OXFORD UNIVERSITY PRESS | International Immunology

Autosomal recessive complete STAT1 deficiency caused by compound heterozygous intronic mutations

Journal:	*International Immunology*
Manuscript ID	INTIMM-20-0079.R1
Manuscript Type:	Original Research
Date Submitted by the Author:	17-Jun-2020
Complete List of Authors:	Sakata, Sonoko; Hiroshima University Graduate School of Biomedical and Health Sciences Health Sciences Major, Department of Pediatrics Tsumura, Miyuki; Hiroshima University Graduate School of Biomedical & Health Sciences, Department of Pediatrics Matsubayashi, Tadashi; Department of Pediatrics, Seirei Hamamatsu General Hospital, Shizuoka, Japan Karakawa, Shuhei; Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, Japan Kimura, Shunsuke; Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, Japan; Department of Pathology, Hematological Malignancies Program, St. Jude Children's Research Hospital, Memphis, TN, USA Tamaura, Moe; Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, Japan, Pediatrics Okano, Tsubasa; Department of Pediatrics and Developmental Biology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan Naruto, Takuya; Department of Pediatrics and Developmental Biology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan Mizoguchi, Yoko; Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, Japan Kagawa, Reiko; Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, Japan Nishimura, Shiho; Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, Japan Imai, Kohsuke; Department of Pediatrics and Developmental Biology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan Le Voyer, Tom; Laboratory of Human Genetics of Infectious Diseases, Necker Branch, INSERM U1163, Necker Hospital for Sick Children, 75015 Paris, France Casanova, Jean-Laurent; St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, The Rockefeller University, New York, NY 10065, USA Bustamante, Jacinta; Laboratory of Human Genetics of Infectious Diseases, Necker Branch, INSERM U1163, Necker Hospital for Sick Children, 75015 Paris, France Morio, Tomohiro; Department of Pediatrics and Developmental Biology,

Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan, Pediatrics and Developmental Biology
Ohara, Osamu; Department of Applied Genomics, Kazusa DNA Research Institute, Kisarazu, Japan
Kobayashi, Masao; Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, Japan
Okada, Satoshi; Hiroshima University Graduate School of Biomedical & Health Sciences, Department of Pediatrics

Keywords: mycobacteria, virus, target RNA sequence, primary immunodeficiency

Title: Autosomal recessive complete STAT1 deficiency caused by compound heterozygous intronic mutations

Authors

Sonoko Sakata[1], Miyuki Tsumura[1], Tadashi Matsubayashi[2], Shuhei Karakawa[1], Shunsuke Kimura[1,3Φ], Moe Tamaura[1], Tsubasa Okano[4], Takuya Naruto[4], Yoko Mizoguchi[1], Reiko Kagawa[1], Shiho Nishimura[1], Kohsuke Imai[4], Tom Le Voyer[5,6], Jean-Laurent Casanova[5,6,7,8,9,10], Jacinta Bustamante[5,6,9,11], Tomohiro Morio[4], Osamu Ohara[12], Masao Kobayashi[1, 13Φ], Satoshi Okada[1]

Corresponding Author

Satoshi Okada, MD, PhD

Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences

1-2-3 Kasumi, Minami-Ku, Hiroshima-Shi, Hiroshima, 734-8551, Japan

Tell: +81-82-257-5212

Fax: +81-82-257-5214

E-mail: sokada@hiroshima-u.ac.jp

Running title: Intron mutations cause STAT1 deficiency

Institutions

1. Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, Japan

2. Department of Pediatrics, Seirei Hamamatsu General Hospital, Shizuoka, Japan

3. Department of Pathology, Hematological Malignancies Program, St. Jude Children's Research Hospital, Memphis, TN, USA

4. Department of Pediatrics and Developmental Biology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan

5. Laboratory of Human Genetics of Infectious Diseases, Necker Branch, INSERM U1163, Necker Hospital for Sick Children, 75015 Paris, France.

6. Paris University, Imagine Institute, Paris, France, EU

7. St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, The Rockefeller University, New York, NY 10065, USA

8. Pediatric Hematology-Immunology Unit, Necker Hospital for Sick Children, Paris, France, EU

9. Study Center of Immunodeficiencies, Necker Hospital for Sick Children, Paris, France, EU

10. Howard Hughes Medical Institute, New York, NY, USA

11. Department of Clinical Immunology, Aarhus University Hospital, Aarhus N, Denmark, EU

12. Department of Applied Genomics, Kazusa DNA Research Institute, Kisarazu, Japan

13. Japan Red Cross, Chugoku-Shikoku Block Blood Center, Hiroshima, Japan

[Φ]current affiliation

Number of words: 3,578 words

Keywords: mycobacteria, virus, target RNA sequence, primary immunodeficiency

Abstract

Autosomal recessive (AR) complete signal transducer and activator of transcription 1 (STAT1) deficiency is an extremely rare primary immunodeficiency that causes life-threatening mycobacterial and viral infections. Only seven patients from five unrelated families with this disorder have been so far reported. All causal *STAT1* mutations reported are exonic and homozygous. We studied a patient with susceptibility to mycobacteria and virus infections, resulting in identification of AR complete STAT1 deficiency due to compound heterozygous mutations, both located in introns: c.128+2 T>G and c.542-8 A>G. Both mutations were the first intronic *STAT1* mutations to cause AR complete STAT1 deficiency. Targeted RNA-seq documented the impairment of *STAT1* mRNA expression and contributed to the identification of the intronic mutations. The patient's cells showed a lack of STAT1 expression and phosphorylation, and severe impairment of the cellular response to IFN-γ and IFN-α. The case reflects the importance of accurate clinical diagnosis and precise evaluation, to include intronic mutations, in the comprehensive genomic study when the patient lacks molecular pathogenesis. In conclusion, AR complete STAT1 deficiency can be caused by compound heterozygous and intronic mutations. Targeted RNA-seq based systemic gene expression assay may help to increase diagnostic yield in inconclusive cases after comprehensive genomic study.

62

63 **Introduction**

64 Signal transducer and activator of transcription 1 (STAT1) is a latent cytoplasmic transcription
65 factor that has a fundamental role in signal transduction from both type I (IFN-α and IFN-β), type II
66 (IFN-γ) and type III (IFN-λ) interferons and also IL-27.(1) In response to IFN-γ, IFN-α/β, or
67 interleukin 27 (IL-27) stimulation, STAT1 forms a homodimer called gamma-interferon activation
68 factor (GAF). GAF translocates to the nucleus and binds to gamma-activating sequences (GAS) to
69 induce the transcription of target genes involved in antimycobacterial immunity.(1,2) STAT1 also
70 forms a heterotrimer with STAT2, and IRF9, which is known as interferon-stimulated gene factor 3
71 (ISGF3), after stimulation by IFN-α/β. ISGF3 binds the interferon-stimulated response element
72 (ISRE) and induces target genes involved in anti-viral immunity.(1,2)

73 Inborn errors in human STAT1 immunity cause at least four types of primary immunodeficiency:
74 i) autosomal recessive (AR) complete STAT1 deficiency; ii) AR partial STAT1 deficiency; iii)
75 autosomal dominant (AD) STAT1 deficiency; and iv) AD STAT1 gain of function.(3) Among them,
76 AR complete STAT1 deficiency is an extremely rare primary immunodeficiency (PID) that causes
77 life-threatening mycobacterial and viral infections. Indeed, only seven patients from five unrelated
78 families with AR complete STAT1 deficiency have been so far reported.(3-7) Those patients show
79 complete functional impairment of STAT1-dependent response to type I and type II interferons.(4)
80 This is a purely recessive disorder and no haplo-insufficiency at the STAT1 locus has been reported

for any of the known cellular or clinical phenotypes.(3) Prognosis of the patients with AR complete STAT1 deficiency is poor, and hematopoietic stem cell transplantation (HSCT) is the only curative treatment. Three patients received HSCT and long-term survival was achieved in two–patients. Overall, five of the seven patients died before 18 months of age from mycobacterial infections (two patients), viral infections (two patients), or multiorgan failure in the course of HSCT (one patient) (summarized in Table 1).(3,7) Therefore, early diagnosis and appropriate therapeutic intervention are necessary to avoid life-threatening events in this disorder.

Here we report a patient with AR complete STAT1 deficiency due to compound heterozygous mutations, c.128+2 T>G/c.542-8 A>G, in *STAT1*. Both mutations were the first intronic *STAT1* mutations reported to cause AR complete STAT1 deficiency. The case reflects the importance of inclusion of non-coding regions and intronic mutations to obtain an accurate clinical diagnosis and precise evaluation in the comprehensive genomic study when the patient lacks molecular pathogenesis. In addition, targeted RNA sequencing (RNA-seq) based systemic gene expression assay may enhance diagnostic yield in inconclusive cases after comprehensive genomic study.

Material and Methods

Case report

The patient is a 6-year old Japanese boy who was born to non-consanguineous parents (**Figure 1A**).

He had no family history of PID. At the age of 1 month, the patient developed respiratory syncytial virus (RSV) bronchiolitis and was treated with non-invasive positive pressure ventilation. He received a BCG vaccination at 9 months. Five weeks later, the patient developed lymphadenitis in the left axillar region. He then presented with a skin rash and fever. Lymph node biopsy was performed at 11 months. The histopathological finding of the lymph nodes showed no granuloma. *Mycobacterium tuberculosis* was detected by PCR from the lymph nodes and bone marrow. The patient was started treatment with isoniazid (INH), rifampicin (RFP) and ethambutol (EMB) with the suspicion of *Mycobacterial tuberculosis* infection. Twenty days later, the vaccine strain BCG was confirmed as the pathogen by Southern blotting. Therefore, the patient was given a diagnosis of disseminated BCG. Laboratory tests showed leukocytosis (47,800/μL) (Reference range: RR 6,000–17,500) and high levels of C-reactive protein (CRP: 10.8 mg/dL). Serum immunoglobulin levels were normal. The patient displayed normal respiratory burst and normal T-lymphocyte activation after phytohemagglutinin (PHA) or concanavalin A (ConA) stimulation. No obvious abnormality was observed in the T and B cell counts, with normal results for T cell receptor excision circles (TRECs) and K-deleting recombination excision circles (KRECs). Deep and comprehensive phenotyping of immune cell subsets detected a decreased frequency of Th1 cells and myeloid DCs (mDCs), and an increased frequency of Th17 cells (**Supplementary Table 1**). Since initial antimycobacterial agents were clinically effective, the treatments with INH, RFP and EMB were

117 continued after diagnosis of disseminated BCG. Cumulatively, the patient received INH, RFP, and
118 EMB for 40, 28, and 2 months from this episode, respectively. The patient was diagnosed with
119 Mendelian susceptibility to mycobacterial diseases (MSMD) and subjected to trio-exome analysis at
120 the age of 11 months. However, trio-exome analysis failed to confirm the genetic etiology at that
121 time.

122 After that episode, he developed severe and recurrent infections, summarized in **Supplementary**
123 **Table 2**. At the age of 1 year and 4 months, he developed acute asthma associated with influenza A
124 infection. The laboratory tests showed an elevated level of serum CRP (12.0 mg/dL) without the
125 findings of pneumonia on radiological imaging. At the age of 2 years and 1 month, he presented with
126 febrile seizure. The blood examination showed thrombocytopenia (platelets 4.0 $\times 10^9$/L), and
127 elevated levels of serum ferritin (8,404 μg/L, RR 20–250) and soluble IL-2 receptor (8,270 U/ml,
128 RR 157–474). Although no obvious splenomegaly nor evidence of hemophagocytosis in the bone
129 marrow were detected, the patient was suspected of having developed hemophagocytic
130 lymphohistiocytosis (HLH) secondary to infection with an unknown pathogen and was successfully
131 treated with dexamethasone palmitate. At the age of 2 years and 11 months, he developed
132 bronchiolitis with a positive result of *Mycoplasma pneumoniae* antigen and treated with
133 azithromycin. At the age of 3 years and 1 month, he developed bilateral tibial osteomyelitis with an
134 elevated level of serum CRP (29.7 mg/dL). No obvious pathogen was identified from a biopsied

specimen by cultivation, and the osteomyelitis did not respond to antibacterial drugs. It gradually improved with a long clinical course. Although there was no evidence, continuous treatment with antimycobacterial agents might have contributed to the improvement of patient's illness. At the age of 3 years and 6 months, he developed pneumonia due to human metapneumovirus (hMPV). He suffered from dyspnea and showed an elevated level of serum CRP (24.5 mg/dL). At the age of 3 years and 9 months, he developed severe enterocolitis with bilious vomiting and paralytic ileus associated with rotavirus infection. At the age of 3 years and 10 months, he suffered from Kawasaki disease-like symptoms (a combination of fever, rash, swelling of the lips and neck lymph nodes, and conjunctival injection) with elevated atypical lymphocytes and serum CRP (10.2 mg/dL). These symptoms improved spontaneously without intravenous immunoglobulin treatment. At the age of 4 years and 2 months, he presented with pneumonia, which responded to antibacterial treatment. The prophylaxis with INH was suspended at the age of 4 years and 4 months. At the age of 4 years and 8 months, 2 weeks after receiving his varicella vaccination, he developed vaccine-strain induced varicella with a typical rash and fever, that was treated with oral acyclovir. At the age of 4 years and 9 months, he developed life-threatening *M. malmoense* mediastinal lymphadenitis and tibial osteomyelitis (**Figure 1B**). The histopathological findings of mediastinal lymph node biopsy showed no granuloma formation (**Figure 1C**). The symptoms gradually improved after starting RFP, EMB, and clarithromycin (CAM). At the age of 5 years and 4 months, he developed influenza A

153 respiratory infection, which was treated with peramivir (CRP 19.3 mg/dL).

154 The systemic gene expression assay with targeted RNA-seq was performed and identified
155 decreased *STAT1* expression with aberrant splicing. The exome data were then reanalyzed, resulting
156 in identification of compound heterozygous intronic mutations, c.128+2 T>G/c.542-8 A>G, in
157 *STAT1*. The patient was thus given a diagnosis of AR complete STAT1 deficiency.

158

159 **Genomic DNA and whole exome sequencing**

160 Genomic DNA was eluted from whole blood with a QIAamp DNA Mini Kit (Qiagen, Hilden,
161 Germany). Whole exome sequencing (WES) library preparation was performed with SureSelect XT
162 or QXT Reagent Kit (Agilent Technologies, Santa Clara, CA, USA) and SureSelect XT Human All
163 Exon V5 Kit (Agilent Technologies). The library was sequenced using the HiSeq1500 system
164 (Illumina, San Diego, CA, USA), and the variants were annotated as previously described.(8) For the
165 first analysis of exome data, we selected the following variants with a global minor allele frequency
166 (GMAF)>0.05 as candidate variants: variants in the cording sequence excluding synonymous
167 variants, variants located within 5 base points from exon–intron boundaries, or variants reported as
168 probably damaging, possibly damaging, disease_causing_automatic, or disease_causing in dbNSPF
169 (https://sites.google.com/site/jpopgen/dbNSFP), or variants reported in OMIM
170 (https://www.ncbi.nlm.nih.gov/omim) or Clinvar (https://www.ncbi.nlm.nih.gov/clinvar/). As for the

171 second analysis (reanalysis) of the exome data, we filtered the variants as previously described.(8)

172

173 **Targeted RNA-seq and expression analysis**

174 Total RNA was extracted from the PBMCs from the patient and 10 other PID patients without
175 genetic etiology after WES. Libraries for the targeted RNA-seq for a PID panel consist of 426
176 immune related genes, which include PID responsible genes reported from International Union of
177 Immunological Societies in 2017 (IUIS 2017), were prepared using an Agilent SureSelect Strand
178 Specific RNA library construction kit with RNAs prepared by the Trizol method (ambion). RNAs
179 derived from PID genes were enriched by hybridization with PID panel probes using a SureSeclect
180 target enrichment system (Agilent Technologies). The enriched libraries were sequenced on an
181 Illumina MiSeq under a 75-base paired-end run mode and the obtained reads were mapped to the
182 human reference genome (NCBI build 37.1) using STAR.(9,10) For expression analysis, the count
183 data were extracted using Rsubread(11) and normalized using DESeq2.(12) To analyze the effect of
184 splice-site mutations, the splicing pattern was assessed manually with IGV software.(13)

185 The detailed method of quantitative PCR and reverse transcription PCR (RT-PCR) are shown in
186 the **Supplementary Methods**.

187

188 **Flow cytometry**

189 The peripheral blood mononuclear cells (PBMCs) were isolated by density gradient centrifugation.
190 The PBMCs were suspended at a density of 10^4 cells/μl in serum-free RPMI1640. The cells were
191 incubated with IFN-γ (1,000 IU/ml) or IFN-α (1,000 IU/ml) for 15 minutes at 37°C in the presence
192 of FITC-conjugated CD14 (BD Biosciences, Franklin Lakes, NJ, USA). They were then washed in
193 RPMI1640 and were fixed and permeabilized according to the BD Phosflow protocol (Protocol III).
194 They were next stained with FITC-conjugated anti-CD14 and PE-conjugated anti-pSTAT1 (pY701)
195 (BD Biosciences), and subjected to flow cytometric analysis to analyze STAT1 phosphorylation.
196 Data were analyzed with FlowJo software (BD Biosciences).
197

198 **Immunoblot analysis and electrophoretic mobility shift assay (EMSA)**

199 The PBMCs or SV40 fibroblasts were incubated in the presence or absence of IFN-γ (1,000 IU/ml)
200 or IFN-α (1,000 IU/ml) for 15 minutes and subjected to immunoblot analysis. Immunoblot analysis
201 was performed as described previously.(14,15) The following primary antibodies used for
202 immunoblotting: an anti-pSTAT1 (pY701) antibody (Cell Signaling Technology, Danvers, MA,
203 USA), an anti-STAT1α antibody against total STAT1 (Santa Cruz Biotechnology, Santa Cruz, CA,
204 USA), and an anti-β-actin antibody (Sigma-Aldrich, St. Louis, MO, USA). EMSA was conducted as
205 previously described.(15,16) Briefly, the cells were stimulated by incubation for 15 minutes with
206 1,000 IU/ml IFN-γ or 1,000 IU/ml IFN-α. We then incubated 10 μg of nuclear extract with

207 ^{32}P-labeled (αdATP) GAS (from *FCGR1* promoter) or ISRE (from *ISG15* promoter) probes for 30
208 minutes and subjected them to analysis.
209

210 **Ethics Statement**

211 We obtained written informed consent for genomic analysis and blood sample based functional
212 studies of the patient, parents and siblings in accordance with the Declaration of Helsinki. The
213 genetic analysis and blood sample based functional studies were approved by the Institutional
214 Review Board of Hiroshima University and Tokyo Medical and Dental University.
215

216 **Results**

217 **Identification of biallelic variants in *STAT1***

218 High-molecular weight genomic DNA was extracted from peripheral blood. Trio-whole-exome
219 sequencing, which was performed at 11 months of age, identified a novel heterozygous variant,
220 c.128+2 T>G, in the *STAT1* gene in the patient and his asymptomatic mother (**Figure 1A, 1D**). Rare
221 variants in other known PID related genes reported in IUIS 2017 (*LRBA*, *NBN*, *NHEJ1*, and *TBX1*)
222 were also identified (**Supplementary Table 3**).(10) However, they were not inferred to be
223 disease-causing based on the clinical manifestations and their inheritance pattern. The c.128+2 T>G
224 *STAT1* variant was confirmed by Sanger sequencing (**Figure 1D**). Monoallelic dominant negative

225 and loss-of-function mutations in *STAT1*, which are normally expressed at the protein level, have
226 been identified in patients with MSMD.(2,15-20) In contrast, a familial study of patients with AR
227 STAT1 complete deficiency, which investigated parents with monoallelic loss-of-expression
228 mutation in *STAT1*, revealed that there is no haploinsufficiency at the *STAT1* locus.(3) The c.128+2
229 T>G variation identified in this study's patient was located at an essential splice site and was
230 suspected to disturb STAT1 protein expression. Based on a previous study that showed a lack of
231 haploinsufficiency, together with the presence of this variant in the asymptomatic mother, the
232 pathogenicity of the c.128+2 T>G variant was missed at the first analysis of the exome data.
233 A systemic gene expression assay using targeted RNA-seq identified decreased STAT1
234 expression and exon skipping at exon 8 of *STAT1*. We thus thoroughly reanalyzed the exome data
235 and identified a novel heterozygous intronic variant, c.542-8A>G, which was missed by the filtering
236 process of first analysis of the exome data (filtering strategy of the first analysis of exome data is
237 detailed in methods). This variant was confirmed by Sanger sequencing and was identified in the
238 asymptomatic father and elder sister (**Figure 1A, 1D**). Therefore, the patient was determined to have
239 compound heterozygous variations, c.128+2 T>G/c.542-8 A>G, in *STAT1*. Neither variant was
240 found in the Single Nucleotide Polymorphism Database (dbSNP), 1000 Genome Projects, the Exome
241 Aggregation Consortium (ExAc) database, or the genome aggregation database (gnomAD).
242

243 **Targeted RNA-seq and qPCR**

244 A targeted RNA-seq based systemic gene expression assay of PBMCs was implemented to
245 investigate the molecular pathogenesis of the patient. This assay ranked *STAT1* as among the top five
246 genes with reduced expression in the patients when we used the other 10 inconclusive PID patients
247 as controls (**Figure 2**). As expected, the splicing pattern assessed manually with IGV software
248 revealed intron retention at exon 3 associated with the c.128+2 T>G variant (**Figure 3A**).
249 Furthermore, we identified abnormal splicing in the form of exon skipping at exon 8 of the *STAT1*
250 gene (**Figure 3B**). We next performed a RT-PCR assay for STAT1 from PBMCs of the patient and
251 two healthy individuals. The RT-PCR assay, which spanned exon 3 and exon 7 of *STAT1*, confirmed
252 the reduced expression of *STAT1* mRNA and the presence of intron retention by detecting
253 approximate 1,350 bp band in the patient (**Supplementary Figure S1**). The RT-PCR assay, which
254 spanned exon 6 and exon 10 of *STAT1*, also confirmed the presence of exon skipping at exon 8 by
255 detecting approximate 300 bp band in the patient. To confirm the result of the targeted RNA-seq and
256 evaluate the impact of the biallelic *STAT1* variations on mRNA synthesis, we carried out qPCR from
257 PBMCs of the patient and a healthy control. This confirmed a severe decrease of *STAT1* mRNA in
258 the patient's cells (9.4% of *STAT1* mRNA compared with the control's) (**Figure 4A**).

259

260 **Impaired STAT1 protein expression and phosphorylation**

261 We performed flow cytometry to analyze the STAT1 function by detecting its phosphorylation
262 (pSTAT1) upon IFN-γ or IFN-α stimulation. The $CD14^+$ monocytes from a healthy control showed
263 pSTAT1 upon IFN-γ or IFN-α (**Figure 4B**). In contrast, the $CD14^+$ monocytes from the patient
264 completely lacked pSTAT1 in response to IFN-γ or IFN-α. To confirm this finding, PBMCs from the
265 patient and a healthy control were stimulated with IFN-γ or IFN-α and subjected to immunoblot
266 analysis. As shown in Figure 4C, PBMCs from the patient showed a complete lack of STAT1
267 protein expression and its phosphorylation upon IFN-γ or IFN-α stimulation. These results are
268 comparable to the patient's clinical manifestations, showing a series of severe mycobacterial and
269 viral infection, together with identifying a lack of granuloma formation associated with
270 mycobacteria infection. Taken together, the biallelic variations identified in the patient were
271 determined to be pathogenic mutations. The patient was thus given a diagnosis of AR complete
272 STAT1 deficiency.
273

274 **STAT1 phosphorylation and DNA-binding ability in SV40-transformed fibroblasts**

275 To confirm the molecular defects observed in the patient's PBMCs, we assessed the STAT1 protein
276 and phosphorylation using SV40-transformed fibroblasts (SV40 fibroblast) from the patient, two
277 healthy controls, and a disease control from a patient with AR complete STAT1 deficiency
278 (STAT1−/−). The SV40 fibroblasts from the patient, as well as STAT1−/− SV40 fibroblasts, showed

a complete lack of STAT1 protein expression and its phosphorylation upon IFN-γ and IFN-α stimulation (**Figure 5A**). Next, the DNA-binding ability of the wild-type (WT) and mutant STAT1 proteins were analyzed by EMSA. The SV40 fibroblasts were stimulated with IFN-γ or IFN-α for 15 minutes and subjected to EMSA. As shown in Figure 5B, SV40 fibroblasts from the patient, as well as the STAT1−/− SV40 fibroblasts, presented a complete loss of DNA-binding activity to GAS in response to IFN-γ or IFN-α stimulation. These results suggested that not only PBMCs but also SV40 fibroblasts from the patient lacked a cellular response to IFN-γ or IFN-α.

Impaired cellular response to IFN-γ and IFN-α

STAT1 plays a nonredundant role in the upregulation of target genes upon IFN-γ and IFN-α stimulation. The $CD14^+$ monocytes from the patient and controls were stimulated with IFN-γ (100 IU/ml) or IFN-α (100 IU/ml) for 6 hours. They were then subjected to RNA-seq. The patient's $CD14^+$ monocyte showed global dysregulation of STAT1 target genes upon the monocyte being stimulated with IFN-γ and IFN-α (**Figure 6**). These results suggest that the patient's cells display impaired responses to IFN-γ and IFN-α.

Discussion

We herein reported a patient with AR complete STAT1 deficiency due to *STAT1* compound

297 heterozygous mutations, both located in introns: c.128+2 T>G and c.542-8 A>G. Both mutations
298 were private and absent from the public databases. Five mutations in *STAT1* have been reported to
299 cause AR complete STAT1 deficiency in previous studies.(3-7) All of the previously reported
300 *STAT1* mutations are located in exonic regions and have been identified in the homozygous state.
301 Therefore, the current case is the first AR complete STAT1 deficiency due to intronic *STAT1*
302 mutations. These mutations resulted in severe impairment of *STAT1* mRNA expression. The c.128+2
303 T>G mutation was identified by WES during the first data analysis, whereas the c.542-8 A>G
304 mutation was missed at that time because the default filtering strategy stringently excluded intronic
305 variants beyond 5 base pairs from the exon–intron boundary. Monoallelic loss-of-function *STAT1*
306 mutations, which results in normal expression and exert a dominant negative effect on WT
307 STAT1-mediated IFN-γ signaling, specifically disturb host immunity to mycobacteria and cause
308 MSMD (called AD partial STAT1 deficiency). In contrast, the previous studies, which analyzed
309 relatives of cases with AR complete STAT1 deficiency, clearly show a lack of haploinsufficiency in
310 human STAT1.(3-7,21-23) The patient in the present study was clinically diagnosed as MSMD
311 when the first exome data analysis was performed. The pathogenicity of the c.128+2 T>G mutations
312 was thus missed at that time because the mutations were inherited from the patient's asymptomatic
313 mother and predicted to disturb STAT1 protein expression by disturbing its splicing.
314 The identification of intronic mutations by comprehensive genomic study can be challenging.

315 Such mutations are easily missed in the filtering process of considerable numbers of variants of
316 unknown significance (VUS) in the analysis of exome data. Reflecting this difficulty, it took nearly
317 4 years to confirm the molecular cause in the current study. To investigate PID related genes with a
318 low expression level effectively, we performed targeted RNAseq, which enriched 426 immune
319 related genes, by focusing on PBMCs. This assay successfully ranked STAT1 as one of the top five
320 genes with reduced expression in the patient when we used another 10 inconclusive PID patients as
321 controls. This assay also detected exon skipping at exon 8 of *STAT1* due to the c.542-8 A>G
322 mutation. The results suggested that targeted RNAseq is a potentially useful diagnostic tool for
323 identifying inconclusive PID patients after WES. Recent studies revealed that RNAseq-based
324 comprehensive transcriptomic analysis is a useful tool for detecting mis-splicing, which improved
325 the diagnostic yield of inconclusive cases by WES. This in turn led to an approximately 10%–35%
326 increase in the detection of pathogenic variants.(24-26) One major difficulty in transcriptomic
327 analysis is tissue-specific expression(26). Regarding this point, PID has an advantage as a target
328 disease for RNA-seq because we can use PBMCs for analysis. However, the introduction of
329 RNA-seq in the diagnosis of PID patients is still in its primitive stage.

330 After the first genetic study, which was performed at the age of 11 months, the patient presented
331 over the following several years with several episodes of severe virus infections. He developed
332 severe influenza A infection (at 1 and 5 years old), hMPV pneumonia (at 3 years old), enterocolitis

333 and paralytic ileus by rotavirus (at 3 years old) and vaccine-strain induced varicella (at 4 years old).
334 Especially, the episodes of paralytic ileus by rotavirus and vaccine-strain induced varicella strongly
335 suggested that the patient was prone to viruses. Retrospectively, these infectious phenotypes,
336 together with the histopathological finding of mycobacterial lymphadenitis, which lacked granuloma
337 formation, suggested AR complete STAT1 deficiency as a differential diagnosis. Indeed, the lack of
338 granuloma formation is a typical finding in patients with AR complete IFN-γR1, IFN-γR2, or
339 STAT1 deficiency.(27,28) However, the clinical rarity and lack of awareness of AR complete
340 STAT1 deficiency makes suspicion of this disorder unlikely. The delayed diagnosis in the current
341 study highlights the importance of clinical information and recognition of the characteristic findings
342 of specific disorders to minimize overlooking pathogenic mutations in WES. Genetic diagnosis
343 brings significant change in management in 25%–37% of patients with PID.(29,30) Indeed, it was
344 decided that the patient in the present study should undergo HSCT after confirming the molecular
345 diagnosis of AR complete STAT1 deficiency. However, the diagnostic yield of next generation
346 sequencing in PID patients ranges from 15%–to 46% (median = 25%)(31) and has room for
347 improvement. The experience of the case presented in this study suggests that the introduction of
348 targeted RNAseq has the potential to improve the diagnostic yield of patients with PID.
349
350

Funding: This work was supported by Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science [16H05355 and 19H03620 to SO], Promotion of Joint International Research from the Japan Society for the Promotion of Science [18KK0228 to SO], and the Practical Research Project for Rare/Intractable Diseases from Japan Agency for Medical Research and development, AMED to SO.

Acknowledgements: The sequence analysis was supported by the Analysis Center of Life Science, Natural Science Center for Basic Research and Development, Hiroshima University. We thank Michael Ciancanelli for assistance.

Author contributions: All authors contributed to the accrual of subjects and/or data. SO contributed to the conception and design of the study. SS, TLV, JB and JLC drafted the manuscript. SS, MT, RK, SN, and YM performed cellular assay and gene expression experiments. TM, SK performed the clinical work and collected data. TO, TN, KI, TM, SK, OO and SO analyzed data obtained by whole exome sequence and RNAseq. All authors have revised the manuscript for important intellectual content and approved the final version.

Conflicts of interest statement: The authors declare that they have no relevant conflicts of interest.

372

373 **References**

374 1 Stark, G. R., Kerr, I. M., Williams, B. R., Silverman, R. H., and Schreiber,
375 R. D. 1998. How cells respond to interferons. *Annual review of biochemistry*
376 67:227.
377 2 Hirata, O., Okada, S., Tsumura, M., et al. 2013. Heterozygosity for the Y701C
378 STAT1 mutation in a multiplex kindred with multifocal osteomyelitis.
379 *Haematologica* 98:1641.
380 3 Boisson-Dupuis, S., Kong, X. F., Okada, S., et al. 2012. Inborn errors of
381 human STAT1: allelic heterogeneity governs the diversity of immunological and
382 infectious phenotypes. *Curr Opin Immunol* 24:364.
383 4 Dupuis, S., Jouanguy, E., Al-Hajjar, S., et al. 2003. Impaired response to
384 interferon-alpha/beta and lethal viral disease in human STAT1 deficiency. *Nat*
385 *Genet* 33:388.
386 5 Chapgier, A., Wynn, R. F., Jouanguy, E., et al. 2006. Human complete Stat-1
387 deficiency is associated with defective type I and II IFN responses in vitro
388 but immunity to some low virulence viruses in vivo. *J Immunol* 176:5078.
389 6 Vairo, D., Tassone, L., Tabellini, G., et al. 2011. Severe impairment of
390 IFN-gamma and IFN-alpha responses in cells of a patient with a novel STAT1
391 splicing mutation. *Blood* 118:1806.
392 7 Burns, C., Cheung, A., Stark, Z., et al. 2016. A novel presentation of
393 homozygous loss-of-function STAT-1 mutation in an infant with
394 hyperinflammation-A case report and review of the literature. *J Allergy Clin*
395 *Immunol Pract* 4:777.
396 8 Naruto, T., Okamoto, N., Masuda, K., et al. 2015. Deep intronic GPR143
397 mutation in a Japanese family with ocular albinism. *Sci Rep* 5:11334.
398 9 Dobin, A., Davis, C. A., Schlesinger, F., et al. 2013. STAR: ultrafast
399 universal RNA-seq aligner. *Bioinformatics* 29:15.
400 10 Picard, C., Bobby Gaspar, H., Al-Herz, W., et al. 2018. International Union
401 of Immunological Societies: 2017 Primary Immunodeficiency Diseases Committee
402 Report on Inborn Errors of Immunity. *J Clin Immunol* 38:96.
403 11 Liao, Y., Smyth, G. K., and Shi, W. 2019. The R package Rsubread is easier,
404 faster, cheaper and better for alignment and quantification of RNA sequencing
405 reads. *Nucleic Acids Res* 47:e47.
406 12 Love, M. I., Huber, W., and Anders, S. 2014. Moderated estimation of fold

407 change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550.
408 13 Robinson, J. T., Thorvaldsdottir, H., Winckler, W., et al. 2011. Integrative
409 genomics viewer. *Nat Biotechnol* 29:24.
410 14 Mizoguchi, Y., Tsumura, M., Okada, S., et al. 2014. Simple diagnosis of STAT1
411 gain-of-function alleles in patients with chronic mucocutaneous candidiasis.
412 *J Leukoc Biol* 95:667.
413 15 Kagawa, R., Fujiki, R., Tsumura, M., et al. 2017. Alanine-scanning
414 mutagenesis of human signal transducer and activator of transcription 1 to
415 estimate loss- or gain-of-function variants. *J Allergy Clin Immunol* 140:232.
416 16 Tsumura, M., Okada, S., Sakai, H., et al. 2012. Dominant-negative STAT1 SH2
417 domain mutations in unrelated patients with Mendelian susceptibility to
418 mycobacterial disease. *Hum Mutat* 33:1377.
419 17 Ueki, M., Yamada, M., Ito, K., et al. 2017. A heterozygous dominant-negative
420 mutation in the coiled-coil domain of STAT1 is the cause of
421 autosomal-dominant Mendelian susceptibility to mycobacterial diseases. *Clin*
422 *Immunol* 174:24.
423 18 Dupuis, S., Dargemont, C., Fieschi, C., et al. 2001. Impairment of
424 mycobacterial but not viral immunity by a germline human STAT1 mutation.
425 *Science* 293:300.
426 19 Chapgier, A., Boisson-Dupuis, S., Jouanguy, E., et al. 2006. Novel STAT1
427 alleles in otherwise healthy patients with mycobacterial disease. *PLoS Genet*
428 2:e131.
429 20 Sampaio, E. P., Bax, H. I., Hsu, A. P., et al. 2012. A novel STAT1 mutation
430 associated with disseminated mycobacterial disease. *J Clin Immunol* 32:681.
431 21 Chapgier, A., Kong, X. F., Boisson-Dupuis, S., et al. 2009. A partial form of
432 recessive STAT1 deficiency in humans. *J Clin Invest* 119:1502.
433 22 Kong, X. F., Ciancanelli, M., Al-Hajjar, S., et al. 2010. A novel form of
434 human STAT1 deficiency impairing early but not late responses to interferons.
435 *Blood* 116:5895.
436 23 Naviglio, S., Soncini, E., Vairo, D., Lanfranchi, A., Badolato, R., and Porta,
437 F. 2017. Long-Term Survival After Hematopoietic Stem Cell Transplantation for
438 Complete STAT1 Deficiency. *J Clin Immunol* 37:701.
439 24 Landrith, T., Li, B., Cass, A. A., et al. 2020. Splicing profile by capture
440 RNA-seq identifies pathogenic germline variants in tumor suppressor genes.
441 *NPJ Precis Oncol* 4:4.
442 25 Deelen, P., van Dam, S., Herkert, J. C., et al. 2019. Improving the

diagnostic yield of exome- sequencing by predicting gene-phenotype associations using large-scale gene expression analysis. *Nat Commun* 10:2837.

26 Marco-Puche, G., Lois, S., Benitez, J., and Trivino, J. C. 2019. RNA-Seq Perspectives to Improve Clinical Diagnosis. *Front Genet* 10:1152.

27 Bustamante, J., Boisson-Dupuis, S., Abel, L., and Casanova, J. L. 2014. Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN-gamma immunity. *Semin Immunol* 26:454.

28 Emile, J. F., Patey, N., Altare, F., et al. 1997. Correlation of granuloma structure with clinical outcome defines two types of idiopathic disseminated BCG infection. *J Pathol* 181:25.

29 Stray-Pedersen, A., Sorte, H. S., Samarakoon, P., et al. 2017. Primary immunodeficiency diseases: Genomic approaches delineate heterogeneous Mendelian disorders. *J Allergy Clin Immunol* 139:232.

30 Rae, W., Ward, D., Mattocks, C., et al. 2018. Clinical efficacy of a next-generation sequencing gene panel for primary immunodeficiency diagnostics. *Clin Genet* 93:647.

31 Yska, H. A. F., Elsink, K., Kuijpers, T. W., Frederix, G. W. J., van Gijn, M. E., and van Montfrans, J. M. 2019. Diagnostic Yield of Next Generation Sequencing in Genetically Undiagnosed Patients with Primary Immunodeficiencies: a Systematic Review. *J Clin Immunol* 39:577.

466

467 **Table.1 Summary of the patients with AR complete STAT1 deficiency**

Family	Pt	Age at onset	Origin	STAT1 mutations	Mycobacterial infections	Other infections	Outcome	Ref.
1	1	3m	Saudi Arabia	c.1757-1758 delAG (homo)	disseminated BCG	Recurrent disseminated HSV-1	Died (16m) disseminated HSV-1 (meningoencephalitis)	(4,18)
2	2	2m	Saudi Arabia	c.1799T>C (p.L600P) (homo)	disseminated BCG	Severe virus infection(suspected)	Died (12m) Viral-like illness	(4)
2	3	3m	Saudi Arabia	c.1799T>C (p.L600P) (homo)	disseminated Mycobacterial disease		Died (3m)	(3)
2	4	3m	Saudi Arabia	c.1799T>C (p.L600P) (homo)	disseminated Mycobacterial disease		Died (3m)	(3)
3	5	3m	Pakistan	c.1928insA (homo)	disseminated BCG	PolioⅢ Parainfluenza Ⅱ Rhinovirus EBV	HSCT (8m) Died (11 m) by fluminant EBV and multiorgan failure	(5)
4	6	10m	Pakistan	c.372G>C p.Q124H (homo)	disseminated *Mycobacterium kansasii*	CMV, HSV1 Sepsis Enterovirus meningitis	HSCT (4y 7m) Alive with multiple and severe complications	(6,23)
5	7	8m	Australia	c.88delA (homo)	no	Multisystem hyper-inflammation MMR vaccine: encephalopathic HHV6 HLH	HSCT (14m) Alive	(7)
6	This case	11m	Japan	c.128+2 T>G /c.542-8 A>G (compound hetero)	disseminated BCG *M. malmoense* positive mediastinal lymphadenitis	RSV, Influenza A, hMPV infections Paralytic ileus by rotavirus Vaccine induced varicella HLH	Alive (6y)	

468 HSV: Herpes simplex virus, EBV: Epstein-Barr virus, HSCT: hematopoietic stem cell
469 transplantation, CMV: Cytomegalovirus, HHV6: Human herpesvirus 6, compound hetero:
470 compound heterozygous, HLH: hemophagocytic lymphohistiocytosis, RSV: respiratory syncytial

virus, hMPV: human metapneumovirus, y: years, m: months, homo: homozygous, hetero: heterozygous

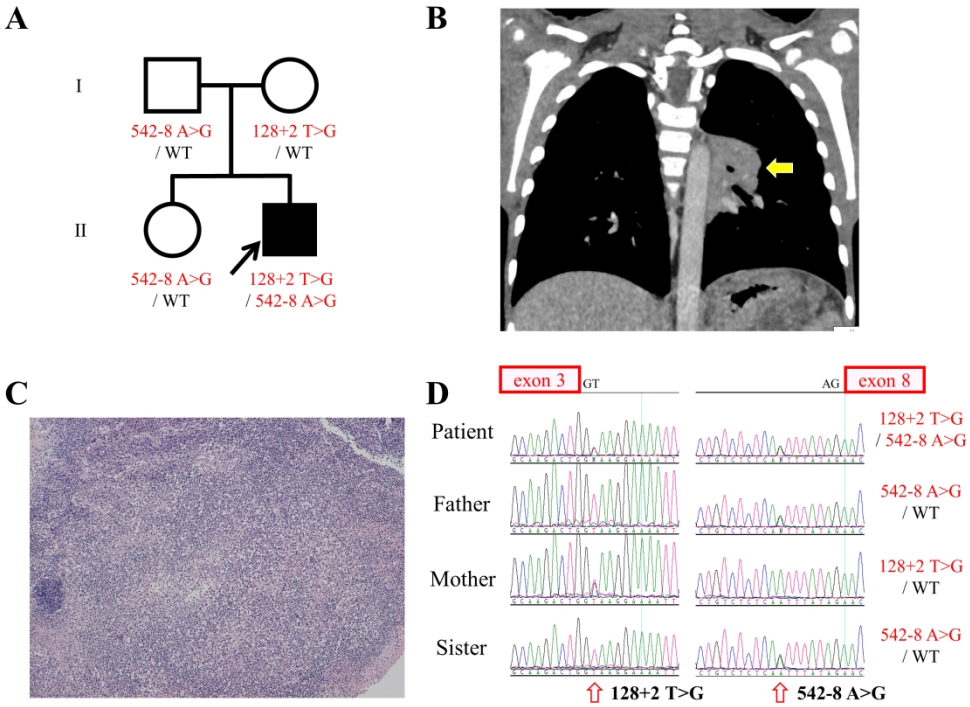

Figure 1
(A) Family pedigree. The proband is indicated with an arrow. Healthy individuals are shown in white. (B) Chest CT at 4 y 9 months. Mediastinal lymphadenitis is shown with a yellow arrow. (C) Hematoxylin and eosin staining of mediastinal lymph node biopsy shows the infiltration of various inflammatory cells, such as lymphocytes, plasmacytes, and neutrophils without granuloma formation. (D) Sanger sequence. Compound heterozygous mutations, c.128+2 T>G and c.542-8 A>G, in STAT1 identified in the patient. His mother carried heterozygous c.128+2 T>G, whereas his father and sister carry the heterozygous c.542-8 A>G mutation.

254x190mm (300 x 300 DPI)

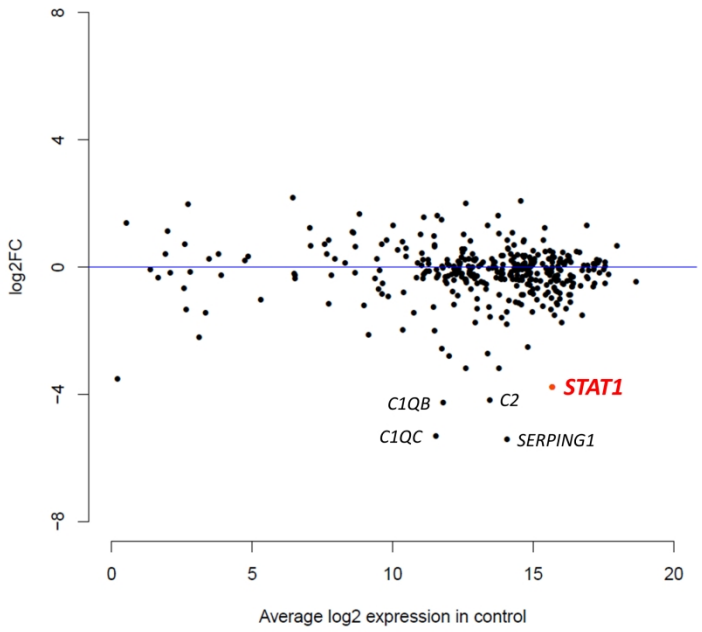

Figure 2 Dysregulated STAT1 expression in the patient's PBMCs. MAplot showing differentially expressed genes in the patient compared with inconclusive PID patients (control, n = 10) analyzed by target RNAseq. Each plot displays genes in the target RNAseq panel (n = 426). Count data of each gene was obtained by Rsubread and normalized with DESeq2. The X axis indicates average log2 expression of normalized counts in control and the y axis represents the log2 fold change (Log2FC) of normalized counts in the patient against average normalized counts of control.

254x190mm (300 x 300 DPI)

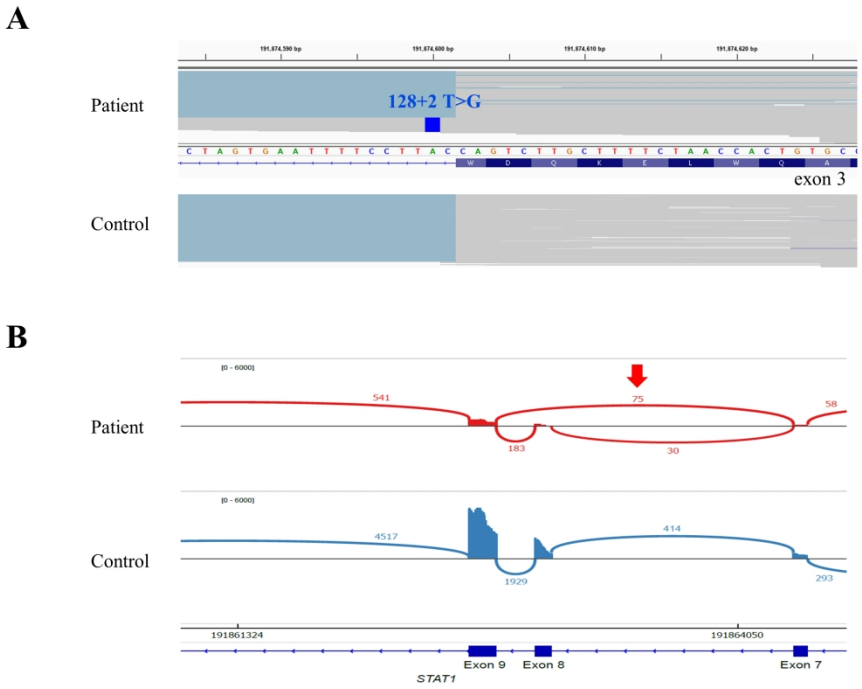

Figure 3
Biallelic abnormal splicing of STAT1 in the patient detected by targeted RNAseq. (A) The c.128+2 T>G STAT1 mutation induced intron retention at exon 3, which was shown by IGV software with the coverage over the intronic region (gray lines) only in the patient sample (upper panel). (B) Exon skipping at exon 8 (red arrow) as a result of the c.542-8 A>G mutation was shown with Sashimi plot. Each line (red line: patient, blue line: normal) and number indicate junctions and read counts, respectively.

254x190mm (300 x 300 DPI)

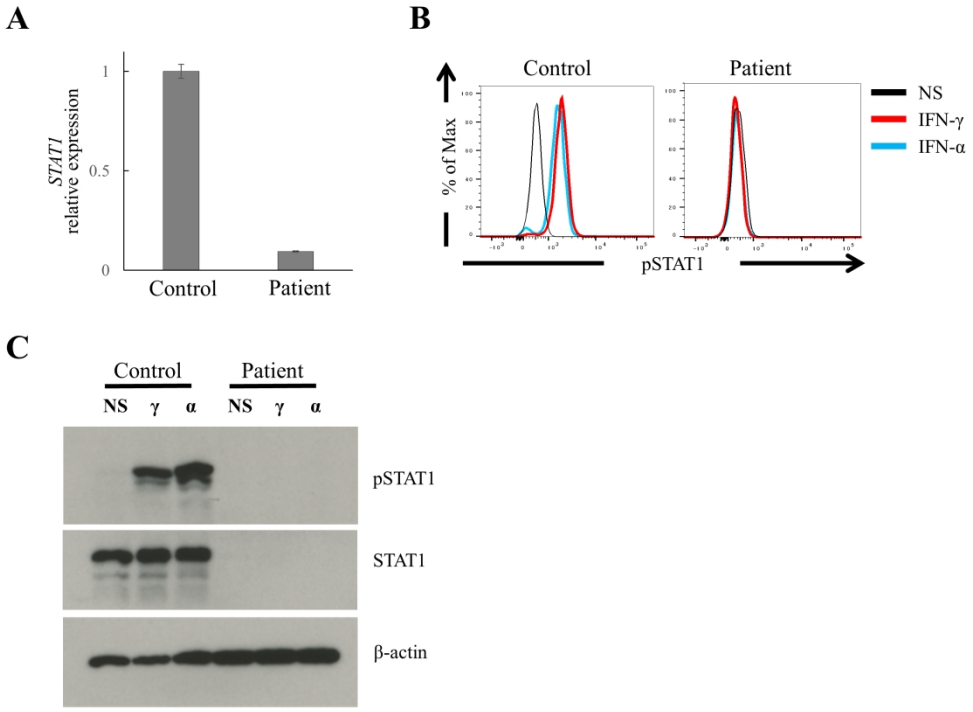

Figure 4
(A) Expression of STAT1 mRNA in PBMCs from the patient and a healthy control. Target gene expression was normalized against GAPDH and presented as n-fold increase over the expression in the healthy control. (B) Flow cytometry analysis of STAT1 phosphorylation in CD14+ cells after stimulation with IFN-γ (1,000 IU/ml, red line) or IFN-α (1,000 IU/ml, blue line). (C) Immunoblot analysis of STAT1 protein and its phosphorylation in PBMCs from the patient and a healthy control. The PBMCs were stimulated with IFN-γ (1,000 IU/ml) or IFN-α (1,000 IU/ml) for 15 minutes for the analysis. NS: no stimulation, γ: IFN-γ, α: IFN-α

254x190mm (300 x 300 DPI)

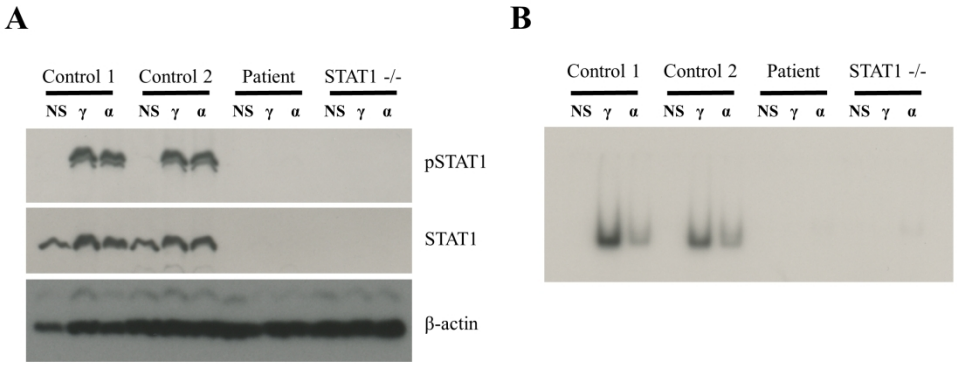

Figure 5
(A) The SV40 fibroblasts from the patient and STAT1−/− SV40 fibroblasts showed complete lack of STAT1 protein expression and its phosphorylation upon IFN-γ and IFN-α. NS: no stimulation, γ: IFN-γ, α: IFN-α (B) The DNA-binding ability of the WT and mutant STAT1 proteins were analyzed by EMSA. The SV40 fibroblasts from the patient, as well as STAT1−/− SV40 fibroblasts, presented complete loss of DNA-binding activity to GAS in response to IFN-γ or IFN-α. NS: no stimulation, γ: IFN-γ, α: IFN-α

254x190mm (300 x 300 DPI)

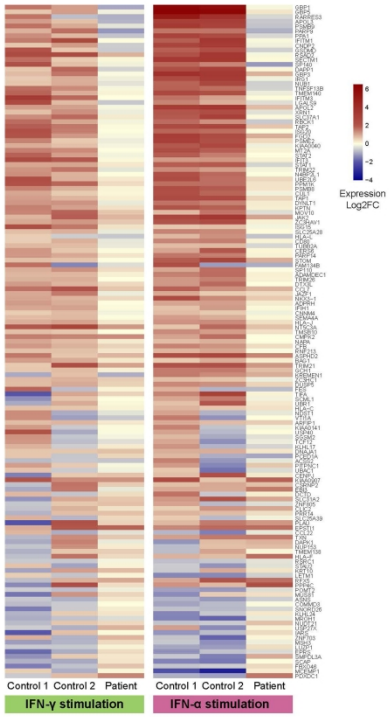

Figure 6
Global dysregulation of STAT1 target genes upon IFN-γ and IFN-α stimulation in patient's CD14+ monocyte. The CD14+ monocyte from the patient and controls are stimulated with IFN-γ (100 IU/ml; left) or IFN-α (100 IU/ml; right) for 6 hours and subjected to RNA sequencing. Log2FC; Log2 fold change.

254x190mm (300 x 300 DPI)