

Berthold Hoppe · Guido A. Heymann ·
Farzaneh Tolou · Holger Kiesewetter ·
Thomas Doerner · Abdulgabar Salama

High variability of peptidylarginine deiminase 4 (PADI4) in a healthy white population: characterization of six new variants of PADI4 exons 2–4 by a novel haplotype-specific sequencing-based approach

Received: 18 June 2004 / Accepted: 19 July 2004 / Published online: 25 August 2004
© Springer-Verlag 2004



BERTHOLD HOPPE received his M.D. degree at the Institute of Biochemistry and Molecular Biology, Freie Universität Berlin, Germany. He is presently a postdoctoral fellow at the Institute of Transfusion Medicine (Charité Campus Virchow-Klinikum, Berlin, Germany). His main scientific interests include molecular hemostaseology and immunology.



ABDULGABAR SALAMA received his M.D. degree at the Institute of Clinical Immunology and Blood Transfusion, University of Gießen, Germany. He is presently Head Professor of Immunohematology and Blood Transfusion at the Institute of Transfusion Medicine (Charité Campus Virchow-Klinikum, Berlin, Germany). His main scientific interests include immunohematology, complement, and transplantation immunology.

The nucleotide sequence data presented in this report have been submitted to the EMBL nucleotide sequence database and were assigned the accession numbers AJ715932–AJ715938. Regarding the localization of the exonic and intronic SNPs, their positions are based on the respective nucleotides in sequences NM_012387 and NT_034376.1.

B. Hoppe (✉) · G. A. Heymann · F. Tolou · H. Kiesewetter ·
T. Doerner · A. Salama
Institute of Transfusion Medicine, Campus Virchow Klinikum,
Charité, Universitätsmedizin Berlin,
Augustenburger Platz 1, 13353 Berlin, Germany
e-mail: berthold.hoppe@charite.de
Tel.: +49-30-450553089
Fax: +49-30-450553988

Abstract Seven single nucleotide polymorphisms (SNPs) of the peptidylarginine deiminase 4 (PADI4) gene have recently been reported to be strongly associated with rheumatoid arthritis in Japanese individuals. These SNPs are located in or close to exons 2–4 of PADI4 and are organized in at least four different haplotypes. However, a detailed sequencing-based characterization of the PADI4 gene in other populations is still lacking. We therefore analyzed exons 2–4 of the PADI4 gene in 102 healthy white Germans individuals by DNA sequencing and characterized new variants and haplotypes by a novel haplotype-specific sequencing-based approach. The haplotypes 2/3 (padi4_89*G, padi4_90*T, padi4_92*G, padi4_94*T, padi4_104*T, padi4_95*C, padi4_96*C), and haplotype 4 (padi4_89*G, padi4_90*T, padi4_92*G, padi4_94*T, padi4_104*C, padi4_95*G, padi4_96*T) conferring susceptibility to rheumatoid arthritis were detected at frequencies of 30.9% and 7.8%, respectively. In addition, three novel coding SNPs in exons 2, 3, and 4, and three SNPs in introns 2 and 3 located near the exon-intron boundaries were identified in 11 individuals (10.8%). The so-called nonsusceptibility haplotype 1 (padi4_89*A, padi4_90*C, padi4_92*C, padi4_94*C, padi4_104*C, padi4_95*G, padi4_96*T) occurred at a frequency of 58.3%. Additionally, we identified a closely related novel haplotype, haplotype 1B (2.9%), that differs from haplotype 1 only by padi4_92*G/padi4_96*C. This haplotype was not described in the Japanese population. Our results indicate that the PADI4 gene exhibits a remarkable variability and a rather complex haplotypic organization. Further studies on disease association of PADI4 should be performed by haplotype-specific sequencing-based approaches to identify the exact genotype of the PADI4 fragment of interest.

Keywords PADI4 · Haplotype · Sequencing · Single nucleotide polymorphism · Rheumatoid arthritis

Abbreviations *PAD*: Peptidylarginine deiminase · *nPADI*: Nonsusceptibility haplotype · *sPADI*: Susceptibility haplotype · *PCR*: Polymerase chain reaction · *RA*: Rheumatoid arthritis · *SNP*: Single nucleotide polymorphism

Introduction

The peptidylarginine deiminases (PADs, EC 3.5.3.15) are enzymes involved in the posttranslational deimination of protein-bound arginine to citrulline [1]. Five different types of PADs encoded by the genes PADI1–4 and PADI6 are currently known [1]. The exact functional significance of these enzymes is unknown. However, evidence suggests that at least PADI4 might have an immunomodulatory function, and that it leads to breakage of tolerance under certain circumstances. Posttranslational deimination of proteins is a phenomenon that occurs under physiological and pathological conditions. Citrulline is found in structural proteins such as filaggrin and in some keratins in terminally differentiating keratinocytes [2]. Recent reports describe the stimulation-dependent citrullination of histones in granulocytes and suggest a possible role of this modification in chromatin remodeling [3]. Moreover, citrulline-modified proteins are thought to be targets of the autoimmune reaction in some autoimmune diseases. For example, enhanced T-cell responsiveness to citrullinated myelin basic protein has been observed in multiple sclerosis [4].

The presence of citrulline-modified target epitopes for autoantibodies is a well known phenomenon in rheumatoid arthritis (RA) [5, 6]. PADs were recently implicated in the generation of anti-cyclic citrullinated peptide antibodies detectable in early stages of RA [5, 6, 7]. The process resulting in anti-cyclic citrullinated peptide antibody formation is thought to play a pivotal role in early stages of disease progression since it is detectable several years before the onset of symptoms in patients with RA [8]. There is evidence that the deimination of arginine at those peptide side-chain positions that interact with the so-called shared epitope of some major histocompatibility complex class II molecules (e.g., HLA-DRB1*0401 or HLA-DRB1*0404) results in the generation of high-affinity peptides, thus inducing a strong *in vitro* T-cell activation [7]. Using gene-based linkage disequilibrium mapping approaches, a Japanese research group identified in 1p36 a genomic region containing the genes PADI1–4, which seemed to be associated with susceptibility to RA. The gene responsible for the association with RA was identified as PADI4, which has four main haplotypes that differ at four exonic single nucleotide polymorphisms (SNPs), with three subsequent amino acid substitutions [9]. While the so-called susceptibility haplotypes (*sPADI*) 2, 3, and 4 were found to be significantly more frequent in Japanese individuals suffering from RA, the nonsusceptibility haplotype (*nPADI*) 1 predominated in healthy individuals [10]. However, another group studying the association between PADI4 and RA in the United Kingdom

did not find a difference in PADI4 haplotype distribution between RA patients and healthy individuals [11]. Thus the relevance of PADI4 variability for susceptibility to RA is still unclear.

PADI4 variability has been tested until now by SNP screening using techniques such as TaqMan 5' allelic discrimination or Invader assays, and the corresponding haplotypes were calculated by the expectation-maximization algorithm [10, 11]. In addition, the identification of SNPs by sequencing-based approaches was limited to screening for heterozygote positions [10]. In other words, techniques that allow an in-depth analysis of PADI4 to determine the exact *cis/trans* linkage of different SNPs and to identify additional novel variants in their exact haplotypic context are still lacking. Consequently we devised a method for sequencing-based characterization of exons 2–4 of the PADI4 gene in a healthy white German population using a novel long-range (5.3 kb) haplotype-specific amplification technique.

Material and methods

Genomic DNA was extracted from whole blood using GenoPrep cartridges B and the GenoM-6 system (GenoVision) following the manufacturer's instructions. Blood samples were withdrawn from healthy, unrelated blood donors who gave their informed consent. The mean age of the 102 individuals studied was 40.6 years (range 19–64 years); 57% of the subjects were women. Cycle sequencing of DNA samples was carried out in two ways (numbering of nucleotides was based on the respective position in sequence NT_034376.1).

First, we sequenced PADI4 exons 2, 3, and 4 separately, which span regions 389,947–390,216 (exon 2), 392,874–393,094 (exon 3), and 395,101–395,353 (exon 4), respectively. Briefly, we amplified DNA on a thermal cycler (GeneAmp PCR system 9700, Perkin Elmer) using primers binding to conserved regions of the respective introns adjacent to the exon of interest (Table 1). The following final concentrations of primers were used: PADI4ex02_01+/PADI4ex02_01– (250 nM), PADI4ex03_01+/PADI4ex03_01– (250 nM), and PADI4ex04_01+/PADI4ex04_05– (1.25 μ M). The polymerase chain reactions (PCRs) were performed in a total volume of 15 μ l containing: 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 2 mM MgCl₂, 0.2 mM deoxyribonucleoside triphosphate (Invitrogen), 50 ng genomic DNA, and 0.6 U Platinum *Taq* DNA polymerase (Invitrogen). Thermal cycling conditions were: denaturation (96°C, 2 min), 10 cycles (96°C, 15 s; 65°C, 1 min), and 20 cycles of (96°C, 10 s; 61°C, 2 min; 72°C, 30 s). PCR products were sequenced on a thermal cycler (25 cycles of 96°C, 10 s; 50°C, 15 s, 60°C, 4 min) using BigDye terminators v. 1.1 (Applied Biosystems) and each of the primers (final concentration: 2.5 nM) used for amplification, separately. Electrophoresis and analysis were performed using an ABI310 capillary sequencer (Applied Biosystems) and Sequencing Analysis Software (Applied Biosystems) or 4Peaks (Mek&Tosj, The Netherlands Cancer Institute).

The second sequencing approach was designed to provide information on the exact haplotypic organization of novel variants and haplotypes of exons 2–4 of PADI4. It was therefore necessary to amplify large DNA fragments (5.3 kb) in a haplotype-specific manner. We designed allele-specific primers for the SNPs *padi4_89*A/G* and *padi4_96*T/C* (Table 1) and performed long-range PCR. Briefly, we performed four amplification reactions using Platinum PCR SuperMix High Fidelity (Invitrogen) and one of the following haplotype-specific primer pairs (Table 1; final concentrations are indicated in parentheses): *padi4_89_F01A* (forward primer, 300 nM)/*padi4_96_R01T* (reverse primer, 300 nM), *padi4_89_F01A* (forward primer, 200 nM)/*padi4_96_R01C* (re-

Table 1 Primers used for sequencing of PADI4. Primers used for amplification and sequencing of exons 2, 3, and 4 of PADI4. The specified locations of the 3'-terminals of the primers are based on the numbering of sequence NT_034376.1 (*F* forward primer, *R* reverse primer)

Designation	Product	Sequence 5'→3' (location of 3' terminal end)
PADI4ex02_1+ (F)	Exon 2	GCCCCCTCCTCACTGCATC (389,946)
PADI4ex02_1- (R)		GATCACGGCAGGACAGAG (390,217)
PADI4ex03_1+ (F)	Exon 3	CCAGTGCCCTACCCCTGAG (392,873)
PADI4ex03_1- (R)		GTAAGCTCCAGGTGTAGTTTC (393,095)
PADI4ex04_1+ (F)	Exon 4	GAAACGACCTGCCCATTC (395,100)
PADI4ex04_5- (R)		ATGGTGATGCTAATTTTCTTTTC (395,354)
padi4_89_F01A (F)	Haplotype 1	TCGTGGATATTGCCACAG (390,046)
padi4_96_R01T (R)		CGGAATACGTGGGACAG (395,518)
padi4_89_F01A (F)	Haplotype 1B	TCGTGGATATTGCCACAG (390,046)
padi4_96_R01C (R)		CGGAATACGTGGGACAGG (395,518)
padi4_89_F01G (F)	Haplotype 4	TCGTGGATATTGCCACAG (390,046)
padi4_96_R01T (R)		CGGAATACGTGGGACAG (395,518)
padi4_89_F01G (F)	Haplotype 2/3	TCGTGGATATTGCCACAG (390,046)
padi4_96_R01C (R)		CGGAATACGTGGGACAGG (395,518)

verse primer, 200 nM), padi4_89_F01G (forward primer, 200 nM)/padi4_96_R01T (reverse primer, 200 nM), and padi4_89_F01G (forward primer, 200 nM)/padi4_96_R01C (reverse primer, 200 nM). The thermal cycle profile for long-range PCR was as follows: denaturation (94°C, 2 min), 15 cycles of (94°C, 30 s; 65°C, 30 s; 68°C, 5.5 min), 15 cycles of (94°C, 30 s; 60°C, 30 s; 68°C, 5.5 min), and 10 cycles of (94°C, 30 s; 55°C, 30 s; 68°C, 5.5 min). The specificity of these primer pairs for the distinct PADI4 haplotypes was tested on the haplotypes calculated using the expectation-maximization algorithm (EH program, available at ftp://linkage.rockefeller.edu/software/eh) based on the results of the sequencing of single exons described above. The reactions resulting in an amplification product were digested using ExoSAP-IT (Amersham Biosciences) following the manufacturer's instructions, and the PCR products were sequenced on a thermal cycler (25 cycles of 96°C, 10 s; 50°C, 10 s; 60°C, 4 min) using BigDye terminators v. 1.1 and one of the following sequencing primers with a final concentration of 2.5 nM (Table 1): PADI4ex02_1- (exon 2, reverse primer), PADI4ex03_1+ (exon 3, forward primer), PADI4ex03_1- (exon 3, reverse primer), or PADI4ex04_1+ (exon 4, forward primer). The designations of the PADI4 haplotypes are in accordance with Suzuki et al. [10].

Results

PADI4 haplotype frequencies (exons 2–4)

Different haplotypes formed by SNPs padi4_89, padi4_90, padi4_92, padi4_94, padi4_104, padi4_95, and padi4_96 were identified by sequencing-based analysis of PADI4 exons 2–4 (Table 2). There were no discrepancies between haplotypes calculated by the expectation-maximization algorithm and those characterized by the haplotype-specific sequencing-based approach. Nonsusceptibility haplotype 1 (58.3%) and susceptibility haplotype 2/3 (30.9%) were the most prevalent haplotypes. Haplotype 4 was found at a frequency of 7.8%. We additionally identified a novel PADI4 haplotype, which is most closely related to haplotype 1. This haplotype designated as 1B (2.9%, accession number: AJ715933) differs from haplotype 1 by padi4_92*G/padi4_96*C. When analyzing the exons 2–4 of PADI4 separately, one individual seemed to exhibit haplotype 1B (padi4_89*A, padi4_90*C, padi4_92*G, padi4_94*C, padi4_104*C, padi4_95*G, padi4_96*C) and an additional novel haplotype. However, when analyzing this DNA sample by the haplotype-specific sequencing-based approach de-

Table 2 Haplotype frequencies of PADI4 (exons 2–4) in white population ($n=102$). Haplotypic organization of exons 2–4 of PADI4 and the haplotype frequencies (*parenthesis*); PADI4 haplotype designations are based on those of Suzuki et al. [10]

	Haplotype			
	1 (58.3%)	1B (2.9%)	2/3 (30.9%)	4 (7.8%)
padi4_89	A	A	G	G
padi4_90	C	C	T	T
padi4_92	C	G	G	G
padi4_94	C	C	T	T
padi4_104	C	C	T	C
padi4_95	G	G	C	G
padi4_96	T	C	C	T

Table 3 PADI4 (exons 2–4) haplotype combinations in a white population ($n=102$): numbers of different PADI4 haplotype combinations

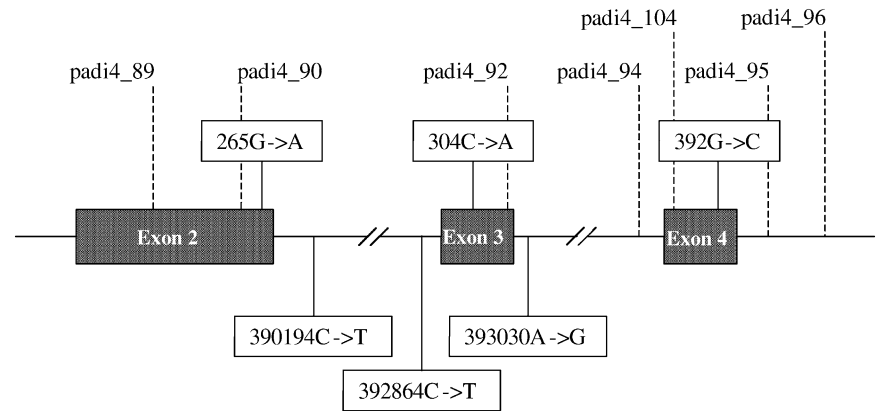
Haplotype I	Haplotype II			
	1	1B	2/3	4
1	35 (34.3%)	5 (4.9%)	35 (34.3%)	9 (8.8%)
1B	0	0	1 (1%)	0
2/3	0	0	12 (11.8%)	3 (2.9%)
4	0	0	0	2 (2%)

scribed in this report, we identified haplotypes 1 and 1B. On haplotype 1 a novel PADI4 variant in intron 2 (AJ715932) was found that was located at the binding site of primer PADI4ex03_1+. Due to this novel variant in this case the amplification of exon 3 of the respective PADI4 haplotype failed using the first sequencing-based technique, resulting in the artificial identification of an additional PADI4 haplotype.

Distribution of PADI4 haplotype combinations

The frequencies of the PADI4 haplotype combinations found in our white population are shown in Table 3. PADI4 haplotype 1 was most prevalent in homozygous form or in combination with the haplotype 2/3 (both 34.3%). The frequencies of all haplotypic constellations are in agreement with a Mendelian distribution.

Fig. 1 Schematic presentation of novel PADI4 variants. Localization of the 3 exonic and 3 intronic variants characterized in this study. Exons 2, 3, and 4 and parts of the neighboring introns are shown



Localization and characterization of six novel PADI4 variants

In the present study we identified six novel PADI4 variants in 11 individuals (10.8%). Three of these SNPs were located in exons 2, 3, and 4 of PADI4 and were found in six individuals (5.9%). The specified positions of these exonic SNPs are indicated based on cDNA sequence NM_012387 (Fig. 1). The substitutions 265G→A ($n=2$, *PADI4h01ex02/01*, accession number AJ715934), 304C→A ($n=2$, *PADI4h02ex03/01*, accession number AJ715937), and 392G→C ($n=2$, *PADI4h02ex04/01*, accession number AJ715935) result in amino acid substitutions D89 N, P102T, and R131T, respectively. The novel PADI4 variants were integrated in the haplotypic context of PADI4 by haplotype-specific sequencing. While PADI4 265A was linked with haplotype 1, the SNPs PADI4 304A and PADI4 392C were linked to haplotype 2/3.

The specified positions of the intronic SNPs identified in 5 individuals (4.9%) are indicated based on the sequence NT_034376.1 (Fig. 1). SNP 390194C→T (*PADI4h02in02/01*, accession number AJ715938) linked to haplotype 2/3 was found in three individuals and is located 38 nucleotides downstream of the boundary of exon 2 and intron 2. SNP 393030A→G (*PADI4h02in03/01*, accession number AJ715936) which was identified in intron 3 of PADI4 haplotype 2/3 ($n=1$) is located 14 nucleotides downstream of exon 3. Linked to haplotype 1 the SNP 392864C→T (*PADI4h01in02/01*, accession number AJ715932) was found 85 nucleotides upstream of exon 3 ($n=1$).

The unambiguous determination of the cis/trans linkage of SNPs 390194C→T and 392G→C by the expectation-maximization algorithm was not possible because both SNPs were identified in individuals presenting uniformly with PADI4 haplotype 1 combined with haplotype 2/3. In these cases haplotype-specific sequencing was necessary to assign the exact haplotypic context.

Discussion

The mechanism by which PADI4 variability affects the breakage of tolerance is still unknown. Initial studies demonstrated different half-lives of mRNA transcribed from sPADI4 and nPADI4 [9, 10]. It was argued that these differences in mRNA stability can result in higher enzymatic activity in cases in which sPADI4 is present, leading to the generation of larger amounts of citrullinated peptides. This could ultimately promote an autoimmunization process. However, we believe that differences in substrate specificities between sPADI4- and nPADI4-encoded enzymes that can result in the formation of specific sPADI4-dependent, citrullinated autoantigens triggering autoimmunization should be considered as well. Similar to the specific binding and presentation of distinct peptide repertoires by different MHC molecules the gene product of sPADI4 could bind and modify peptide motifs that are not compatible for the interaction with nPADI4-encoded proteins. To verify this hypothesis the PADI4 gene should be characterized using techniques capable of identifying the cis/trans linkage of SNPs directly, thus allowing one to determine the exact haplotypic organization of PADI4, including the detection and characterization of novel polymorphisms and their haplotypic linkage.

We devised a corresponding approach allowing an exact analysis of PADI4 (exons 2–4) haplotypes using a haplotype-specific amplification protocol covering the whole region of interest. This approach utilizes a PCR assay designed to amplify large DNA fragments containing *Taq* DNA polymerase and the proofreading enzyme *Pyrococcus* species *GB-D* polymerase. In principle, the development of a haplotype- or allele-specific amplification procedure using allele-specific forward and/or reverse primers could be hampered by the 3'→5' exonuclease activity of the proofreading enzyme *GB-D* polymerase. However, the technique described here permits unambiguous haplotype-specific amplification and sequence analysis (Fig. 2). The applicability of the described approach for an allele-specific amplification might be explained by at least two reasons. First, there is an enormous preponderance of allele-specific primers over those pri-

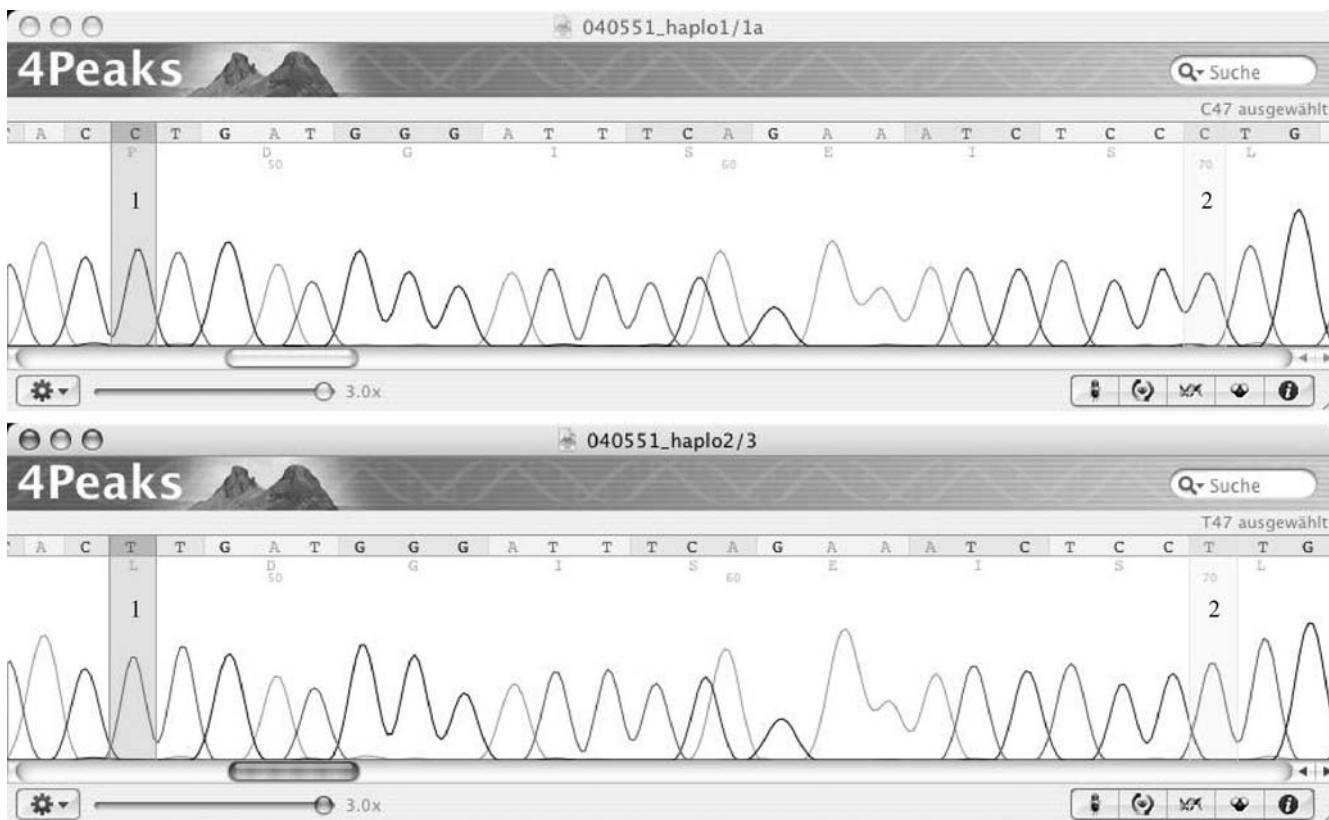


Fig. 2 Haplotype-specific sequencing of PADI4. Sequences of DNA sample 040551 are shown. The haplotypes 1 and 2/3 were amplified and sequenced separately. The presented sequences cover

the boundary between intron 3 and exon 4 of PADI4. 1: SNP padi_94; 2: SNP padi_104

mers whose allele-specific 3'-terminal ends were cut by 3'→5' exonuclease activity during the amplification procedure. Second, due to the allele-specificity of both the forward and reverse primers used for haplotype characterization, the portion of such DNA strands synthesized by the concerted action of forward and reverse primers that are both cut at their 3'-terminal ends is negligible.

The main PADI4 haplotypes in our white German population exhibited a distribution similar to those in Japanese and British studies [10, 11]: The most prevalent forms were haplotype 1 (padi_89*A, padi_90*C, padi_92*C, padi_94*C, padi_104*C, padi_95*G, padi_96*T) and haplotype 2/3 (padi_89*G, padi_90*T, padi_92*G, padi_94*T, padi_104*T, padi_95*C, padi_96*C; Germany 58%/31%; Japan 60%/29%; United Kingdom 56%/32%). We did not discriminate between PADI4 haplotypes 2 and 3 because SNP padi_102, which differentiates between haplotypes 2 and 3, is located more than 11 kb downstream of the region of interest. Haplotype 4 (padi_89*G, padi_90*T, padi_92*G, padi_94*T, padi_104*C, padi_95*G, padi_96*T) was about twice as frequent in Germany and the United Kingdom as in Japan (Germany 8%; Japan 4%; United Kingdom 9%).

An exact comparison of the frequency of the haplotype 1B identified in this study with this of previously published studies was not possible. No SNP constellation

comparable to haplotype 1B was described in the Japanese population [10]. In the British study only the SNPs padi_89, padi_90, padi_92, and padi_104 were determined [11]. However, when considering the constellation padi_89*A, padi_90*C, padi_92*G, and padi_104*C, which is common to haplotypes 1B, the frequencies reported in the UK (2.2%) and in the present study (2.9%) are largely similar.

The most remarkable finding in our study was the large number of additional novel variants identified (Fig. 1). More than 10% of the individuals studied presented with previously unknown polymorphisms. All of the exonic variations result in amino acid substitutions (265G→A, D89 N; 304C→A, P102T; 392G→C, R131T) that alter the charge of the respective amino acids (D89 N, R131T), or that may affect the steric arrangement of the neighboring amino acids (P102T). Because the novel intronic variations are located near the exon-intron boundaries—390194C→T, 392864C→T, and 393030A→G are located 38 bp downstream of exon 2, 85 bp upstream of exon 3, and 14 bp downstream of exon 3—one may speculate that both variations affect the process of splicing. However, intronic variants located more distantly from intron-exon boundaries may also affect the results of disease association studies. Further studies should address the questions of the functional relevance of the described amino acid substitutions and of the influence of intronic variants on

PADI4 splicing. Studies focusing on the structural analysis of PADI4 by X-ray crystallographic analysis are under way [12]. A complete structural analysis of PADI4 will help to understand the way by which PADI4 interacts with the respective substrates and how variations of PADI4 could modify substrate specificity.

A further interesting observation is that four out of six novel variants present in 8 out of 11 individuals were found to be in cis linkage with the susceptibility haplotype 2/3. This finding is all the more interesting when one considers that haplotype 2/3 is about half as frequent as haplotype 1. The linkage of the newly described PADI4 variants with the susceptibility haplotype 2/3 raises the question of whether the phenomenon of association with RA is affected, not only by the SNP constellation characterizing the so-called susceptibility PADI4 haplotypes 2/3 and 4 themselves but also by additional variants predominantly found in linkage with these susceptibility haplotypes. Such additional variants cannot be identified by simple SNP diagnostic procedures such as amplification refractory mutation system, TaqMan 5' allelic discrimination assays or Invader assays. We therefore emphasize that further studies on disease association of PADI4 should be performed using sequencing-based approaches that allow the identification of novel variants and the characterization of their exact haplotypic context.

In view of the variability of PADI4 and the need for a correct attribution of novel variants to the respective PADI4 haplotypes we feel it is necessary to establish a PADI4 nomenclature allowing a clearcut description of PADI4 variants.

Acknowledgements We thank Gisela Diederich for excellent technical assistance.

References

1. Chavanas S, Mechin MC, Takahara H, Kawada A, Nachat R, Serre G, Simon M (2004) Comparative analysis of the mouse and human peptidylarginine deiminase gene clusters reveals highly conserved non-coding segments and a new human gene, PADI6. *Gene* 330:19–27
2. Senshu T, Kan S, Ogawa H, Manabe M, Asaga H (1996) Preferential deimination of keratin K1 and filaggrin during the

- terminal differentiation of human epidermis. *Biochem Biophys Res Commun* 225:712–719
3. Nakashima K, Hagiwara T, Yamada M (2002) Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. *J Biol Chem* 277:49562–49568
4. Tranquill LR, Cao L, Ling NC, Kalbacher H, Martin RM, Whitaker JN (2000) Enhanced T cell responsiveness to citrulline-containing myelin basic protein in multiple sclerosis patients. *Mult Scler* 6:220–225
5. Girbal-Neuhauser E, Durieux JJ, Arnaud M, Dalbon P, Sebbag M, Vincent C, Simon M, Senshu T, Masson-Bessiere C, Jolivet-Reynaud C, Jolivet M, Serre G (1999) The epitopes targeted by the rheumatoid arthritis-associated antifilaggrin autoantibodies are posttranslationally generated on various sites of (pro)filaggrin by deimination of arginine residues. *J Immunol* 162:585–594
6. Masson-Bessiere C, Sebbag M, Girbal-Neuhauser E, Nogueira L, Vincent C, Senshu T, Serre G (2001) The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin. *J Immunol* 166:4177–4184
7. Hill JA, Southwood S, Sette A, Jevnikar AM, Bell DA, Cairns E (2003) Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. *J Immunol* 171:538–541
8. Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH, Habibuw MR, Vandenbroucke JP, Dijkmans BA (2004) Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum* 50:380–386
9. Yamada R, Suzuki A, Chang X, Yamamoto K (2003) Peptidylarginine deiminase type 4: identification of a rheumatoid arthritis-susceptible gene. *Trends Mol Med* 9:503–508
10. Suzuki A, Yamada R, Chang X, Tokuhiko S, Sawada T, Suzuki M, Nagasaki M, Nakayama-Hamada M, Kawaida R, Ono M, Ohtsuki M, Furukawa H, Yoshino S, Yukioka M, Tohma S, Matsubara T, Wakitani S, Teshima R, Nishioka Y, Sekine A, Iida A, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet* 34:395–402
11. Barton A, Bowes J, Eyre S, Spreckley K, Hinks A, John S, Worthington J (2004) A functional haplotype of the PADI4 gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population. *Arthritis Rheum* 50:1117–1121
12. Arita K, Hashimoto H, Shimizu T, Yamada M, Sato M (2003) Crystallization and preliminary X-ray crystallographic analysis of human peptidylarginine deiminase V. *Acta Crystallogr D Biol Crystallogr* 59:2332–2333