Is the T-G-CT-G SNRNP70 haplotype another proof that mixed connective tissue disease is distinct from systemic lupus erythematosus and systemic sclerosis? A novel gene variant in SNRNP70 gene

A. Paradowska-Gorycka[1], A. Wajda[1], I. Stolarek[2], A. Felis-Giemza[3], M. Walczyk[3], J. Nałecz-Janik[3], M. Olesińska[3]

[1]*Department of Molecular Biology, National Institute of Geriatrics, Rheumatology and Rehabilitation, Warsaw;*
[2]*Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan;*
[3]*Department of Connective Tissue Diseases, National Institute of Geriatrics, Rheumatology and Rehabilitation, Warsaw, Poland.*

Agnieszka Paradowska-Gorycka, PhD
Anna Wajda, PhD
Ireneusz Stolarek, PhD
Anna Felis-Giemza, PhD
Marcela Walczyk, MD
Jolanta Nałecz-Janik, PhD
Marzena Olesińska, MD, PhD

Please address correspondence to:
Dr Agnieszka Paradowska-Gorycka, Department of Molecular Biology, National Institute of Geriatrics, Rheumatology and Rehabilitation, Spartańska 1, 02-637 Warszawa, Poland.
E-mail: paradowska_aga@interia.pl; agnieszka.paradowska-gorycka@spartanska.pl

Received on April 3, 2019; accepted in revised form on May 27, 2019.

Clin Exp Rheumatol 2019; 37 (Suppl. 119): S32-S40.

Key words: SNRNP70, anti-U1-70K, MCTD, phenotype, genetics

Funding: this work was supported by a grant from the Polpharma Scientific Foundation.

Competing interests: none declared.

ABSTRACT

Objective. *U1-70K, encoded by the* SNRNP70 *gene, is a key early immunogen in connective tissue disease. The aim of the study was the genetic analysis of the* SNRNP70 *gene in mixed connective tissue disease (MCTD), systemic lupus erythematosus (SLE) and systemic sclerosis (SSc) patients.*

Methods. SNRNP70 *genetic variants were detected using 3730 DNA Analyzer.* SNRNP70 *rs560811128 G/A (c.476-252 G/A), rs78616533delCT (c.475+130_475+131delCT) and rs117167710 T/C (c.393+326 T/C) variants were genotyped using the technique of sequence-specific hybridisation probe binding assays.* SNRNP70 *393_47 G/A mutation was detected using TaqMan SNP genotyping assay.*

Results. *We found one novel c.393+47G>A and three, c.476-252 G/A, c.475+130_475+131delCT and c.393+326 T/C, previously recorded variants. The present study revealed that T-G-CT-G haplotype demonstrated significantly higher frequencies in MCTD patients than in SLE and SSc patients. In MCTD patients c.475+130_475+131delCT distribution of genotype was gender-dependent and showed association with thrombo-/leukocytopenia. Mutation at position c.476-252G>A was predicted to possibly have an impact on splicing of the SNRNP70 transcript and it was present only in one MCTD patient.*

Conclusion. *Our results demonstrated that the T-G-CT-G* SNRNP70 *haplotype is another proof that MCTD may be distinct from SLE and SSc. The novel c.476-252G>A mutation in* SNRNP70 *gene created a new acceptor splice site and may potentially alert of splicing of the* SNRNP70 *transcript.*

Introduction

In 1990, Spritz *et al*. (1) introduced the sequence of *SNRNP70* gene, which is mapped on chromosome 19q13.33. *SNRNP70* gene encoding the human U1-70K protein belonging to the U1 small nuclear ribonucleoprotein (snRNP also known as U1-RNP) complex, which is a key component of the spliceosome (2-4). U1-70K protein contains an N-terminal domain, an 80 amino acid RNA binding domain (RBP) with two characteristics RNA recognition motifs (RRMs) and a C-terminal domain with two Arg-rich domains. U1-70K protein through RNP-80 motif directly binds to stem-loop I of U1-RNA as an individual molecule (5-9).

The U1-70K protein and U1-RNA involved immune cells, such as B and T cells, as well as their receptors in complex interactions leading to the production of autoantibodies, inflammation and tissue damage (8, 10). Moreover, protein components of the U1-snRNP complex are targeted by autoantibodies present in patients with systemic lupus erythematosus (SLE), systemic sclerosis (SSc) and mixed connective tissue disease (MCTD) (11-14). Anti-U1-snRNP antibodies are observed in 30–35% of SLE patients, 2-14% of SSc patients and in nearly all, if not all, MCTD patients (8, 11, 12, 15). Antibodies against U1-70K develop as first, in the early immune response, and participate in the development of other antibodies. They are detected in 75–90% of MCTD patients but in contrast, they have detected only in 20–50% of SLE patients whose sera is anti-RNP positive (8, 12, 16, 17). A prerequisite for the diagnosis of MCTD is the presence of anti-U1-RNP antibodies. In patients with high titres of these antibodies who do not

meet the criteria for MCTD diagnosis or any other connective tissue disease, MCTD usually develops within 2 years. U1-70K is a major early immunogen might be important for triggering of the immune response to U1-snRNP complex (13, 16, 17).

However, the precise role of the U1-70K protein and especially *SNRNP70* gene in the pathogenesis of MCTD, SLE and/or SSc is unknown. To verify whether genetic variation located in the SNRNP70 gene may be a factor explaining phenotype or characteristics above connective tissue diseases (CTD), we used gene sequencing for detection of all polymorphisms/mutation located in this gene.

Materials and methods

Study populations

A total of 504 individuals were selected to participate in this study. Peripheral blood samples from 135 MCTD patients, 135 SLE patients, 81 SSc patients, and 153 healthy subjects were obtained. All patients were recruited from the Department of Connective Tissue Diseases of the National Institute of Geriatrics, Rheumatology and Rehabilitation. Healthy control participants were recruited from the Regional Center for Blood Donation and Blood Treatment in Warsaw. Healthy subjects did not have a history of autoimmune and/or inflammatory disease at the time of sampling. MCTD patients, SLE patients, SSc patients, and healthy subjects were of Caucasian origin. Characteristics of all participants were summarised in Table I, where we presented only those clinical parameters which are shared by the three diseases. Objective methods were constructed based on the tools employed in SLE; an MCTD activity index (MCTD-AI) was modeled on the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) score. The most common symptoms of the active disease were: decreased number of leukocytes and/or platelets, swelling of the hands, increased of erythrocyte sedimentation rate (ESR) and/or C-reactive protein (CRP) levels, hypergammaglobulinaemia and skin rashes. MCTD Damage index (MCTD-DI) was a counterpart of Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index (SLICC/ACR DI). Damage most commonly affected skin, musculoskeletal and cardiovascular systems and lungs. Autoantibodies to Sm, Ro, La, RibP, PCNA, CENPB, Scl-70, Jo-1, His, dsDNA were measured in sera using DOT-blot tests (recomLine ANA/ENA, Mikrogen Diagnostik, Neuried, Germany). The identification of antinuclear antibodies (ANA) was performed by indirect immunofluorescence (IF) on Hep2 cell lines (Euroimmun Polska, Wroclaw, Poland), with median a titre of 1:5,840 (range 1:80 – 1:40,960). The presence of anti-U1-RNP was determined by electrochemiluminescence (ECLIA) using streptavidin-coated paramagnetic beads (UNICAP100, Phadia, Sweden). The presented clinical investigation was approved by the Research Ethics Committee of the National Institute of Geriatrics, Rheumatology and Rehabilitation in Warsaw. All participants, patients, and healthy donors gave informed written consent and research was conducted in accordance with the Declaration of Helsinki.

Table I. Patients' characteristics.

Characteristics		MCTD n=83 mean values ± SD	SLE n=97 mean values ± SD	SSc n=71 mean values ± SD
Age [years]		42.71 ± 13.40	40.58 ± 12.73	55.06 ± 13.47
Disease duration [years]		9.75 ± 8.52	3.85 ± 5.97	4.99 ± 7.50
Gender		67 (81%) female 16 (19%) male	90 (93%) female 7 (7%) male	53 (75%) female 18 (25%) male
Damage index [median range (min-max)]		MCTD-DI 5 (0-68)	SLICC 4 (0-17)	2 (0-15)
Activity index [median range (min-max)]		MCTD-AI 9.4 (0-32)	SELENA_SLEDAI 4 (0-17)	0.5 (0-9)
ESR [mm/h]		27.58 ± 23	32.62 ± 26.89	19.37 ± 16.54
CRP [mg/L]		11.63 ± 14.80	25.7 ± 57.62	8.28 ± 11.21
			n (%)**	
anti-dsDNA		12 (14%)	58 (60)	1 (1%)
anti-SSA	*Ro60*	8 (10%)	37 (38%)-SSA	1 (1%)
	Ro52	23 (28%)		1 (1%)
anti-SSB		1 (1%)	12 (12%)	0 (0%)
anti-SM	*SmB*	27 (33%)	14 (14%)	0 (0%)
	SmD	5 (6%)		0 (0%)
anti-Rib		3 (4%)	10 (10%)	0 (0%)
anti-Jo		2 (2%)	0 (0%)	0 (0%)
anti-histon		12 (14%)	22 (23%)	1 (1%)
anti-Scl-70		2 (2%)	1 (1%)	5 (7%)
anti-U1-RNP		82 (99%)	14 (14%)	0
anti-70K		60 (72%)	-	-
anti-A		68 (82%)	-	-
anti-C		64 (77%)	-	-
Sjögren		18 (22%)	16 (16%)	-
Raynaud's phenomenon		80 (96%)	-	47 (66%)
Facial erythema		36 (43%)	44 (45%)	-
arthritis		78 (94%)	62 (64%)	24 (34%)
Pulmonary fibrosis		28 (34%)	5 (5%)	21 (30%)
Leukopenia/trombocytopenia		51 (61%)	52 (50%)	-

MCTD: mixed connective tissue disease; SLE: systemic lupus erythematosus; SSc: systemic sclerosis; SD: standard deviation; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; MCTD-DI: MCTD damage index; SLICC: Systemic Lupus International Collaborating Clinics; MCTD-AI: MCTD activity index; SELENA-SLEDAI: Systemic Lupus Erythematosus Disease Activity Index.

Mixed connective tissue disease patients

MCTD patients met Alarcon-Segovia and Villarreal classification criteria, which were chosen due to their very high specificity for MCTD. Patients with MCTD, who met the classification criteria for two connective tissue diseases at the same time were excluded from the study. Since guidelines to evaluate the clinical state of MCTD patients have not been yet established, a set of objec-

tive and subjective methods to assess the damage infelicity by MCTD as well as disease activity had to be devised specifically for the purpose of this study. The median value of MCTD-AI was 9.4 points (lowest 0, highest 32 points). MCTD-AI at or below 6 points reflected a cut-off point for low disease activity. Patients with a value above 6 points had symptoms of organ involvement. The most common symptoms of the active disease were: decreased number of leukocytes and/or platelets, swelling of the hands, increased of ESR and/or CRP levels, hypergammaglobulinaemia and skin rashes. Median of MCTD-DI was 5 (lowest value 0, highest 68). The clinical picture of MCTD was variable. Some symptoms such as skin lesions typical for SLE, swelling of the hands, arthropathy, inflammation and involvement of the muscle were mostly occurring around the onset of the disease. Furthermore, skin lesions typical for SS, the involvement of the lungs or oesophagus were diagnosed at various times during the course of MCTD. Virtually all MCTD patients presented with Raynaud's phenomenon right at the onset of the disease. Pulmonary arterial hypertension (PAH) was detected in 32% of the patients at the time of the study. The presence of PAH is routinely screened in MCTD patients. Antinuclear antibodies in titre >1:320 (ANA) were detected in 88% of MCTD patients. 99% of MCTD patients had anti-U1-RNP antibodies: anti-70K was detected in 72 % of patients, anti-A in 82 % of patients and anti-C in 77 % of patients.

Systemic lupus erythematosus patients

SLE patients met the American College of Rheumatology/Systemic Lupus International Collaborating Clinics (ACR/SLICC) 2012 classification criteria. The mean disease duration of SLE patients was 4 years. Disease activity assessed by SELENA-SLEDAI index was 4 points (lowest 0, highest 12), while damage index assessed by the SLICC index was 4 points (lowest 0, highest 17). ANA in titre >1:320 were detected in 48% of SLE patients, furthermore, anti-U1-RNP antibodies were observed in 14% of SLE patients.

Systemic sclerosis patients

SSc patients met the ACR/European League Against Rheumatism (EULAR) 2013 classification criteria. The mean disease duration of SSc patients was 2 years (range 0-16 years). Disease activity was measured by the European Scleroderma Research Group Activity Index EScSG-AI. The disease activity score in SSc patients were 0,5 points (lowest 0, highest 9). The damage index assessed by the Damage Index was 2 points (lowest 0, highest -15). ANA in titre >1:320 were detected in 75% of SSc patients, while anti-U1-RNP antibodies do not were observed in SSc patients.

DNA extraction

Genomic DNA was extracted from peripheral blood lymphocytes using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) in accordance with the manufacturer's instructions.

Genetic analysis

DNA of 20 individuals (7 MCTD, 3 SLE, 3 SSc and 7 controls) was used for mutations/polymorphisms detection in the *SNRNP70* gene. Genotyping was performed using 3730 DNA Analyzer (Thermo Fisher Scientific); this analysis was made by a private company (it was a commissioned service). SNRNP70 rs560811128 G/A (c.476-252 G/A), rs78616533delCT (c.475+130_475+131delCT) and rs117167710 T/C (c.393+326 T/C) variants were directly genotyped using the technique of sequence-specific hybridisation probe binding assays in LightCycler 480 II instrument (Roche, Germany). SNRNP70 393_47 G/A mutation was performed by Quant Studio 5 detection system (Life Technologies).

Real-time PCR

rs560811128 G/A (c.476-252 G/A), rs78616533delCT (c.475+130_475+131delCT) and rs117167710 T/C (c.393+326 T/C) were directly genotyped using the technique of sequence-specific hybridisation probe binding assays in LightCycler 480 II instrument (Roche, Germany). LightSNiP assay (contains primer and probe, TIBMOLBIO, Germany) and LightCycler 480 Genotyping Master (Roche Diagnostics GmbH) was used for the reaction. The reaction was performed in a total volume of 10ul and contains 50 ng of DNA, 1 ul of Genotyping Master Mix and 1 µl of LightMix. The PCR conditions were: 10 min at 95ºC followed by 45 cycles of 95ºC at 10 s, 10 s at 60ºC and 15 s at 72ºC. Melting curve analysis was performed with an initial denaturation at 95°C for 30 seconds, 2 min at 40°C, slow heating to 75°C with a ramping rate of 0.2°C/s and continuous fluorescence detection. Wild-type, heterozygous and polymorphic alleles were detected by specific melting temperature (Tm) of the obtained amplicons. Tm values for the rs560811128 G/A were 62.93°C for allele G and 68.99°C for allele A, for the rs78616533delCT were 55.57°C for CT allele and 61.06°C for allele with delCT and for the rs117167710 T/C were 57.96°C for allele T and 65.55°C for allele C.

Determination of genetic variants at position 393_47 G/A was performed by Quant Studio 5 detection system (Life Technologies). The custom TaqMan SNP Genotyping Assay was used. The reaction was performed in a total volume of 10ul and contains: 10 ng of DNA, 5 ul of TaqMan Genotyping Master Mix 2x (Applied Biosystems) and 0,5 µl of TaqMan genotyping assay mix 20x (Applied Biosystems). Thermal cycle conditions were as follows: denaturation at 95°C for 10 min, followed by 40 cycles of denaturation at 92°C for 15 s, and annealing and extension at 60°C for 1 min.

Negative controls and duplicate randomly selected samples were included to check the accuracy of genotyping.

Bioinformatics analysis

The potential effects of c.393+47G>A and c.476-252G>A, T.475+130_475+131del and c.393+326 T>C mutations on the splicing sites modifications in the *SNRNP70* gene were studied. Human Splicing Finder (v. 3.1) was used to analyse splicing sites in the ENST00000221448 transcript of the *SNRNP70* gene.

Statistical analysis

The statistical analysis was conducted

using Statistica 10 Software Package (StatSoft, Poland) and the R program (R Development Core Team (2008). Clinical and demographic characteristic of MCTD, SLE and SSc patients were presented as mean and standard deviation. Categorical variables were presented as percentages. The ORs and 95% CIs, adjusted by gender and age, were calculated with an unconditional logistic regression. Differences in genotypes, adjusted by gender and age, between examined groups (OR and 95% confidence intervals, *p*-value) was calculated with logistic regression. 4 genetic models (codominant, dominant and recessive) were analysed. Shapiro-Wilk test was used to determine normality distribution of continuous variables. The association between genotype and clinical parameters was analysed using Mann-Whitney U-test or unpaired t-test. In the case of categorical variables analysis was conducted using χ^2 with Yates correction. Statistical significance was set at $p<0.05$. Hardy-Weinberg equilibrium (HWE) was performed using an online calculator (Michael H. Court [2005-2008]). Linkage disequilibrium (LD), the coefficient (D´ and r2) for haplotypes and their frequencies were performed using SHEsis software (http://shesisplus.bio-x.cn/SHEsis.html) (18, 19). Clinical and demographic characteristic of MCTD, SLE and SSc patients were presented as mean and standard deviation. Categorical variables were presented as percentages. The ORs and 95% CIs, adjusted by gender and age, were calculated with an unconditional logistic regression. Differences in genotypes, adjusted by gender and age, between examined groups (OR and 95% confidence intervals, *p*-value) was calculated with logistic regression. Shapiro-Wilk test was used to determine normality distribution of continuous variables.

Results

Identification of novel genetic variant (mutation) in SNRNP70 gene

To identify possible polymorphisms/mutations in the *SNRNP70* gene we sampled 20 subjects from Poland (7 MCTD, 3 SLE, 3 SSc and 7 controls). Genetic variants were identified in 6 out

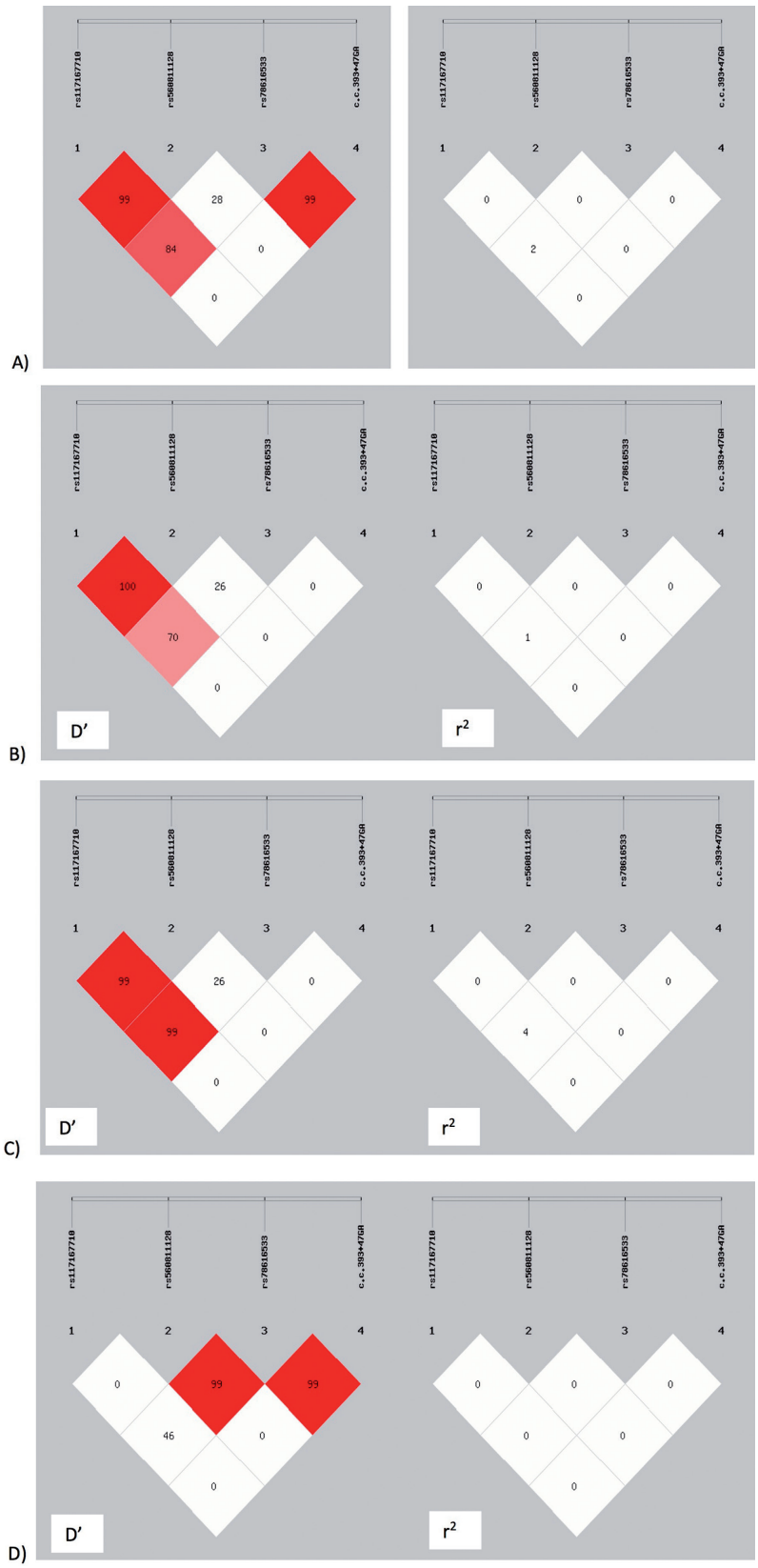


Fig. 1. Linkage disequilibrium (LD) map of the SNRNP70 gene. The plot illustrates pairwise LD between genetic variants based on D' (left) and R-squared (r2; right) scores. Values approaching zero indicate absence of LD, and those approaching 100 indicate complete LD. D' can be very important for population genetics, but for substituting or tagging one SNP for another r2 is the important measure. **A)** LD between all patients and controls; **B)** LD between MCTD patients and SLE patients; **C)** LD between MCTD patients and SSc patients; **D)** LD between SLE patients and SSc patients.

of 20 subjects. In those six subjects we found four different genetic variants located in introns: one novel genetic variant and three variants, rs560811128 G/A (c.476-252 G/A), rs78616533delCT (c.475+130_475+131delCT) and rs117167710 T/C (c.393+326 T/C), previously recorded in the public database, but not reported by previous studies. The genetic variants at position c.393+47G>A, located in intron 4, was *de novo* discovery in our investigation and it is not recorded in the public database. In our study groups, we do not find genetic variants located in exons and/or in the 3'- or 5' untranslated region (5'-UTR) of the *SNRNP70* gene.

Distribution of the SNRNP70 genetic variants in MCTD, SLE and SSc patients and healthy subjects

To investigate if the genetic variants, rs560811128, rs78616533, rs117167710 and c.393+47G>A of *SNRNP70* gene are associated with MCTD, SLE and/or SSc we genotyped a population of 504 Polish Caucasian individuals. The genotyping success was greater than 73% in all cases. We observed that only one genetic variant at position rs78616533delCT belongs to the polymorphic loci (Suppl. Table S1 in the supplementary file). However, there were no significant differences in genotype distribution between patients and controls in all genetic models (dominant and recessive models not shown). Genotype distributions of rs78616533delCT were in Hardy-Weinberg Equilibrium (HWE) in all study groups. Furthermore, other examined genetic variants, rs560811128, rs117167710 and c.393+47G>A, do not belong to the polymorphic loci because their frequencies are extremely low in the population; these genetic variants should be interpreted as mutations. The rs117167710 TC genotype was present in only 3 patients with MCTD, 3 patients with SLE, 6 patients with SSc and in 7 healthy controls. The rs117167710 CC genotype was not found in our study group. The rs560811128 GA genotype was present only in one MCTD patient. Other subjects (patients and controls) were wild-type homozygous for this genetic variant (rs560811128 GG genotype). c.393+47GA genotype was found only in one healthy individual, other subjects have c.393+47GG genotype (wild-type genotype).

Linkage disequilibrium (LD) between SNRNP70 genetic variants

LD was calculated in all possible variants (Fig. 1). We observed a very high D' between rs117167710 and rs560811128, between rs117167710 and rs78616533 as well as between rs78616533 and c.393+47 G/A (D'=99, D'=0.84 and D'=0.99, respectively). This high D' suggested that examined genetic variants are in a strong LD. However, D' value is not suitable for LD prediction. r^2 is preferred when the focus is on the predictability of one polymorphism given the other and that's way r^2 score is much better for predicting LD. Our analysis has shown that r^2 score is very low ($r^2 = 0.0$) which suggests that examined genetic variants are not good predictors of each other.

LD calculated between MCTD patients and SLE patients as well as MCTD patients and SSc patients give a high D' between rs117167710 and rs560811128 and between rs117167710 and rs78616533 (Fig. 1). Furthermore, LD calculated between SLE patients and SSc patients give a very high D" between rs560811128 and rs78616533 (D'=0.99) and between rs78616533 and c.393+47 G/A (D'=0.99). r^2 value is 0.0 in all our correlation. Our analysis demonstrated that studied *SNRNP70* genetic variants cannot substitute each other and that they are in indistinct LD.

SNRNP70 haplotype and risk of MCTD, SLE and SSc.

Haplotypes with frequency <0.03 were ignored. The most frequent haplotype identified in our patients with connective tissue diseases (Suppl. Table S2) was T-G-C-G with a frequency of 45%. Furthermore, in controls the most frequent haplotype was T-G-CT-G with a frequency of 39%. Testing the association of *SNRNP70* haplotypes with our diseases risk indicated that T-G-C-G haplotype may be associated with increased CTD risk in the Polish population (OR=1.372, 95% CI=1.008-1.867, p=0.04). Secondly, we also compared the frequency of SNRNP70 haplotypes between individual groups of patients (Table II). We found that T-G-CT-G haplotype demonstrated significantly higher frequencies in MCTD patients than in patients with SLE and SSc (both p=0.0002). T-G-C-G haplotype was significantly more frequent in SLE (49%) and SSc (50%) patients compared to patients with MCTD (32%) (p=0.0003 and p=0.002, respectively). Above correlation was not observed between SLE patients and SSc patients. The presented results may be another proof that MCTD is a separate disease, distinct from SLE and SSc.

Correlation between rs78616533delCT polymorphisms and MCTD, SLE and SSc phenotype

In the present study, we also analysed the clinical data from the MCTD, SLE and SSc patients to explore if the SNRNP70 rs78616533delCT polymorphism may be associated with clinical phenotypes. We observed that in MCTD patients the distribution of genotype was gender-dependent (Table III). The rs78616533 CT/CT genotype was more frequent in the female, while the rs78616533 -/CT genotype was more frequent in male (p=0.03). In MCTD patients with rs78616533 CT/CT genotype thrombo-/ leukocytopenia was more frequently observed compared with MCTD patients with rs78616533 -/CT genotype (p=0.005). There were no significant differences in the distribution of other MCTD clinical parameters between rs78616533 CT/CT genotype and rs78616533 -/CT genotype. In patients with SLE and SSc, we do not observe an association between *SNRNP70* rs78616533delCT SNP and clinical parameters (Suppl. Table S3 and S4).

Splice site modifications in the SNRNP70 gene

Human Splicing Finder (HSF, v. 3.1) was used to analyse splicing sites in the ENST00000221448 transcript of the *SNRNP70* gene. Summary of the mutational effects on the transcript splicing is presented in Table IV. Genetic variants at position rs560811128 G/A (c.476-252G>A) affects the acceptor-

Table II. *SNRNP70* haplotypes created between patients (MCTD, SLE, SSc). Loci chosen for hap-analysis: rs117167710 T/C, rs560811128 G/A, rs78616533 delCT, c.393+47 G/A.

Haplotype	MCTD 2n=184 (%)	SLE 2n=256 (%)	*p**	OR (95% CI)
C-G-delCT-G	3.98 (2)	0.53 (0.1)	-	-
C-G-CT-G	0.07 (0)	0.99 (0.4)	-	-
C-G-C-G	5.95 (3)	1.49 (0.5)	**0.03**	0.172 (0.028 to 1.046)
T-G-delCT-G	23.02 (13)	39.47 (15)	0.4	1.249 (0.718 to 2.174)
T-G-CT-G	91.93 (50)	86.01 (34)	**0.0002**	0.486 (0.329 to 0.719)
T-G-C-G	58.05 (32)	126.51 (49)	**0.0003**	2.075 (1.393 to 3.090)
Haplotype	SSc 2n = 132 (%)	MCTD 2n= 184 (%)	*p**	OR (95% CI)
C-G-delCT-G	0.00 (0)	3.98 (2)	-	-
C-G-CT-G	4.99 (4)	0.07 (0)	**0.02**	96.557 (6.293 to 1481.562)
C-G-C-G	0.01 (0)	5.95 (3)	0.06	0.001 (0.000 to 0.025)
T-G-delCT-G	21.00 (16)	23.02 (13)	0.4	1.282 (0.676 to 2.430)
T-G-CT-G	40.01 (30)	91.93 (50)	**0.0002**	0.412 (0.257 to 0.661)
T-G-C-G	65.99 (50)	58.05 (32)	**0.002**	2.083 (1.311 to3.310)
Haplotype	SSc 2n = 132 (%)	SLE 2n= 256 (%)	*p**	OR (95% CI)
C-G-delCT-G	0.00 (0)	0.53 (0.1)	-	-
C-G-CT-G	4.99 (4)	0.99 (0.4)	**0.01**	10.035 (1.619 to 62.189)
C-G-C-G	0.01 (0)	1.49 (0.5)	-	-
T-G-delCT-G	21.00 (16)	39.47 (15)	0.9	1.023 (0.575 to 1.822)
T-G-CT-G	40.01 (30)	86.01 (34)	0.5	0.844 (0.537 to 1.328)
T-G-C-G	65.99 (50)	126.51 (49)	1	1.000 (0.656 to1.523)

*Fisher's test. *p*-values in bold face are considered significant. MCTD: mixed connective tissue disease; SLE: systemic lupus erythematosus; SSc: systemic sclerosis; OR: odds ratio, CI: Confidence interval.

Table III. Association of SNRNP70 rs78616533 with MCTD clinical parameters.

Characteristics rs78616533	rs78616533 CT/CT		rs78616533 -/CT	*p*-value
		n=57	n=25	
Age [years]		41 (18-70)	44 (21-88)	0.56[a]
Disease duration [years]		8 (0.92-52)	8 (0.8-35)	0.26[a]
Gender	*female*	50 (88%)	16 (64%)	0.03[b]
	male	7 (12%)	9 (36%)	
ESR-OB [mm/h]		19 (8-68)	23 (5-95)	0.88[a]
CRP [mg/L]		7 (3-24)	8.5 (4-71)	0,60[a]
		n (%)		
anti-U1-RNP		56 (98%)	25 (100%)	0,69[b]
anti-dsDNA		9 (16%)	3 (12%)	0,91[b]
anti-SSA	*Ro60*	5 (9%)	2 (8%)	0,78[b]
	Ro52	16 (28%)	6 (24%)	0,91[b]
anti-SSB		0 (0%)	1 (4%)	0,66[b]
anti-SM	*SmB*	19 (33%)	8 (32%)	*0,84[b]*
	SmD	4 (7%)	*1 (4%)*	*0,98[b]*
anti-Rib		1 (2%)	*2 (8%)*	*0,45[b]*
antiJo		1 (2%)	1 (4%)	*0,86[b]*
anti-histon		8 (14%)	4 (16%)	*0,97[b]*
anti-Scl70		0 (0%)	*2 (8%)*	*0,17[b]*
anti-CENP B		0 (0%)	0 (0%)	
Sjögren syndrome		12 (21%)	5 (20%)	0,96[b]
Raynaud's phenomenon		56 (98%)	23 (92%)	0,45[b]
facial erythema		23 (40%)	13 (52%)	0,46[b]
arthritis		54 (98%)	23 (92%)	0,98[b]
pulmonary fibrosis		22 (39%)	6 (24%)	0,30[b]
thrombo-/ leucocytopenia		41 (72%)	9 (36%)	0,005[b]

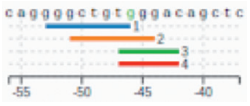
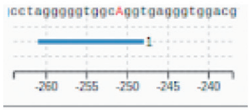
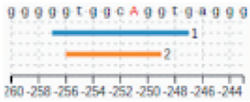
sites at the 3' end of intron 7 and may cause activation of an intronic cryptic acceptor site, which could potentially have an impact on splicing of the transcript (Table IV and V). Genetic variants c.393+47G>A alters an intronic Exonic Splicing Silencers (ESS) site and creates an intronic Exonic Splicing Enhancers (ESE) site, but these changes have no impact on the splicing. ESE sites are specific short nucleotide sequences targeted essentially by Serine/Arginine-rich (SR) proteins leading to exon definition. Conversely, ESS sites help the spliceosome to ignore pseudoexons and they act as binding sites for proteins which then promote exon exclusion (20). Also no potential branch points, enhancer or silencer motifs were created by this genetic variant. Genetic variants at position rs78616533delCT (c.475+130_475+131delCT) create an intronic ESE site, however, the changes probably have no impact on the splicing. This genetic variant also removes a CTCF binding site (Table IV). Genetic variants at position rs117167710 (c.393+326 T>C) alters an intronic ESS site without impact on the splicing.

Discussion

The main results of the present study are as follows: 1) novel mutation, c.393+47G>A, was found in *SNRNP70* gene, 2) *SNRNP70* haplotypes may be another proof that MCTD is distinct from SLE and SSc, 3) rs560811128 G/A (c.476-252G>A) mutation was present only in one MCTD patient; this patient has a heterozygous variant of this mutation, 4) mutation at position c.476-252G>A was predicted to have an impact on splicing of the *SNRNP70* transcript.

Nowadays we know that components of U1-snRNP complex are important for triggering immune responses and that anti-U1RNP may play a pathogenic role in CTD pathogenesis. Previous data showed that anti–U1-70K autoantibodies have a strong association with MCTD (21). These autoantibodies are also present in patients with SLE (15, 22) and patients with SSc (15, 23). We detected anti-U1RNP antibodies in all MCTD patients and in 14% of SLE patients what is consistent with the previ-

Table IV. Summary of the mutational effects on the ENST00000221448 transcript splicing.

Mutation	Predicted signal	Prediction algorithm	cDNA Position	Interpretation
c.393+47G>A	ESS Site broken	1 - Sironi *et al*. - Motif 2 2 - ESR Sequences from Goren *et al*. 3 - IIEs from Zhang *et al*.		Alteration of an intronic ESS site. Probably no impact on splicing.
	New ESE Site	1 - ESE-Finder - SC35 2 - PESE Octamers from Zhang & Chasin 3 - RESCUE ESE Hexamers 4 - EIEs from Zhang *et al*.		Creation of an intronic ESE Site. Probably no impact on splicing.
c.476-252G>A	New Acceptor Site	1 - HSF Matrices		Activation of an intronic cryptic acceptor site. Potential alteration of splicing.
	ESS Site broken	1 - IIEs from Zhang *et al*. 2 - Sironi et al. - Motif 2		Alteration of an intronic ESS site. Probably no impact on splicing.
T.475+130_475+131del	New ESE site	1 - ESE-Finder - SF2/ASF 2 - ESE-Finder - SF2/ASF(Ig) 3 - ESR Sequences from Goren *et al*. 4 - ESE-Finder - SC35		Creation of an intronic ESE Site. Probably no impact on splicing. Removal of the CTCF binding site
c.393+326 T>C	New ESE site	1 - ESE-Finder - SRp40 2 - ESR Sequences from Goren *et al*. 3 - ESE-Finder - SRp55 4 - ESE-Finder - SF2/ASF 5 - ESE-Finder - SF2/ASF(Ig)		Creation of an intronic ESE Site. Probably no impact on splicing.

ESE: Exonic splicing enhancer; ESS: Exonic splicing silencer.

Table V. Potential splice sites created by the c.476-252G>A in the ENST00000221448 transcript.

Sequence Position	cDNA Position	Splice site type	Motif	New splice site	Wild Type	Mutant	If cryptic site use, exon length variation	Variation (%)
39	-262	Acceptor	aggggggtggcgggt	aggggggtggcagGT	47.8	76.75	+250	New site +60.56
41	-260	Donor	ggggtggcg	GGGgtggca	69.94	69.54	+248	-0.57
44	-257	Acceptor	gtggcgggtgaggg	gtggcaggtgagGG	69.37	69.42	+245	+0.07
48	-253	Donor	cgggtgagg	CAGgtgagg	92.05	96.91	+241	+5.28

ous study. But in contrast, we did not detect anti-U1RNP antibodies in SSc patients. Among all diagnosed anti-U1-RNP antibodies the strongest relationship with MCTD revealed anti-U1-70K antibodies (16, 24). These antibodies were detected in 72% of our MCTD patients. Anti-U1-70K serum levels increased with exacerbation of the disease (25), but decreased in periods of disease remission in patients with MCTD (24).

To date, no studies have focused on the genetic analysis of *SNRNP70* gene. In the present study, using gene sequencing method, we found four different genetic variants located in *SNRNP70* gene. All detected genetic variants are located in the introns. We observed that only one genetic variant at position rs-78616533delCT belongs to the polymorphic loci. Furthermore, other examined genetic variants, rs560811128, rs117167710 and c.393+47G>A, do not belong to the polymorphic loci because their frequencies are extremely low in the population; these genetic variants should be interpreted as mutations. Although we do not have clinical information from all collected patients, for genotyping analysis we decided to use all collected materials to increase the power of the study.

The genetic variants at position c.393+47G>A, which is located in intron 4 of the *SNRNP70* gene, it is de novo discovery mutation and it is not recorded in the public database. This mutation was identified only in one study subject from healthy controls. This is a 57-year old woman and she has a heterozygous variant for this mutation. Bioinformatic analysis showed that c.393+47G>A alters an intronic ESS site and creates an intronic ESE site. Likely the changes have no impact on the splicing. Also no potential branch points, enhancer or silencer motifs were created by the mutation.

U1-RNP is a highly conserved nuclear RNA-protein complex involving in the mRNA splicing, which is an essential

step in the gene expression (26). Moreover, U1-RNP is also involved in other cellular processes such as mRNA degradation or transcription initiation (27). U1-RNP binding to the 5' exon-intron junction of pre-mRNA is a central step in the formation of the early splicing complex (28). It is generally accepted that mutations affected pre-mRNA splicing are significant players in the occurrence of genetic diseases and they have an impact on clinical practice (27). The most common consequence of these mutations is skipping of the exon, activation of cryptic or *de novo* splice sites and retention of the introns in mRNA (29). In the present study, we found a mutation rs560811128 G/A (c.476-252G>A) located in intron 6 of the *SNRNP70* gene that was present only in one MCTD patient. Mutation c.476-252G>A created a new acceptor splice site, lead to activation of an intronic cryptic acceptor site and potentially alert of splicing of the *SNRNP70* transcript. Cryptic splice sites are used only when the use of a natural splice site is disrupted by mutation elsewhere in the gene and they are not detectably used in wild-type context. MCTD patient with c.476-252G>A mutation has a heterozygous variant, it was w women and she has the most active form of MCTD from all patients participating in the study. She has antinuclear antibodies (ANA) in titre 1: 163,800 and she has anti-U1RNP, anti-U1-70K as well as anti-U1-A antibodies. We suggested that this mutation may lead to the creation of a protein product with different properties from a naturally occurring product, which could disturb the interaction between U1-RNP and other factors of the spliceosome in the splicing process. And in consequence c.476-252G>A mutation can modify the phenotype of MCTD and strongly suggests a significant role for splicing in phenotypic differences between individuals. However, our hypothesis requires further testing.

The concept of MCTD as a separate new rheumatic disease syndrome has persisted for more than four decades, but it is still controversial. In the present study, we observed that T-G-CT-G haplotype was more frequent in MCTD patients (50%) than in patients with SLE (34%) and SSs (30%). While T-G-C-G haplotype was significantly more frequent in SLE (49%) and SSc (50%) patients compared to patients with MCTD (32%). These results may be another proof that MCTD is a separate connective tissue disease, distinct from SLE and SSc.

In summary, we identified a new c.393+47G>A substitution encoded in *SNRNP70*, which alters an intronic ESS site and creates an intronic ESE site, and we show that c.476-252G>A mutation created a new acceptor splice site and potentially alert of splicing of the *SNRNP70* transcript. Moreover, our data highlight that the T-G-CT-G *SNRNP70* haplotype is another proof that MCTD may be distinct from SLE and SSc.

Acknowledgments

The technical assistance of Wieslawa Frankowska and Teresa Golaszewska is gratefully acknowledged. We are also grateful to all of the MCTD, SLE and SSc patients and healthy subjects whose cooperation made this study possible.

References

1. SPRITZ RA, STRUNK K, SUROWY CS, MOHRENWEISER HW: Human Ul-70K Ribonucleoprotein Antigen Gene: Organization, Nucleotide Sequence, and Mapping to Locus 19q13.3. *Genomics* 1990; 8: 371-9.
2. GREINDINGER EL, FOECKING MF, RANATUNGA S, HOFFMAN RW: Apoptotic U1- 70 kd is antigenicall distinct from the intact from of the U1- 70-kd molecule. *Arthritis Rheum* 2002; 46: 1264-9.
3. HOF D, CHEUNG K, DE ROOIJ DJ *et al*.: Antibodies specific for apoptotic U1-70K are superior serological markers for mixed connective tissue disease. *Arthritis Res Ther* 2005; 7: R302-R309.
4. OSNES LT, NAKKEN B, BODOLAY E, SZODORAY P: Assessment of intracellular cytokines and regulatory cells in patients with autoimmune diseases and primary immunodeficiencies – novel tool for diagnostics and patient follow-up. *Autoimmun Rev* 2013; 12: 967-71.
5. NELISSEN RL, WILL CL, VAN VENROOIJ WJ, LÜHRMANN R: The association of the U1-specyfic 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. *EMBO* 1994; 13: 4113-25.
6. ROMAC JMJ, GRAFF DH, KEENE JD: The U1 small nuclear ribonucleoprotein (snRNP) 70K protein is transported independently of U1 snRNP particles via a nuclear localization signal in the RNA-binding domain. *Mol Cell Biol* 1994; 14: 4662-70.
7. RICHTER C, SIMON T, ASEN I, BRENNER-WEISS G, HUBBUCH J: Soluble full-length expression and characterization of snRNP protein U1-68/70K. *Protein Expr Purif* 2014; 104: 65-70.
8. KATTAH NH, KATTAH MG, UTZ PJ: The U1-snRNP complex: structural properties relating to autoimmune pathogenesis in rheumatic diseases. *Immunol Rev* 2010; 233: 126-45.
9. MURAKAMI A, KOJIMA K, OHYA K, IMAMURA K, TAKASAKI Y: A new conformational epitope generated by the binding of recombinant 70-kD protein and U1 RNA to anti-U1 RNP autoantibodies in sera from patients with mixed connective tissue diseases. *Arthritis Rheum* 2002; 46: 3273-82.
10. PARADOWSKA-GORYCKA A: U1-RNP and Toll-like receptors in the pathogenesis of mixed connective tissue disease. Part II. Endosomal TLRs and their biological significance in the pathogenesis of mixed connective tissue disease. *Reumatologia* 2015; 53 :1-9.
11. SWANTON J, ISENBERG D: Mixed connective tissue disease: still crazy after all these years. *Rheum Dis Clin N Am* 2005; 31: 421-36.
12. GREIDINGER EL, HOFFMAN RW: Autoantibodies in the pathogenesis of mixed connective tissue disease. *Rheum Dis Clin N Am* 2005; 31: 437-50.
13. DEGEN WGJ, PIEFFERS M, WELIN-HENRIKSSON E, VAN DEN HOOGEN FH, VAN VENROOIJ WJ, RAATS JM: Characterization of recombinant human autoantibody fragments direct toward the autoantigenic U1-70K protein. *Eur J Immunol* 2000; 30: 3029-38.
14. KEITH MP, MORATZ C, TSOKOS GC: Amti-RNP immunity: implications for tissue injury and the pathogenesis of connective tissue disease. *Autoimmune Rev* 2007; 6: 232-6.
15. SOBANSKI V, GIOVANNELLI J, LYNCH BM *et al*.: Characteristics and survival of anti-U1 RNP antibody-positive patients with connective tissue disease-associated pulmonary arterial hypertension. *Arthritis Rheumatol* 2016; 68: 484-93.
16. CIANG NC, PEREIRA N, ISENBERG DA: Mixed connective tissue disease-enigma variations? *Rheumatology* (Oxford) 2017; 56: 326-33.
17. HOF D, RAATS JMH, PRUIJN GJM: Apoptotic modification affects the autoreactivity of the U1-snRNP autoantigens. *Autoimmune Rev* 2005; 4: 380-8.
18. SHI YY, HE L: SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res* 2005; 15: 97-8.
19. LI Z, ZHANG Z, HE Z *et al*.: A partition-ligation-combination-subdivision EM algorithm for haplotype inference with multiallelic markers: update of the SHEsis (http://analysis.bio-x.cn). *Cell Res* 2009; 19: 519-23.
20. DESMET FO, HAMROUN D, LALANDE M, COLLOD-BÉROUD G, CLAUSTRES M, BÉROUD C: Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Research* 2009; 37: e67.
21. TAKEDA Y, WANG GS, WANG RJ *et al*.:

Enzyme-linked immunosorbent assay using isolated (U) small nuclear ribonucleoprotein polypeptides as antigens to investigate the clinical significance of autoantibodies to these polypeptides. *Clin Immunol Immunopathol* 1989; 50: 213-30.

22. DIMA A, JURCUT C, BAICUS C: The impact of anti-U1-RNP positivity: systemic lupus erythematosus versus mixed connective tissue disease. *Rheumatol Int* 2018; 38: 1169-78.
23. MEHRA S, WALKER J, PATTERSON K, FRITZLER MJ: Autoantibodies in systemic sclerosis. *Autoimmun Rev* 2013; 12: 340-54.
24. BURDT MA, HOFFMAN RW, DEUTSCHER SL, WANG GS, JOHNSON JC, SHARP GC: Long-term outcome in mixed connective tissue disease: longitudinal clinical and serologic findings. *Arthritis Rheum* 1999; 42: 899-909.
25. HABETS WJ, DE ROOIJ DJ, HOET MH, VAN DE PUTTE LB, VAN VENROOIJ WJ: Quantitation of anti-RNP and anti-Sm antibodies in MCTD and SLE patients by immunoblotting. *Clin Exp Immunol* 1985; 59: 457-466.
26. WEST S: The increasing functional repertoire of u1 snRNA. *Biochem Soc Trans* 2012; 40: 846-9.
27. BURATTI E, BARALLE D: Novel roles of U1 snRNP in alternative splicing regulation. *RNA Biol* 2010; 7: 412-9.
28. KONDO Y, OUBRIDGE C, VAN ROON AM, NAGAI K: Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle reveals the mechanism of 5' splice site recognition. *eLIFE* 2015; 4: e04986.
29. VORECHOVSKY I: Aberrant 30 splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res* 2006; 34: 4630-41.