicism as follows. In previous studies, the number of reads with the sequence error of a certain category in a sequence position was modelled based on the Poisson distribution with two parameters λ and k where the expected number of reads containing an error and the observed number of reads containing a sequence alteration, respectively, are as shown below[18]:

$$\mathrm{Pois}(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}. \qquad (1)$$

This model assumes that the error rate is constant across the different sequence regions but our data described above pointed out that the sequence error rate varies with the sequence content.[19] Thus, we introduced a position- and strand-specific error rate $q_{i,j,d}$ for a certain error category j in amplicon position i with strand d based on the sequencing data from 50 control samples. With the error rate $q_{i,j,d}$, the upper probability (P) that the number of reads (R) with a certain sequence alteration of category j in position i is equal or greater than the number of observed reads r out of N reads with a sequenced direction d for an unknown sample was defined as:

$$P(R \geq r_{i,j,d} | \lambda_{i,j,d}) = 1 - \sum_{k=0}^{r-1} \frac{\lambda_{i,j,d}^k e^{-\lambda_{i,j,d}}}{k!}, \qquad (2)$$

where, $\lambda_{i,j,d} = N_{i,d} \times q_{i,j,d}$.

For the mismatch error rate, we did not consider the type of base substituted in an amplicon position in this study. We took $(1 - P)$ as a measure of the statistical confidence of the data and conventionally set a threshold of the statistical confidence to be 99.9%. In other words, if P-value was <0.001, the sequence alteration was considered to be a real sequence variation, not an error. For the final identification of real genetic variation with low-level somatic mosaicism, we determined that both of the P-values for the ith residue in the upper and lower strands must be smaller than the threshold.

To evaluate the lower detection limit for the allele frequencies of somatic mosaicism based on the statistical formulation shown above, we generated a series of known allele frequencies by diluting DNA from CAPS patients carrying heterozygous *NLRP3* mutations (c.1043C>T, c.1316C>T, and c.1985T>C) with DNA from normal donors carrying the wild-type *NLRP3* gene. In the dilution series, the mutant allele frequencies were adjusted to be 10, 5, 3, 2, 1, and 0.5% (Table 2). The data indicated that somatic mosaicism at these sites and at an allele frequency $\geq$1% could be convincingly detected with statistical significance ($P < 0.001$) if more than 350 reads for each strand were obtained for an amplicon. We also applied this statistical method to detect somatic mosaicism in patients with known low-level mosaic mutations described above and confirmed that all of

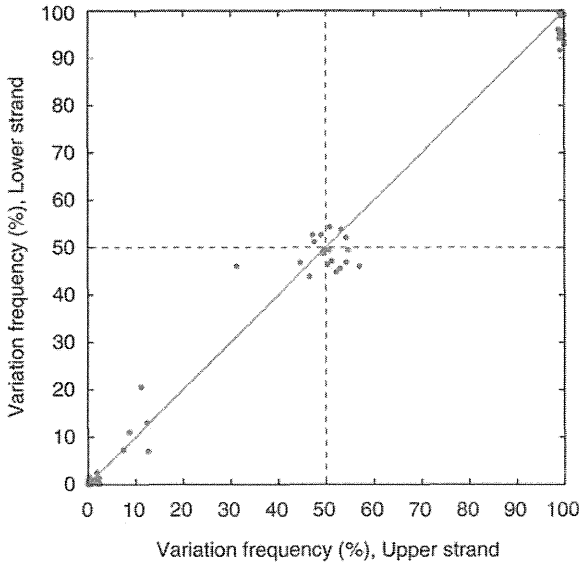

Figure 2. Scatter plot of the observed frequency variation in both strands. The colours depict known SNPs (green), heterozygous and mosaic mutations (orange) and errors (grey).

Table 2. Evaluation of the lower detection limit for mosaicism with three sets of dilution series

Mutation	Dilution (%)	Upper strand				Lower strand			
		Total reads	Mutant reads	%Mutant	*P*-value	Total reads	Mutant reads	%Mutant	*P*-value
c.1043C>T; p.Thr348Met	10.0	724	61	8.43	8.62E−130	520	57	10.96	1.73E−117
	5.0	453	24	5.30	2.86E−47	372	15	4.03	1.26E−25
	3.0	876	27	3.08	1.16E−46	757	21	2.77	6.83E−32
	2.0	737	10	1.36	1.05E−14	645	7	1.09	8.68E−09
	1.0	715	9	1.26	4.73E−13	624	4	0.64	1.11E−04
	0.5	1025	7	0.68	1.15E−14	756	3	0.40	3.22E−03[a]
c.1431C>A; p.Asn477Lys	10.0	542	65	11.99	1.22E−113	346	24	6.94	6.84E−49
	5.0	491	30	6.11	1.13E−44	356	17	4.78	2.42E−32
	3.0	487	21	4.31	1.26E−28	374	19	5.08	1.78E−36
	2.0	577	18	3.12	2.78E−22	495	9	1.82	4.57E−14
	1.0	491	4	0.82	9.17E−04	354	5	1.41	7.34E−08
	0.5	483	0	0	NA	424	3	0.71	NA
c.1985T>C; p.Met662Thr	10.0	658	79	12.01	1.13E−179	643	74	11.51	4.64E−167
	5.0	643	31	4.82	2.56E−59	608	33	5.43	9.96E−65
	3.0	777	27	3.48	4.65E−48	704	29	4.12	1.26E−53
	2.0	929	21	2.26	7.59E−34	835	15	1.80	3.92E−23
	1.0	735	17	1.09	2.74E−11	709	9	1.27	4.06E−13
	0.5	702	2	0.29	3.90E−03[a]	590	1	0.17	1.37E−01[a]

[a]Not significant.

Table 3. Potential mosaic mutations detected in patients with unknown mutations

Patient ID	Amplicon #	Variation		% Variation frequency		*P*-value		dbSNP	State
				Forward	Reverse	Forward	Reverse		
P1	Exon3_2	c.907G>C	p.Asp303His	7.12	11.56	3.0E−44	1.7E−84	rs121908153	Known
P2	Exon3_5	c.1699G>A	p.Glu567Lys	5.94	5.79	2.0E−69	8.9E−47	—	Known
P3	Exon3_5	c.1699G>A	p.Glu567Lys	18.28	15.33	0.0E+00	1.0E−312	—	Known
P4	Exon3_2	c.906C>A	p.Phe302Leu	9.78	9.70	1.7E−86	2.2E−122	—	Novel

the mutations could be detected with statistical significance without any false positives (data not shown).

3.3. *Detection and characterization of* NLRP3 *somatic mosaicism using the MPS platform*

To demonstrate the power of this approach in practice, we applied our new pipeline for 10 CINCA/NOMID patients in whom we failed to detect mutations in the *NLRP3* gene using a conventional direct DNA sequencing approach. The mutations detected by the analysis formulated using the MPS platform in this study are listed in Table 3. We successfully identified four out of the 10 patients with *NLRP3* somatic mosaicism, which was confirmed by subcloning and Sanger sequencing. The nucleotide substitutions were as follows (parentheses indicate the corresponding amino acid change): c.907G>C (p.Asp303His), c.1699G>A (p.Glu567Lys) in two patients, and c.906C>A (p.Phe302Leu). The frequencies of mosaicism identified in these patients by the MPS approach were consistent with those that were identified by the subcloning and subsequent capillary DNA sequencing method (data not shown). Both c.907G>C and c.1699G>A variants were reported as CINCA/NOMID-associated mutations in Infevers database (http://fmf.igh.cnrs.fr/ISSAID/infevers/) and in the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/).[8]

Because the NLRP3 p.Phe302Leu mutation was novel and not detected in the 50 healthy controls, we performed an *in vitro* functional analysis to see the effect of p.Phe302Leu on the protein function. We used two different *in vitro* transfection experiments,

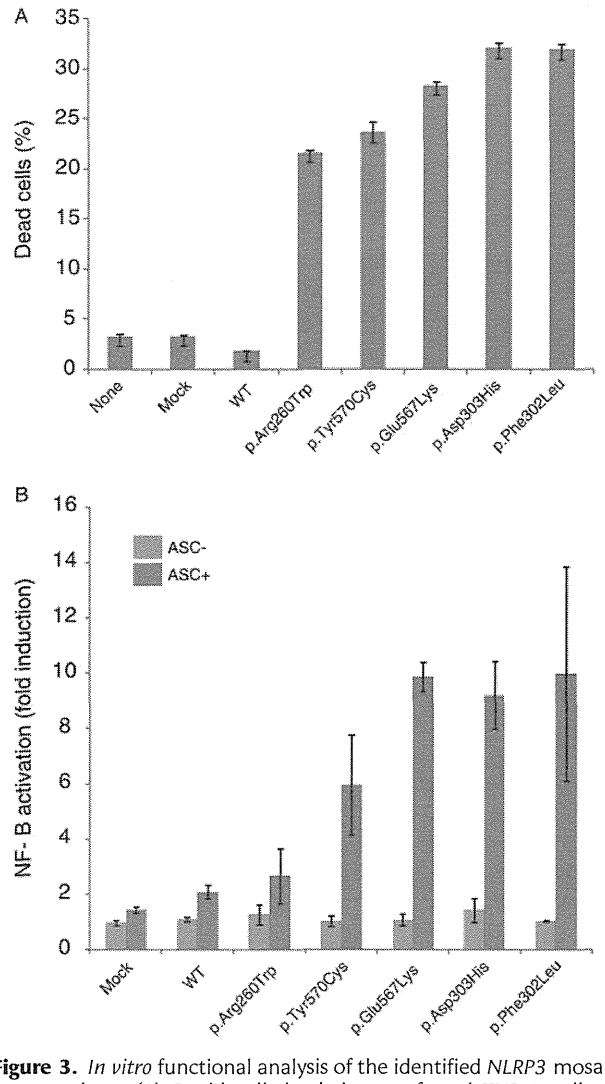

Figure 3. *In vitro* functional analysis of the identified *NLRP3* mosaic mutations. (A) Rapid cell death in transfected THP-1 cells. A GFP-fused wild-type or mutant *NLRP3* was transfected into THP-1 cells and incubated with PMA (10 ng/ml) for 4 h. The percentage of dead cells (7-amino-Actinomycin D [7-AAD]-positive) among the GFP-positive cells is shown. Data represent the means ± SD of triplicate experiments and are representative of two independent experiments. The data for previously reported mutations as well as the mutations found in this study are shown. (B) ACS-dependent NF-κB activation in transfected HEK293FT cells. HEK293FT cells were co-transfected with wild-type or mutant *NLRP3* in the presence or absence of ASC. NF-κB induction is shown as the fold-change compared with cells that were transfected with a control vector without ASC (set equal to one). Values are the means ± SD of triplicate experiments, and the data are representative of three independent experiments. The data for previously reported mutations (p.Arg260Trp and p.Tyr570Cys) and the mutations found in this study are shown. For each mutation, the data obtained in the presence and absence of ASC are shown. These findings identified p.Phe302Leu as a novel disease-causing mutation.

the rapid cell death in transfected THP-1 cells and the ASC-dependent NF-κB activation in transfected HEK293FT cells (Fig. 3A and B, respectively). Both assays clearly showed that p.Phe302Leu was a disease-causing mutation similar to known CINCA/NOMID-associated pathogenic mutations (p.Asp303His and p.Glu567Lys).[9]

4. Discussion

Although the somatic mutation rate at the nucleotide level *in vivo* was difficult to quantitatively measure due to the complexity of the genome and laborious molecular detection processes, recent advances in MPS technologies have allowed us to directly quantitate somatic mutations in human genome.[20–22] The current estimate for the somatic (*de novo*) mutation rate is $1-2 \times 10^{-8}$ residues/generation/haploid, and this estimate is sufficiently low that we would expect to never observe somatic mosaicism in the *NLRP3* gene by chance; although the error rate of the high-fidelity DNA polymerase used to produce the amplicons is two orders of magnitude larger than the somatic mutation rate,[23,24] we could not detect PCR-generated mosaicism higher than 1% in the 454 sequencing error maps. Based on the literature, the single base substitutions are the most frequent type of somatic mutations (~500 times more frequent than short insertions/deletions)[25] and protein-coding sequences are less mutagenic than sequences in non-coding regions, assuming that the somatic mutation spectrum in malignant cells is the same as in normal cells. Somatic mosaicism is thought to result from *de novo* gain-of-function-type mutations that are introduced at a very early and limited stage of development, and it is reasonable to focus our efforts on detecting base substitutions for somatic mosaicism in the *NLRP3* gene.

It is challenging but highly important in many areas of research, such as cancer, to detect low-level somatic mutations, which we designated as somatic mosaicism in this study, from apparently mutation-negative samples by conventional sequencing. Subcloning followed by the capillary DNA sequencing has been a *de facto* standard to identify somatic mosaicism, but this is not the method of choice for routine diagnostics because it is laborious, time consuming, and costly. Thus, it is reasonable for us to explore MPS as a new tool for this purpose. Although previous studies have used MPS technology to detect somatic mosaicism, it was unclear how sensitive this method is to detect a low-level somatic mosaicism using the MPS platform because this platform is generally error-prone. To address this challenge, we developed a new pipeline to detect low-level somatic mosaicism with statistical confidence using base position- and strand-specific error rate maps for the *NLRP3* amplicons to be studied. Whereas the

detection limit of somatic mosaicism depends on the base position and the read depth of the amplicons, the limit of detection could be as low as 1% allele frequency with no false positives for substitutions (the precision is higher than 99.9%). Our error map shows that 98.1% of base positions (3343 out of 3407 target positions) in the *NLRP3* exonic amplicons can be detected with ~1% mosaicism when more than ~350 reads were accumulated for each strand. Although the remaining region (64 base positions out of 3407 target positions) was too error-prone (the error rate ranged from 0.1 to 1.7% in either the upper or lower strand) to detect low-level mosaicism by MPS, medium-level mosaicism (5% or high) could be identified in all base positions in the target region with the same significance level. Based on this pipeline, we successfully identified four cases of somatic mosaicism among 10 apparently mutation-negative CINCA/NOMID patients. These results were subsequently confirmed by functional analysis and subcloning followed by capillary DNA sequencing method.

As described above, we revealed that a read depth of ~350 for each strand of each amplicon would be sufficient to detect somatic mosaicism as low as 1% with statistical confidence. This means that an analysis of somatic mosaicism (detection limit of 1% allele frequency) of the *NLRP3* gene for one sample requires $350 \times 2 \times 14 = 9800$ reads with the 454 GS-FLX sequencer, which has a capacity to obtain 1 000 000 reads per run. Thus, we could analyse ~100 patient samples with a single run (~10 h) using this MPS platform. For this purpose, a miniaturized 454 sequencer might be more convenient because it could analyse 10 patient samples at once with a reasonably reduced running cost.

The approach used to detect somatic mosaicism is very similar to that for low-frequency alleles in pooled DNA samples, for which MPS applications have been reported by many groups.[18,26,27] However, the main aim of these previous studies was to screen for a rare allele in a population. Thus, the discovery phase on the MPS platform must be followed by an evaluation phase using conventional methods. Therefore, when diagnosing somatic mosaicism of the *NLRP3* gene based solely on the MPS platform, we could not use the same approach to detect rare alleles in a population due to its low accuracy. The sequencing error rate on the Roche MPS platform was sufficiently stable and low enough as shown in this study. Using our pipeline, we were able to detect 1% somatic mosaicism in the *NLRP3* gene with 99.9% confidence. Although another research group recently used a similar approach with a short-read MPS,[28] the Roche long-read MPS is more suitable as a diagnostic tool mainly because of the short run time. If we could diagnose somatic mosaicism of the *NLRP3* gene within a reasonable time with low labour and costs as shown in this study, the success rate of CINCA/NOMID genetic diagnosis will increase from 60 to 80% or higher,[9] which will greatly advance the health and care of these patients and prevent irreversible bone and neurological complications of disease.

This pipeline would also be efficient to detect somatic mosaicism in mutation-negative patients with other diseases, including cancer. The error rate map for a given gene should be constructed from authentic plasmids, and used to detect somatic mosaicism of other genes as well as rare alleles in various populations.

Supplementary data: Supplementary data are available at www.dnaresearch.oxfordjournals.org.

Funding

This study was supported by the Japanese Ministry of Education, Science, Sports, and Culture, and the Japanese Ministry of Health, Labor, and Welfare.

Acknowledgements: We thank all patients who participated in the study. We are grateful to Ms. Yuki Takaoka at the Department of Pediatrics, Kyoto University Graduate School of Medicine and Mr. Takashi Watanabe at the Department of Human Genome Research, Kazusa DNA Research Institute for their technical assistance.

References

1. Prieur, A.M. and Griscelli, C. 1981, Arthropathy with rash, chronic meningitis, eye lesions, and mental retardation, *J. Pediatr.*, **99**, 79–83.
2. Hassink, S.G. and Goldsmith, D.P. 1983, Neonatal onset multisystem inflammatory disease, *Arthritis Rheum.*, **26**, 668–73.
3. Torbiak, R.P., Dent, P.B. and Cockshott, W.P. 1989, NOMID—a neonatal syndrome of multisystem inflammation, *Skeletal Radiol.*, **18**, 359–64.
4. Feldmann, J., Prieur, A.M., Quartier, P., et al. 2002, Chronic infantile neurological cutaneous and articular syndrome is caused by mutations in CIAS1, a gene highly expressed in polymorphonuclear cells and chondrocytes, *Am. J. Hum. Genet.* United States, 198–203.
5. Aksentijevich, I., Nowak, M., Mallah, M., et al. 2002, De novo CIAS1 mutations, cytokine activation, and evidence for genetic heterogeneity in patients with neonatal-onset multisystem inflammatory disease (NOMID)—a new member of the expanding family of pyrin-associated autoinflammatory diseases, *Arthritis Rheum.*, **46**, 3340–8.

6. Hoffman, H.M., Mueller, J.L., Broide, D.H., Wanderer, A.A. and Kolodner, R.D. 2001, Mutation of a new gene encoding a putative pyrin-like protein causes familial cold autoinflammatory syndrome and Muckle-Wells syndrome, *Nat. Genet.*, **29**, 301–5.
7. Goldbach-Mansky, R. 2011, Current status of understanding the pathogenesis and management of patients with NOMID/CINCA, *Curr. Rheumatol. Rep.*, **13**, 123–31.
8. Milhavet, F., Cuisset, L., Hoffman, H.M., et al. 2008, The infevers autoinflammatory mutation online registry: Update with new genes and functions, *Hum. Mutat.*, **29**, 803–8.
9. Tanaka, N., Izawa, K., Saito, M.K., et al. 2011, High incidence of NLRP3 somatic mosaicism in patients with chronic infantile neurologic, cutaneous, articular syndrome: results of an International Multicenter Collaborative Study, *Arthritis Rheum.*, **63**, 3625–32.
10. Qin, W., Kozlowski, P., Taillon, B.E., et al. 2010, Ultra deep sequencing detects a low rate of mosaic mutations in tuberous sclerosis complex, *Hum. Genet.*, **127**, 573–82.
11. Campbell, P.J., Pleasance, E.D., Stephens, P.J., et al. 2008, Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing, *Proc. Natl. Acad. Sci. USA*, **105**, 13081–6.
12. Rohlin, A., Wernersson, J., Engwall, Y., Wiklund, L., Bjoerk, J. and Nordling, M. 2009, Parallel sequencing used in detection of mosaic mutations: Comparison with four diagnostic DNA screening techniques, *Hum. Mutat.*, **30**, 1012–20.
13. Saito, M., Nishikomori, R., Kambe, N., et al. 2008, Disease-associated CIAS1 mutations induce monocyte death, revealing low-level mosaicism in mutation-negative cryopyrin-associated periodic syndrome patients, *Blood*, **111**, 2132–41.
14. Kent, W.J. 2002, BLAT—the BLAST-like alignment tool, *Genome Res.*, **12**, 656–64.
15. Fujisawa, A., Kambe, N., Saito, M., et al. 2007, Disease-associated mutations in CIAS1 induce cathepsin B-dependent rapid cell death of human THP-1 monocytic cells, *Blood*, **109**, 2903–11.
16. Gilles, A., Meglecz, E., Pech, N., Ferreira, S., Malausa, T. and Martin, J.F. 2011, Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing, *BMC Genomics*, **12**, 245.
17. Margulies, M., Egholm, M., Altman, W.E., et al. 2005, Genome sequencing in microfabricated high-density picolitre reactors, *Nature*, **437**, 376–80.
18. Altmann, A., Weber, P., Quast, C., Rex-Haffner, M., Binder, E.B. and Muller-Myhsok, B. 2011, vipR: variant identification in pooled DNA using R, *Bioinformatics*, **27**, i77–84.
19. Dohm, J.C., Lottaz, C., Borodina, T. and Himmelbauer, H. 2008, Substantial biases in ultra-short read data sets from high-throughput DNA sequencing, *Nucleic Acids Res.*, **36**, e105.
20. Lee, W., Jiang, Z., Liu, J., et al. 2010, The mutation spectrum revealed by paired genome sequences from a lung cancer patient, *Nature*, **465**, 473–7.
21. Awadalla, P., Gauthier, J., Myers, R.A., et al. 2010, Direct Measure of the de novo mutation rate in Autism and Schizophrenia Cohorts, *Am. J. Hum. Genet.*, **87**, 316–24.
22. Conrad, D.F., Keebler, J.E.M., DePristo, M.A., et al. 2011, Variation in genome-wide mutation rates within and between human families, *Nat. Genet.*, **43**, 712–4.
23. Cha, R.S. and Thilly, W.G. 1993, Specificity, efficiency, and fidelity of PCR, *PCR Methods Appl.*, **3**, S18–29.
24. Vandenbroucke, I., Marck, H.V., Verhasselt, P., et al. 2011, Minor variant detection in amplicons using 454 massive parallel pyrosequencing: experiences and considerations for successful applications, *BioTechniques*, **51**, 167–77.
25. Pleasance, E.D., Cheetham, R.K., Stephens, P.J., et al. 2010, A comprehensive catalogue of somatic mutations from a human cancer genome, *Nature*, **463**, 191–6.
26. Fakhrai-Rad, H., Zheng, J.B., Willis, T.D., et al. 2004, SNP discovery in pooled samples with mismatch repair detection, *Genome Res.*, **14**, 1404–12.
27. Bansal, V. 2010, A statistical method for the detection of variants from next-generation resequencing of DNA pools, *Bioinformatics*, **26**, i318–24.
28. Flaherty, P., Natsoulis, G., Muralidharan, O., et al. 2012, Ultrasensitive detection of rare mutations using next-generation targeted resequencing, *Nucleic Acids Res.*, **40**, e2.