

ORIGINAL ARTICLE

# Aggregation of rare/low-frequency variants of the mitochondria respiratory chain-related proteins in rheumatoid arthritis patients

Shigeki Mitsunaga<sup>1</sup>, Kazuyoshi Hosomichi<sup>2,3</sup>, Yuko Okudaira<sup>1</sup>, Hirofumi Nakaoka<sup>2</sup>, Yasuo Suzuki<sup>4</sup>, Masataka Kuwana<sup>5,6</sup>, Shinji Sato<sup>4</sup>, Yuko Kaneko<sup>6</sup>, Yasuhiko Homma<sup>7</sup>, Akira Oka<sup>8</sup>, Takashi Shiina<sup>1</sup>, Hidetoshi Inoko<sup>1</sup> and Ituro Inoue<sup>2</sup>

Exome sequencings were conducted using 59 patients having rheumatoid arthritis (RA) and 93 controls. After stepwise filtering, 107 genes showed less than 0.05 of *P*-values by gene-burden tests. Among 107 genes, *NDUFA7* which is a subunit of the complex I in the mitochondrial respiratory chain was selected for further analysis based on previous reports. A case–control study was performed on the three single-nucleotide variants (SNVs) of *NDUFA7* with 432 cases and 432 controls. An association was observed between *NDUFA7* and RA with severe erosive arthritis. These results together with previous reports suggested the involvement of reactive oxygen species (ROS) in the pathogenesis of RA. In the next step, four SNVs from three genes related to the mitochondrial respiratory chain were selected, which is a major source of ROS, and conducted a case–control study. An association was observed based on a pathway-burden test comprising *NDUFA7*, *SDHAF2*, *SCO1* and *ATP5O*: *P* = 1.56E-04, odds ratio = 2.16, 95% confidence interval = 1.43–3.28. Previous reports suggested the involvement of ROS in the pathogenesis of RA. The aggregation of SNVs in the mitochondria respiratory chain suggests the pivotal role of those SNVs in the pathogenesis of RA with severe erosive arthritis.

*Journal of Human Genetics* (2015) 60, 449–454; doi:10.1038/jhg.2015.50; published online 28 May 2015

## INTRODUCTION

Rheumatoid arthritis (RA) is a multifactorial and systemic autoimmune disease and characterized by synovial inflammation and hyperplasia, autoantibody production, cartilage and bone destruction and systemic features including cardiovascular, pulmonary, psychological and skeletal disorders.<sup>1</sup> Although the causality of RA is not completely understood, genetic factors contribute to the onset. About 60% of the RA risk is genetic<sup>2</sup> and one-third of the genetic risk of RA is attributed to the major histocompatibility complex (human leukocyte antigens in human) region where HLA-DRB1 is strongly associated. Several genome-wide association studies based on the common disease–common variant hypothesis as well as meta-analyses have been conducted to elucidate the genetic risk of RA. In a recent meta-analysis using greater than 100 000 subjects, the 100 RA risk loci outside the human leukocyte antigen region explained only 5.5% and 4.7% of heritability in Europeans and Asians, respectively.<sup>3</sup> On the other hand, common disease–multiple rare variants hypothesis has been proposed to explain the remaining genetic risk. Namely, a significant portion of the inherited susceptibility to common diseases

may be due to summation of the effects of a series of low-frequency dominantly and independently acting variants of a variety of different genes.<sup>4,5</sup> These variants cannot be detected by conventional genome-wide association studies, thereby exome sequencing is an alternative tool.

Deep exon sequencing has been used for detecting rare/low-frequency variants related to common diseases. In the DNA sequence of the exons plus exon/intron boundaries of *LPL* of 313 patients with hypertriglyceridemia, 20 rare variants were detected including previously reported seven variants.<sup>6</sup> Regarding RA, deep exon sequencing of 25 biological candidate genes which were identified by genome-wide association studies revealed that the aggregated contribution of low-frequency and common coding variants to the risk of RA using 10 609 RA cases and 35 605 controls.<sup>7</sup> Aggregation of rare coding variants in *PLD3* of the patients with Alzheimer's disease has also been reported.<sup>8</sup>

Exome sequencing has been introduced to identify causative- or susceptible- rare/low-frequency variants in Mendelian diseases and common diseases, and numerous successful results have been

<sup>1</sup>Department of Molecular Life Sciences, Division of Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, Isehara, Japan; <sup>2</sup>Division of Human Genetics, Department of Integrated Genetics, National Institute of Genetics, Mishima, Japan; <sup>3</sup>Department of Bioinformatics and Genomics, Graduate School of Medical Sciences, Kanazawa University, Kanazawa, Japan; <sup>4</sup>Division of Rheumatology, Department of Internal Medicine, Tokai University School of Medicine, Isehara, Japan; <sup>5</sup>Department of Allergy and Rheumatology, Nippon Medical School Graduate School of Medicine, Tokyo, Japan; <sup>6</sup>Division of Rheumatology, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan; <sup>7</sup>Department of Clinical Health Science, Tokai University School of Medicine, Isehara, Japan and <sup>8</sup>The Institute of Medical Science, Tokai University School of Medicine, Isehara, Japan

Correspondence: Professor I Inoue, Division of Human Genetics, Department of Integrated Genetics, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan. E-mail: itinoue@nig.ac.jp

Received 3 December 2014; revised 14 April 2015; accepted 16 April 2015; published online 28 May 2015

reported.<sup>9,10</sup> We also performed exome sequencing to identify genetic risk factors of RA and reported *BTNL2* as a RA-susceptible gene.<sup>11</sup> Here, we focused on rare/low-frequency variants aggregated in a gene, which are related to RA. We identified the aggregation of rare/low-frequency variants in the mitochondrial respiratory chain-related genes. Oxidative stress has been considered to be involved in the onset of RA.<sup>12</sup> Indeed, levels of H<sub>2</sub>O<sub>2</sub> and O<sub>2</sub><sup>-</sup> in RA patients' plasma were significantly higher than those in controls.<sup>13</sup> Elevated levels of urinary 8-hydroxyguanine (8-OHdG), which is a standard biomarker for oxidative stress have also been reported in RA patients.<sup>14,15</sup> It is likely that the single-nucleotide variants (SNVs) associated with RA contribute to the increase of electron leakage in the mitochondrial respiratory chain and production of reactive oxygen species (ROS) through reducing oxygen molecules to the superoxide anion by the leaked electrons. The aggregation of rare/low-frequency variants in the mitochondrial respiratory chain suggests genetic bases of oxidative stress in RA patients.

## MATERIALS AND METHODS

### Subjects

A total of 432 Japanese RA patients (male:  $n=72$ ; average age=62.1, female:  $n=360$ ; average age=58.8) were enrolled from outpatients of the Division of Rheumatology, Tokai University Hospital and Division of Rheumatology, Keio University Hospital. All RA patients fulfilled the 1987 revised criteria of the American College of Rheumatology.<sup>16</sup> A total of 432 unrelated healthy Japanese controls without a history of autoimmune diseases (male:  $n=202$ ; average age=58.7, female:  $n=230$ ; average age=45.6) were recruited at the Health Evaluation and Promotion Center of Tokai University Hospital.

The RA patients were classified into three groups based on the number of joints with erosion and the time course of erosion:<sup>17</sup> a subset with mutilating disease (MUD: 17 cases), a subset with more erosive disease (MES: 276 cases) and a subset with least erosive disease (LES: 139 cases). The order of Sharp score is expected to be MUD>MES>LES. The MUD group was combined with the MES group as severe erosive cases for statistical analyses because of the small number of MUD cases.

All subjects gave written informed consent for genetic analyses. Ethical approvals for this study were obtained from the ethics committee of Tokai University School of Medicine and from the ethics committee of Keio University School of Medicine.

### Next Generation DNA Sequencing

DNAs were extracted from peripheral blood using a DNA extraction kit, Genomix (Biologica, Nagoya, Japan). Exome sequencing was performed for 59 patients with severe erosive RA comprising 7 MUD cases and 52 MES cases together with 93 controls without a history of autoimmune diseases using the SureSelect Human All Exon kit (Agilent Technologies, Santa Clara, CA, USA), and the enriched libraries were sequenced by using an Illumina HiSeq 2000 sequencer (Illumina, San Diego, CA, USA) with 2 × 100 bp paired-end module. The reads were mapped to the reference genome (UCSC hg19, NCBI GRCh37) using bwa v.0.7.5.<sup>18</sup> BWA generated sam files were sorted and indexed using Samtools v.0.1.18.<sup>19</sup> Then duplicated reads were marked with Picard v.1.102 (<https://github.com/broadinstitute/picard>). The files obtained in bam format were analyzed using Genome Analysis Toolkit v.2.7 (GATK)<sup>20,21</sup> following their best practices guidelines. In brief, the bam files were first subject to indel realignment, base quality score recalibration and then SNVs and small indels calling with the UnifiedGenotyper walker, to obtain the potential variants in a vcf file. Variants in vcf files were annotated using script in ANNOVAR (version 2013jul21).<sup>22</sup>

### SNV filtering, narrow down and statistical analyses of exome-derived SNVs

SNVs detected in both of cases and controls by the exome sequencing were filtered out based on the following criteria: SNVs with phred-scaled variant quality provided by the UnifiedGenotyper of GATK<sup>20,21,23</sup> <100, SNVs with

supporting read <8, synonymous SNVs, SNVs on segmental duplications, SNVs with no reads in greater than 5% of samples. To focus on the rare/low-frequency variants, the genes having SNV(s) with minor allele frequency (MAF) >0.05 in control were excluded. In this step, genes harboring common SNV(s) and rare/low-frequency variant(s) as well as genes harboring only common SNV(s) with greater than 5% of MAF were excluded, because it is likely that the RA-susceptible genes harboring common SNV(s) were already identified by genome-wide association studies and to narrow down candidate genes. A total of 6012 SNVs (3806 genes) were obtained after filtering. In the next step, a gene-burden test<sup>24,25</sup> was performed using filtered SNVs from 59 RA samples and 93 controls to prioritize pathways in which RA-susceptible genes involved. Fisher's exact test was used for the gene-burden test based on the number of cases or controls with or without greater than one SNV of interest in each gene. Among the genes which showed significance ( $P<0.05$ ) in a gene-burden test, we addressed a gene in the mitochondrial respiratory chain based on the previous reports and observation of SNVs in the gene with a relatively high proportion among RA patients. And then, a case-control study followed by a pathway-burden test was conducted. Steps of SNV filtering and narrow down SNVs were shown by a flow chart (Supplementary Figure 1).

Case-control studies were conducted with the SNVs contained in the genes of the pathway (the mitochondrial respiratory chain) using Sanger sequencing of SNVs on the 432 cases and 432 controls. One SNV in *SCO1* at the position 10600796 on chromosome 17, which was newly detected in the amplicon for Sanger sequencing of the SNV at the position 10600794 on chromosome 17 in the case-control study, was included in the following statistical analysis. The gene-burden test on each gene was performed. The pathway-burden test was performed by integrating the results on selected genes in the mitochondrial respiratory chain. Namely, a 2 × 2 contingency table based on the number of cases or controls with or without greater than one SNV in the genes was used for Fisher's exact test.

## RESULTS

In the exome sequencing, mapping rate of the alignment was 90.9%, duplication rate of mapped reads was 17.3%, and the average coverage depth was 143.1 × in 59 patients with 96.2% of the target bases covered by at least eight reads. In each sample, 18 171 exonic SNVs and 404 exonic indels were identified on average. The SNV transition/transversion rate was 2.92. After filtering the data, we performed gene-burden test for obtaining candidate RA-susceptible genes. Gene-burden tests indicated that 107 genes showed differences between cases and controls ( $P<0.05$ ). Top 20 out of 107 genes were shown in Table 1. 27.1% of the RA patients carried rare/low-frequency variants on *NDUFA7*, which showed a  $P$ -value 6.68E-03, and we focused on the gene in the following study based on several lines of evidence (i) *NDUFA7* is a subunit of mitochondrial respiratory chain complex I, (ii) ~90% of ROS is produced in the mitochondria,<sup>24</sup> (iii) oxidative stress has been considered to be involved in the onset of RA.<sup>12</sup> We performed a case-control study using 432 Japanese RA patients and 432 Japanese healthy controls on the association of three SNVs in *NDUFA7* (positions 8381416, 8381435 and 8385744 on chromosome 19, Supplementary Figure 2) with RA based on the gene-burden test, which were observed in exome sequencing of RA samples. Nominally significant at  $P<0.05$  association was observed between SNVs of *NDUFA7* with severe erosive RA (MUD plus MES):  $P=0.011$ , odds ratio (OR) = 1.81, 95% confidence interval (CI) = 1.14–2.90, although significance was not reached for overall RA:  $P=0.134$ , OR = 1.41, 95% CI = 0.90–2.21.

If SNVs of *NDUFA7* result in increasing electron leakage followed by ROS production through reducing oxygen molecules to superoxide anion and ROS from mitochondria is involved in RA onset, other SNVs related to mitochondria respiratory chain might be associated with RA. Then, we re-explored those genes: subunits of mitochondrial respiratory chain complexes I–V (<http://www.genenames.org/>

**Table 1** Rheumatoid arthritis-susceptible candidate genes obtained from exome sequencing followed by filtering<sup>a</sup>

Gene	Chr	Number of SNV in cases	Number of SNV in controls	Number of shared SNV	Number of cases having SNV	Number of control having SNV	P-value <sup>b</sup>	OR <sup>b</sup>	95% CI <sup>b</sup>
WNT10A	2	4	0	0	9	0	1.34 × 10 <sup>-4</sup>	—	—
PLIN5	19	3	0	0	9	0	1.34 × 10 <sup>-4</sup>	—	—
PALB2	16	4	1	1	11	2	5.96 × 10 <sup>-4</sup>	10.4	2.22–49.0
SEMA6D	15	2	1	1	8	1	2.33 × 10 <sup>-3</sup>	14.4	1.76–119
TMEM132E	17	6	1	0	8	1	2.33 × 10 <sup>-3</sup>	14.4	1.76–119
CDKN2A	9	3	2	2	12	4	2.47 × 10 <sup>-3</sup>	5.68	1.74–18.6
MYOC	1	4	3	2	11	3	2.59 × 10 <sup>-3</sup>	6.88	1.83–25.8
SLC9A9	3	4	2	2	11	3	2.59 × 10 <sup>-3</sup>	6.88	1.83–25.8
MFSD3	1	2	0	0	6	0	2.91 × 10 <sup>-3</sup>	—	—
ZNF556	19	3	0	0	6	0	2.91 × 10 <sup>-3</sup>	—	—
CSPP1	8	5	2	2	14	6	2.98 × 10 <sup>-3</sup>	4.51	1.62–12.5
UHRF1BP1L	12	6	1	1	9	2	3.46 × 10 <sup>-3</sup>	8.19	1.70–39.4
PRR16	5	1	1	1	11	4	5.29 × 10 <sup>-3</sup>	5.10	1.54–16.9
MLL7	17	2	1	1	7	1	5.79 × 10 <sup>-3</sup>	12.4	1.48–103
TMUB2	17	2	1	1	7	1	5.79 × 10 <sup>-3</sup>	12.4	1.48–103
COBLL1	2	6	5	4	16	9	6.68 × 10 <sup>-3</sup>	3.47	1.42–8.50
NDUFA7	19	3	1	1	16	9	6.68 × 10 <sup>-3</sup>	3.47	1.42–8.50
ALCAM	3	4	4	3	12	5	7.01 × 10 <sup>-3</sup>	4.49	1.49–13.5
FPGT-TNNI3K	1	3	0	0	5	0	7.91 × 10 <sup>-3</sup>	—	—
ZNHIT6	1	2	0	0	5	0	7.91 × 10 <sup>-3</sup>	—	—

Abbreviations: 95% CI, 95% confidence interval; OR, odds ratio; SNV, single-nucleotide variant.

<sup>a</sup>The top 20 candidate genes ranked by P-value are listed.<sup>b</sup>P-value, OR and 95% CI indicate the results of gene-burden test.**Table 2** Rheumatoid arthritis-susceptible candidate genes and SNVs selected

Gene	Chr	position	SNV			Positive case <sup>a</sup>		MAF	
			rs number	Base changes (reference/observed)	Amino acid changes	RA	Control	RA	Control
NDUFA7	19	8381416	—	G/A	S72L	2	0	0.0169	0
		8381435	rs2288415	G/C	P66A	11	9	0.0932	0.0484
		8385744	—	T/C	K33R	3	0	0.0254	0
ATP5O	21	35281421	rs4842	T/C	K98R	2	2 <sup>b</sup>	0.0169	0.0161
		35286771	—	A/G	F24L	1	0	0.0085	0
SDHAF2	11	61205535	—	G/A	R107H	3	0	0.0254	0
SCO1	17	10600794	—	G/T	V11I	2	1	0.0169	0.0054
		10900796	—	C/A <sup>c</sup>	R10L	0	0	0	0

Abbreviations: RA, rheumatoid arthritis; SNV, single-nucleotide variant.

<sup>a</sup>Number of SNV-positive cases among 59 RA cases or 93 controls.<sup>b</sup>One control subject is homozygote of the SNV.<sup>c</sup>Observed in the case-control study but not in the exome sequencing.

genefamilies/mitocomplex), and mitochondrial respiratory chain complex assembly factors (<http://www.genenames.org/genefamilies/MITOAF>). A total of 105 SNVs in 59 genes were observed from 59 exome sequencing results of RA patients: 17 SNVs in 13 genes among 38 complex I genes, no SNVs in 4 genes of complex II, 3 SNVs in 3 genes among 9 complex III genes, 2 SNVs in 2 genes among 16 complex IV genes, 11 SNVs in 8 genes from 17 complex V genes and 27 SNVs in 14 genes of 35 mitochondrial respiratory chain complex assembly factors (Supplementary Tables 1 and 2). We selected the genes with greater than 1 of OR for further analysis. Three genes were selected other than *NDUFA7*: *SDHAF2*, *SCO1* and *ATP5O* (Table 2). The results of the gene-burden test using 432 cases and 432 controls were shown in Table 3. Based on the results of the gene-burden tests, genes were subjected to the pathway-burden test. Sixty-eight individuals of 292 RA patients with severe erosion and 53 individuals out of

431 controls had greater than one SNV in the mitochondrial respiratory chain related genes (pathway) containing *NDUFA7*, *SDHAF2*, *SCO1* and *ATP5O*. Four RA patients with severe erosion and no controls had two SNVs in those genes. An association was observed with severe erosive RA:  $P = 1.56E-04$ , OR: 2.16, 95% CI = 1.43–3.28 (Table 3). The allele frequencies of respective SNVs in 432 RA cases and 432 controls are summarized in Supplementary Table 3.

## DISCUSSION

Rare variants for common diseases are highlighted owing to the available massive sequencing data. However, single-variant tests with rare variants have less power comparing with common variants. Accordingly, aggregation tests<sup>24,25</sup> evaluating cumulative effects of multiple rare variants in a gene or a pathway, which are associated with a disease, have increased power to detect genetic causality.<sup>26</sup>

Filtering and narrowing down of SNVs for quality control and focusing on rare/low-frequency variants, 3806 genes were obtained as RA-susceptible candidate genes (Supplementary Figure 1). It indicated that the significant level of *P*-value was  $1.32 \times 10^{-5}$  by gene-burden test after Bonferroni Correction and the *P*-values of gene-burden tests using the exome-sequencing data did not reach the significant level. However, we hypothesized that significant association based on a pathway would be detected by integrating multiple rare/low-frequency variants which were observed in the exome-sequencing but not reached the significant level after Bonferroni correction because of small number of rare SNVs in each gene. In this context, gene-burden test was useful for prioritizing of pathways in which RA-susceptible genes involved. Consequently, a significant association between severe erosive RA and mitochondrial respiratory complex-related genes which were major site of ROS generation was detected. Owing to the low frequency of each SNV, the variance explained by each of the SNVs in *NDUFA7*, *ATP5O*, *SCO1* and *SDHAF2* was at most 0.13% by using a liability threshold model<sup>27</sup> in which the prevalence of RA was assumed to be 1.0%. Increment of sample size and the functional analyses of SNVs in the mitochondrial respiratory chain are subjects of future investigations.

Mitochondrial respiratory complexes I and III have been considered to be the major ROS generation site because large changes in the potential energy of the electrons occur in the sites.<sup>28</sup> However, other complexes in the mitochondrial respiratory chain have been also

shown to be related with ROS production using inhibitors and siRNA.<sup>29–31</sup> Therefore, it is likely that SNVs in the genes related to those complexes elicit ROS production. Five complexes of mitochondrial respiratory chain are constituted with proteins encoded 97 genes and 35 assembly factors are involved in the formation of the complexes. Mitochondrial flavoenzymes such as  $\alpha$ -ketoglutarate dehydrogenase (*OGDH*) and glycerol phosphate dehydrogenase (*GPD2*) are also important sources of ROS.<sup>32</sup> In this study, a limited number of rare/low-frequency variants were observed because of a relatively small number of exome-sequenced cases used. Some particular rare/low-frequency variants in other genes, which were not detected in this study, may be involved in the ROS production in both of severe erosive RA and least erosive RA.

Oxidative stress has been considered to be involved in the onset of RA.<sup>12</sup> RA is characterized by synovial inflammation and hyperplasia, autoantibody production, cartilage and bone destruction and systemic features.<sup>2</sup> Previous reports suggest ROS is underlying some of these characteristics.

Anti-cyclic citrullinated peptides antibody (Ab) has been utilized as a biomarker for RA diagnosis because of its high specificity and sensitivity to RA.<sup>33</sup> *PADI4* has been reported to be a RA-susceptible gene.<sup>34</sup> *PADI4*, one of the members of peptidylarginine deiminase (PADI) that catalyzes the conversion of arginine residues into citrulline, was expressed in hematological and RA synovial tissues.<sup>35</sup> It was reported that the mRNA level of *PADI4* was increased in U373MG cells and the citrullinated proteins were also increased in U-2 OS cells by transfection of *p53*-expression vector.<sup>36</sup> ROS activates *p53* through the activation of *p38MAPK*<sup>37</sup> or damaging DNA. The overexpression of *p53* in synovial tissue from patients with RA was reported.<sup>38</sup> In addition, the somatic mutations in *p53* have been found within the RA synovium and greater than 80% of the mutations are G/A and T/C transitions, which are characteristic of oxidative deamination by NO or oxygen radicals.<sup>39</sup> Both wild-type *p53* and mutated *p53* effectively induce NF- $\kappa$ B.<sup>40</sup> Therefore, it is likely that the SNVs aggregated in the mitochondrial respiratory chain-related proteins increase the leakage of ROS followed by activation of *p53* and overexpression of *PADI4*. Aberrantly produced citrullinated proteins will activate immune system and cause overproduction of pro-inflammatory cytokines such as TNF- $\alpha$  and IL-6 (Figure 1).

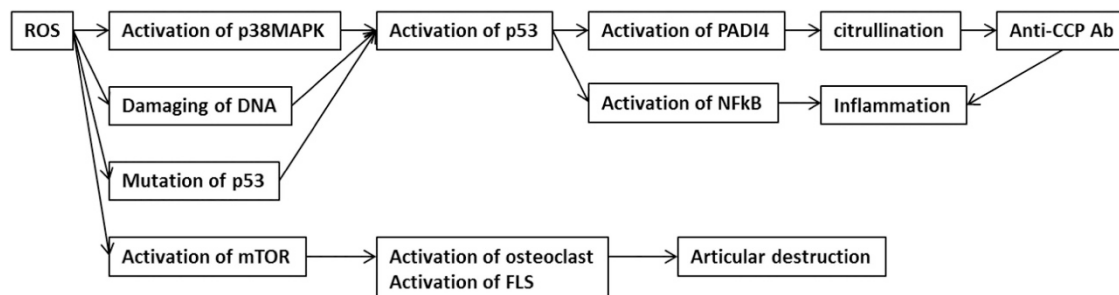
ROS activates mTOR<sup>41</sup> which activates osteoclasts and fibroblast-like synoviocytes. Inhibition of mTOR by sirolimus or everolimus reduced synovial osteoclast formation and protected against local bone erosions and cartilage loss in human tumor necrosis factor transgenic

**Table 3 Association analysis of rheumatoid arthritis with selected genes**

	Overall			Severe erosive RA (MUD+MES)		
	P-value	OR	95% CI	P-value	OR	95% CI
<i>Gene-burden test</i>						
<i>NDUFA7</i>	0.134	1.41	0.90–2.21	0.011	1.81	1.14–2.90
<i>ATP5O</i>	0.049	2.43	1.00–6.49	0.028	2.65	1.02–7.39
<i>SDHAF2</i>	0.624	3.00	0.24–158	0.309	4.44	0.36–233
<i>SCO1</i>	0.064	4.57	0.94–43.7	0.035	5.26	0.99–52.2
<i>Pathway-burden test<sup>a</sup></i>						
	6.58E-03	1.70	1.15–2.53	$1.56 \times 10^{-4}$	2.16	1.43–3.28

Abbreviations: CI, confidence interval; MUD, mutilating disease; MES, more erosive disease; OR, odds ratio; RA, rheumatoid arthritis.

<sup>a</sup>Pathway-burden test was performed by combining the SNVs in *NDUFA7*, *ATP5O*, *SDHAF2* and *SCO1*.



**Figure 1** Involvement of ROS in the onset of rheumatoid arthritis: bibliographic consideration. ROS activates *p53* through activation of *p38MAPK*, damaging DNA and mutating *p53*. Activated *p53* promotes aberrant citrullination of proteins through activation of *PADI4* and results in production of anti-cyclic citrullinated peptides antibody. Activated *p53* also promotes inflammation through activation of NF $\kappa$ B as well as immunoreactions caused by anti-cyclic citrullinated peptides antibody. On the other hands, ROS causes articular destruction through activation of mTOR followed by activation of osteoclast and fibroblast-like synoviocytes.

mice.<sup>42</sup> Rapamycin treatment reduces the phosphorylation of mTOR and decreased fibroblast-like synoviocytes invasion in RA tissues which correlate with radiographical and histological damage in RA.<sup>43</sup> Thus, ROS is also involved in cartilage and bone destruction through activation of osteoclast and fibroblast-like synoviocytes (Figure 1).

Recently, resveratrol that is a natural anti-inflammatory antioxidant has been reported to be a promising strategy in controlling synovial inflammatory response.<sup>44</sup> It supports our findings in this study.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank the DNA donors and supporting medical staff for making this study possible. We thank Miyako Nakagawa and Tomomi Ito (Division of Rheumatology, Tokai University School of Medicine), Masayuki Tanaka and Hideki Hayashi (Department of Bioinformatics, Support Center for Medical Research and Education, Tokai University) for their excellent technical assistance. We also thank Dr Atsushi Toyoda (Comparative Genomics Laboratory, National Institute of Genetics), Dr Shoji Tsuji, Dr Jun Mitsui and Dr Hiroyuki Ishiura (Department of Neurology, Division of Neuroscience, Graduate School of Medicine, The University of Tokyo), Dr Shinichi Morishita and Dr Jun Yoshimura (Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo) for their supporting based on the Grant-in-Aid-for Scientific Research on Innovative Areas 'Genome Science' (No.221S0002) from Ministry of Education, Culture, Sports, Science and Technology of Japan. This work was supported in part by, a Grant-in-Aid-for Scientific Research on Innovative Areas 'Genome Science' (No.221S0002) from Ministry of Education, Culture, Sports, Science and Technology of Japan and a Grant-in-Aid for Scientific Research from the Japanese Ministry of Health, Labour and Welfare.

- McInnes, I. B. & Schett, G. The pathogenesis of rheumatoid arthritis. *N. Engl. J. Med.* **365**, 2205–2219 (2011).
- MacGregor, A. J., Snieder, H., Rigby, A. S., Koskenvuo, M., Kaprio, J., Aho, K. *et al.* Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum.* **43**, 30–37 (2000).
- Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K. *et al.* (2014) Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**, 376–381 (2014).
- Bodmer, W. & Bonilla, C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat. Genet.* **40**, 695–701 (2008).
- Cirulli, E. T. & Goldstein, D. B. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat. Rev. Genet.* **11**, 415–425 (2010).
- Evans, D., Arze, J., Aberle, J. & Beil, F. U. Rare variants in the lipoprotein lipase (LPL) gene are common in hypertriglyceridemia but rare in Type III hyperlipidemia. *Atherosclerosis* **214**, 386–390 (2011).
- Diogo, D., Kurreeman, F., Stahl, E. A., Liao, K. P., Gupta, N., Greenberg, J. D. *et al.* Rare, low-frequency, and common variants in the protein-coding sequence of biological candidate genes from GWASs contribute to risk of rheumatoid arthritis. *Am. J. Hum. Genet.* **92**, 15–27 (2013).
- Cruchaga, C., Karch, C. M., Jin, S. C., Benitez, B. A., Cai, Y., Guerreiro, R. *et al.* Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. *Nature* **505**, 550–554 (2014).
- Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C. *et al.* Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272–276 (2009).
- Ng, S. B., Buckingham, K. J., Lee, C., Bigham, A. W., Tabor, H. K., Dent, K. M. *et al.* Exome sequencing identifies the cause of a mendelian disorder. *Nat. Genet.* **42**, 30–35 (2010).
- Mitsunaga, S., Hosomichi, K., Okudaira, Y., Nakaoka, H., Kunii, N., Suzuki, Y. *et al.* Exome sequencing identifies novel rheumatoid arthritis-susceptible variants in the BTNL2. *J. Hum. Genet.* **58**, 210–215 (2013).
- Phillips, D. C., Dias, H. K., Kitas, G. D. & Griffiths, H. R. Aberrant reactive oxygen and nitrogen species generation in rheumatoid arthritis (RA): causes and consequences for immune function, cell survival, and therapeutic intervention. *Antioxid. Redox Signal.* **12**, 743–785 (2010).
- Veselinovic, M., Barudic, N., Vuletic, M., Zivkovic, V., Tomic-Lucic, A., Djuric, D. *et al.* Oxidative stress in rheumatoid arthritis patients: relationship to diseases activity. *Mol. Cell. Biochem.* **391**, 225–232 (2014).
- Rall, L. C., Roubenoff, R., Meydani, S. N., Han, S. N. & Meydani, M. Urinary 8-hydroxy-2'-deoxyguanosine (8-OHdG) as a marker of oxidative stress in rheumatoid arthritis and aging: effect of progressive resistance training. *J. Nutr. Biochem.* **11**, 581–584 (2000).
- Kageyama, Y., Takahashi, M., Ichikawa, T., Torikai, E. & Nagano, A. Reduction of oxidative stress marker levels by anti-TNF-alpha antibody, infliximab, in patients with rheumatoid arthritis. *Clin. Exp. Rheumatol.* **26**, 73–80 (2008).
- Arnett, F. C., Edworthy, S. M., Bloch, D. A., McShane, D. J., Fries, J. F., Cooper, N. S. *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.* **31**, 315–324 (1988).
- Ochi, T., Iwase, R., Yonemasu, K., Matsukawa, M., Yoneda, M., Yukioka, M. *et al.* Natural course of joint destruction and fluctuation of serum C1q levels in patients with rheumatoid arthritis. *Arthritis Rheum.* **31**, 37–43 (1988).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A. *et al.* The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
- Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
- Li, H., Ruan, J. & Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* **18**, 1851–1858 (2008).
- Lee, S., Abecasis, G. R., Boehnke, M. & Lin, X. Rare-variant association analysis: study designs and statistical tests. *Am. J. Hum. Genet.* **95**, 5–23 (2014).
- Morgenthaler, S. & Thilly, W. G. A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). *Mutat Res.* **615**, 28–56 (2007).
- Nakaoka, H., Cui, T., Tajima, A., Oka, A., Mitsunaga, S., Kashiwase, K. *et al.* A systems genetics approach provides a bridge from discovered genetic variants to biological pathways in rheumatoid arthritis. *PLoS One* **6**, e25389 (2011).
- Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* 4 edn (Longmans Green, Harlow, Essex, UK, 1996).
- Balaban, R. S., Nemoto, S. & Finkel, T. Mitochondria, oxidants, and aging. *Cell* **120**, 483–495 (2008).
- Siebels, I. & Dröse, S. Q-site inhibitor induced ROS production of mitochondrial complex II is attenuated by TCA cycle dicarboxylates. *Biochim. Biophys. Acta* **1827**, 1156–1164 (2013).
- Jayanthi, S., Lewis, B. D. & Cadet, J. L. Fas-induced apoptosis of glioma cells is associated with down-regulation of the hSCO1 protein, a subunit of complex IV. *Brain Res. Mol. Brain Res.* **91**, 131–136 (2001).
- Natera-Naranjo, O., Kar, A. N., Aschrafi, A., Gervasi, N. M., Macgibeny, M. A., Gioio, A. E. *et al.* Local translation of ATP synthase subunit 9 mRNA alters ATP levels and the production of ROS in the axon. *Mol. Cell. Neurosci.* **49**, 263–270 (2012).
- Tahara, E. B., Navarete, F. D. & Kowaltowski, A. J. Tissue-, substrate-, and site-specific characteristics of mitochondrial reactive oxygen species generation. *Free Radic. Biol. Med.* **46**, 1283–1297 (2009).
- Schellekens, G. A., Visser, H., de Jong, B. A., van den Hoogen, F. H., Hazes, J. M., Breedveld, F. C. *et al.* The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum.* **43**, 155–163 (2000).
- Iwamoto, T., Ikari, K., Nakamura, T., Kuwahara, M., Toyama, Y., Tomatsu, T. *et al.* Association between PADI4 and rheumatoid arthritis: a meta-analysis. *Rheumatology (Oxford)* **45**, 804–807 (2006).
- Chang, X., Yamada, R., Suzuki, A., Sawada, T., Yoshino, S., Tokuiro, S. *et al.* Localization of peptidylarginine deiminase 4 (PADI4) and citrullinated protein in synovial tissue of rheumatoid arthritis. *Rheumatology (Oxford)* **44**, 40–50 (2005).
- Tanikawa, C., Ueda, K., Nakagawa, H., Yoshida, N., Nakamura, Y. & Matsuda, K. Regulation of protein citrullination through p53/PADI4 network in DNA damage response. *Cancer Res.* **69**, 8761–8769 (2009).
- Lan, A., Liao, X., Mo, L., Yang, C., Yang, Z., Wang, X. *et al.* Hydrogen sulfide protects against chemical hypoxia-induced injury by inhibiting ROS-activated ERK1/2 and p38MAPK signaling pathways in PC12 cells. *PLoS One* **6**, e25921 (2011).
- Tak, P. P., Smeets, T. J., Boyle, D. L., Kraan, M. C., Shi, Y., Zhuang, S. *et al.* p53 overexpression in synovial tissue from patients with early and longstanding rheumatoid arthritis compared with patients with reactive arthritis and osteoarthritis. *Arthritis Rheum.* **42**, 948–953 (1999).
- Firestein, G. S., Echeverri, F., Yeo, M., Zvaifler, N. J. & Green, D. R. Somatic mutations in the p53 tumor suppressor gene in rheumatoid arthritis synovium. *Proc. Natl Acad. Sci. USA* **94**, 10895–10900 (1997).
- Bohuslav, J., Chen, L. F., Kwon, H., Mu, Y. & Greene, W. C. p53 induces NF-kappaB activation by an IkkappaB kinase-independent mechanism involving phosphorylation of p65 by ribosomal S6 kinase 1. *J. Biol. Chem.* **279**, 26115–26125 (2004).
- Chen, L., Xu, B., Liu, L., Luo, Y., Zhou, H., Chen, W. *et al.* Cadmium induction of reactive oxygen species activates the mTOR pathway, leading to neuronal cell death. *Free Radic. Biol. Med.* **50**, 624–632 (2011).
- Cejka, D., Hayer, S., Niederreiter, B., Sieghart, W., Fuereder, T., Zwerina, J. *et al.* Mammalian target of rapamycin signaling is crucial for joint destruction in experimental

- arthritis and is activated in osteoclasts from patients with rheumatoid arthritis. *Arthritis Rheum.* **62**, 2294–2302 (2010).
- 43 Laragione, T. & Gulko, P. S. mTOR regulates the invasive properties of synovial fibroblasts in rheumatoid arthritis. *Mol. Med.* **16**, 352–358 (2010).
- 44 Valcárcel-Ares, M. N., Riveiro-Naveira, R. R., Vaamonde-García, C., Loureiro, J., Hermida-Carballo, L., Blanco, F. J. *et al.* Mitochondrial dysfunction promotes and aggravates the inflammatory response in normal human synoviocytes. *Rheumatology (Oxford)* **53**, 1332–1343 (2014).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)