Saudi Journal of Biological Sciences 29 (2022) 1227–1233

Contents lists available at ScienceDirect

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Novel functional polymorphism on *PADI-4* gene and its association with arthritis onset

Maryam Mukhtar Dr Ph. D. [a,b,*], Nadeem Sheikh Dr. Ph. D. [a,*], Andleeb Batool Dr. Ph. D. [c], Muhammad Babar Khawar Dr. Ph. D. [a,d], Naz Fatima Dr. Ph. D. [a], Rabia Mehmood Dr. Ph. D. [a]

[a] *Cell and Molecular Biology Lab, Department of Zoology, University of the Punjab, Lahore, Pakistan*
[b] *Rhumatology Domain, Center of Molecular Medicine, Karolinska Institute, Department of Medicine, Stockholm, Sweden*
[c] *Department of Zoology, Government College University, Lahore, Pakistan*
[d] *Department of Zoology, Faculty of Sciences, University of Central Punjab (UCP), Lahore, Pakistan*

ARTICLE INFO

Article history:
Received 1 August 2021
Revised 10 September 2021
Accepted 13 September 2021
Available online 17 September 2021

Keyword:
Rheumatoid arthritis
Osteoarthritis
PADI-4 gene polymorphism
Genotyping

ABSTRACT

Background: Citrullinated proteins formed by peptidyl arginine deiminases (PADIs) deimination of arginine residues in proteins are of particular interest in arthritis pathogenesis. Polymorphisms on the *PADI-4* gene lead to the malfunctioning of PADIs leading to the onset of arthritis.
Objective: The present study was conducted to determine the polymorphisms on the *PADI-4* gene and their association with rheumatoid arthritis (RA) as well as Osteoarthritis (OA).
Methodology: To achieve the above-mentioned objective a case-control study was conducted. Blood samples were collected from RA, OA, and control subjects. DNA was extracted from each blood sample by modified organic method and was quantified as well as qualified by DNA gel electrophoresis and Nanodrop. Patients were tested for rs874881, rs11203366, rs11203367, rs2240336, rs2240337, rs2240339, rs1748033 and rs2240340 polymorphic sites by amplifying targeted regions through PCR with site-specific primers. Genotyping was performed by Restriction Fragment Length Polymorphism and direct sequencing method. Mutations were identified by analyzing sequences on BioEdit software. Allelic, genetic, and multiple site analysis were performed by SHEsis and PLINK software. Change in the amino acid sequence was identified by MEGA 6.0 software.
Results: Polymorphisms were identified on all targeted polymorphic sites except rs2240337 in both RA and OA individuals. In addition, two novel mutations were also identified in exon 4 identified i-e SCV000804840: c.218T > C and SCV000807675: c.241G > T. All the SNPs except rs11203366 were found to be significantly associated with RA at an allelic level whereas all SNP's have been significant risk factors in the onset of OA. At genotypic level rs874881, rs11203366, rs2240339, SCV000804840 and SCV000807675 were significantly associated to RA development whereas rs874881, rs11203366, rs11203367, rs2240339, SCV000804840 and SCV000807675 were genetic risk factors in OA onset. Haplotype analysis indicated that GACCACGCC and GACCACGCT were highly significant in disease development. Polymorphisms identified altered the functioning of PADIs by altering their amino acid sequence.
Conclusion: In conclusion, it was found that *PADI-4* gene polymorphism was not only involved in the onset of RA but was also found to be a significant risk factor in OA onset.

1. Introduction

Arthritis is a disease characterized by inflammation of joints and was a leading cause of the limited working efficiency of individuals. Rheumatoid arthritis (RA) and osteoarthritis (OA) are the two major types of arthritis. An autoimmune and systematic inflammatory RA disease was characterized by articular manifestation as well as extra-articular manifestation as fatigue, anemia, weight loss, and malaise (Mirza, 2016) whereas, OA is a nonsystematic disease characterized by non-inflammatory an inherent

* Corresponding authors at: Cell and Molecular Biology Lab, Department of Zoology, University of the Punjab, Lahore, Pakistan.
E-mail addresses: m.mukhtar29@gmail.com (M. Mukhtar Dr), s_nadeem77@yahoo.com (N. Sheikh Dr.).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.sjbs.2021.09.037
1319-562X/

disease of movable joints. OA involves articular joint deterioration and new bones formation at joint margins and surfaces (Hough, 2001). The etiology of the disease is still not fully understood but it is evident that both RA and OA are multifactorial diseases characterized by the interaction of environmental and genetic factors.

The linkage between arthritis and HLA is very well established and it accounts for about 1/3 of the genetic factor of RA (MacGregor et al., 2000). So far many non-HLA genes were also investigated to determine new RA as well as OA susceptible loci. Peptidylarginine deiminases 4 (*PADI-4*) is one member of the *PADI* gene family located on the 1p36 chromosome and encodes for catalytic enzymes involved in the posttranslational modification of arginine residues into citrulline. If subsequent conformation changes caused proteins to be citrullinated there is the possibility that such proteins may cause breakdowns in immunologic tolerance, due to variation of their antigenicity (Zhou and Ménard, 2002). There was evidence that *PADI-4* mRNA stability increased with its association with RA susceptibility *PADI-4* haplotype (Suzuki et al., 2003). The increasing stability leads to the *PADI-4* protein accumulation with citrullinated proteins subsequent increase which ultimately enhanced autoantibodies production against citrullinated peptides (Cha et al., 2007).

Multiple susceptible genes to RA have been identified as a result of genome-wide association studies (GWAS) (Goulielmos et al., 2016; Okada et al., 2014). Positive association of *PADI-4* gene and RA was reported in Korean and French populations (Kang et al., 2006; MacGregor et al., 2000). However, no association was observed in the Spanish population, UK resident Caucasian subjects, and French Caucasian families (Caponi et al., 2005; Barton et al., 2004; Martinez et al., 2005). Taken together, these data indicate that *PADI-4* may be considered as one of the strong candidate genes for RA susceptibility.

PADI-4 Functional haplotypes that affect transcripts stability, as well as anti-CCP antibody level, have been linked with RA in the Asian population (Suzuki et al., 2003). The *PADI-4_94*, *_104*, *_92*, *_90*, and *_89* polymorphisms have been widely examined in multiple research studies on *PADI-4* polymorphisms in arthritis (Zavala-Cerna et al., 2013; Li et al., 2013; Panati et al., 2012; Cheng et al., 2012; Hassine et al., 2012; El-Gabalawy et al., 2011; Chen et al., 2011; Burr et al., 2010) but still the results from those studies are controversial. The difference might be because of the change in ethnic groups or genetic makeup. Majorly the *PADI-4* gene polymorphism was investigated with RA onset in different ethnic groups other than Pakistani individuals. To my knowledge, no study was conducted to find its association in OA subjects. Therefore the present study was conducted to investigate the linkage of *PADI-4* gene polymorphism in association to RA as well as OA onset in Pakistani individuals.

2. Materials and methods

2.1. Sampling

A Case-control study was conducted after getting ethically approved by the Bioethical Committee of the University of the Punjab, Lahore. The study was conducted in the rheumatology and orthopedic center of public and semi-Government Hospitals of Punjab, Pakistan. The government-registered ortho, rheumatology centers, and outpatient clinics were selected to collect clinical data of patients and self-reported cases of RA and OA were excluded. The detailed inclusion criteria for RA subjects include patients suitable for Rituximab therapy and with positive RF factor; 2010 ACR / EULAR criteria for classification/diagnosis of disease; patients of age either 18years or above and patient able of signing informed consent Performa; keenness and capacity to follow with scheduled visits, plans of treatment and laboratory tests, and other study procedures. The inclusion criteria for OA subjects include patients which were overweight/Obese (Overweight: BMI $\geq$ 25; Obese: BMI $\geq$ 30); patients with a previous knee injury or surgery and with history of knee pain; patients whom parents or sibling who had a knee replacement. All patients clinically diagnosed by the physician were included in the study. Patients that were self-diagnosed and have incomplete laboratory tests reports were excluded from the study. Gout patients were excluded. Patients with an additional viral disease like hepatitis or tuberculosis were also excluded from the study. All control subjects were age and gender-matched to the cases. All control subjects were healthy and with a negative family history of arthritis. The summarized methodology has been presented in Fig. 1.

2.2. Collection of blood samples

Blood samples were collected after taking written consent from 300 RA, 316 OA, and 412 controls individuals in EDTA coated tubes by using BD syringes. The samples were stored at 4 °C before DNA extraction.

2.3. DNA extraction

DNA was manually isolated by the modified organic method. A whole blood sample of about 500 μl was taken in Eppendorf and mixed with 500 μl TE buffer. The mixture was centrifuged at 12,500 rpm in a refrigerated centrifuge. The supernatant was discarded and the pellet was washed with 1000 μl TE buffer 3–4 times until the pellet get white-colored. Pellet was incubated with 0.5 μl of proteinase K, 375 μl of 3 M sodium acetate, and 15 μl of sodium dodecyl sulfate for 24 h at 37©C. The mixture of isoamyl alcohol and chloroform was added and DNA layer was separated from debris. The DNA was precipitated with chilled absolute alcohol. The DNA pellet was washed with 70% alcohol and then DNA pellet was dissolved in DEPC water and was stored at −80 °C.

2.4. Genotyping

Polymorphisms were determined for rs874881, rs11203366, rs11203367, rs2240336, rs2240337, rs2240339, rs1748033 and rs2240340 polymorphic sites by PCR-RFLP and direct sequencing method. The primers concerning the targeted polymorphic sites along with their respective endonucleases enzymes (Thermo Scientific) and conditions were summarized in Table 1.

2.5. Statistical analysis

Polymorphisms were identified by analyzing the sequences on BioEdit software. The data was tabulated and genetically analyzed. The whole data was passed through Hardy-Weinberg Equilibrium (HWE). For allelic and genetics frequencies Fisher's P test was applied. Linkage disequilibrium (L.D) and haplotypes were calculated by SHEsis and PLINK software. Alterations in amino acid sequences were determined by Mega 6 software.

3. Result

The selected gene and SNP's were analyzed for their susceptibility to OA and RA. A case-control study was conducted comprised of 300 cases of RA, 316 of OA, and 412 controls. Blood samples were collected along with the clinical characters. The clinical characteristics of the diseased groups are tabulated in Table 2.

As a result of sequencing, it was observed that on rs874881 the polymorphic site C allele was replaced by the G allele. On

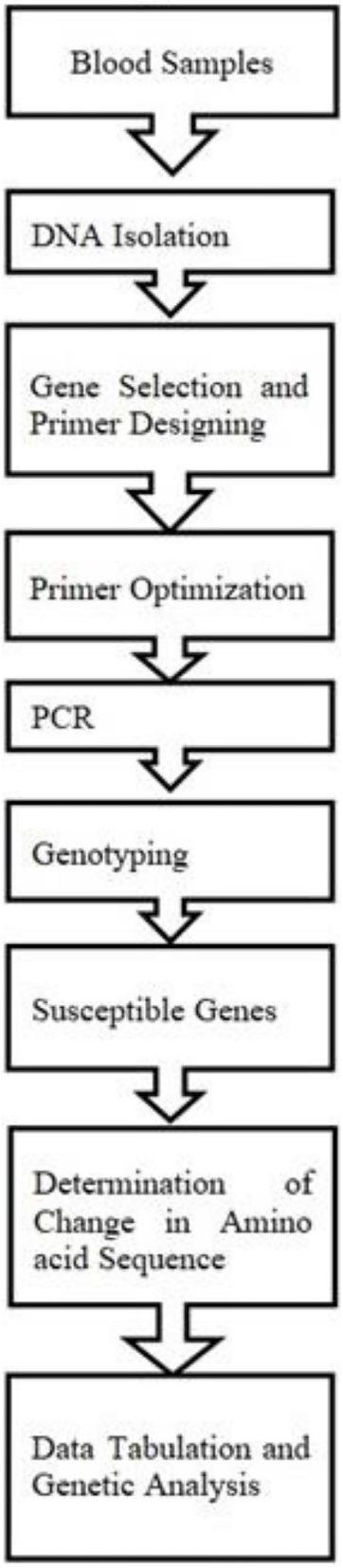

Fig. 1. Flow chart Summarizing Methodology.

rs011203366 and rs11203367 polymorphic sites on exon 2 of *the PADI-4* gene wild alleles A and T were replaced by mutant alleles G and C in both OA and RA individuals respectively. At rs2240336 and rs2240339 alleles, T and C were more prevalent among arthritis patients as compared to alleles C and A respectively. However, no change was observed on the rs2240337 site. On rs1748033 allele T replaced allele C whereas on rs2240340 T allele was more prevalent in arthritis patients as compared to allele C. Two novel SNPs on exon 4 of *PADI-4* gene were also identified i-e SCV00807675: g.17157671 G > T and SCV000804840: c.341–15T > C.

It was observed that all the targeted polymorphic sites on *PADI-4* were followed HWE as p = 1.00. The allelic and genotypic analysis was performed and it was observed that allelic frequency of rs874881, rs11203367, rs2240336, rs2240339, rs1748033, SCV00807675, rs2240340 and SCV000804840 were significantly varied among RA and controls with O.R = 0.010, 95% CI = [0.006–0.017]; O.R = 15.328, 95% CI = [11.258–20.868]; O.R = 0.117, 95% CI = [0.087–0.156]; O.R = 0.062, 95% CI = [0.042–0.091]; O.R = 0.114, 95% CI = [0.082–0.159]; O.R = 0.035, 95% CI = [0.024–0.052]; O.R = 16.073, 95% CI = [11.424–22.615]; O.R = 49.00; 95% CI = [17.528–136.981] as well as in OA and controls with O.R = 0.013 95% CI = [0.008–0.020]; O.R = 9.814, 95% CI = [7.250–13.284]; O.R = 0.987, 95% CI = [0.074–0.132]; O.R = 0.045, 95% CI = [0.029–0.069]; O.R = 0.090, 95% CI = [0.065–0.126]; O.R = 0.035, 95% CI = [0.024–0.052]; O.R = 10.210, 95% CI = [7.276–14.334]; O.R = 49.00; 95% CI = [17.528–136.981] respectively. However, the allelic frequency of rs11203366 was significantly varied in OA individuals but not in RA subjects (Table 3). At genotypic level, rs874881, rs11203366, rs2240339, SCV00807675 and SCV000804840 were found to be significant risk factor in OA and RA onset (Table 4).

The L.D in detail for both RA and OA were presented in Figs. 2 and 3 (a, b) respectively. It was observed that in both cases all SNP's together were significant risk factors in the onset of disease to the next generation as for RA D′ = 0.100; r2 = 0.311 and for OA D′ = 0.100; r2 = 0.311 respectively. SCV000807675 and SCV000804840 were also involved in disease development.

Both in RA and OA as a result of haplotype analysis it was observed that all haplotypes were significant ($p < 0.01$) for disease onset as their frequency was higher in patients as compared to control subjects (Table 5).

As a whole *PADI-4* gene was found to be a significant risk factor in the onset of both types of disease with a global chi-square value of 201.704 and fisher's $p < 0.001$.

Protein alignment was performed for each functional polymorphic site and was summarized in Table 6.

4. Discussion

To identify polymorphisms influencing the risk of arthritis development, the genetic association is a vital statistical tool. But it is difficult to interpret from such studies which genes show exact true association in the development of disease because of inconsistent results across multiple studies. To resolve this issue present study was designed to test alleles as well as genotypes with an increasing probability of disease development. Many population-specific *PADI-4* gene association studies were conducted with RA development However up to the best of my knowledge, this is the -first-ever study that reported the association of polymorphism on the *PADI-4* gene with not only RA as well as OA development among Pakistani individuals.

The current study demonstrated that rs874881, rs11203367, rs2240336, rs2240339, rs1748033, rs2240340 were significantly associated with the onset of RA in the Pakistani population, However rs11203366 was not a significant risk factor to disease onset in the studied population. Similarly, two consecutive studies based on the Iranian population also revealed that rs874881, rs11203367,

Table 1
Primers and RFLP Enzymes of targeted PADI-4 gene polymorphic sites along with their optimized conditions.

SNP ID	Primers	A.T (°C)	RFLP enzyme	I.T (°C)	Time (Hours)	T.D (°C)
rs874881	F: AGCTTTTTGCTTTCCCTCCATT R: GCATGGTTTCTAGCCAGTCAGAC	56.4	HpaII	37	6	65
rs11203366	F: CCTCACTGCATCCTCTGCTTTC R: GAAGCCCATCCACACTGCCAC	63.5	BsuRI (HaeIII)	37	16	80
rs11203367			MlsI (MscI)	37	6	65
rs2240336	F: CTGGCCCAGGCACCACCAG R: AGGGTTTCGGCAGCTGTGCC	65	–	–	–	–
rs2240337			PasI	55	16	80
rs2240339			–	–	–	–
rs1748033	F: CATCACAGTTGTGGCCCCG R: GCGGGTGATGTCTGCGCCC	62.5	RsaI	37	6	80
SCV000807675			–	–	–	–
rs2240340			–	–	–	–
SCV000804840			–	–	–	–

A.T: Annealing Temperature of Primers; RFLP: Restriction Fragment Length Polymorphism; I.T: Incubation Temperature of Enzyme; T.D: Thermal Deactivation Temperature of Enzymes for 20 mins

Table 2
Clinical Characteristics of the patients included for serum and genetic analysis.

Sr. No.	Characteristics	RA (n = 300)		OA (n = 316)	
		Male (n = 71)	Females (n = 229)	Male (n = 100)	Females (n = 216)
1.	Age (Years)	38.62 (35.94 ± 41.30)	38.98 (37.42 ± 40.55)	55.27 (52.88 ± 57.66)	49.19 (47.79 ± 50.58)
2.	BMI (Kg/m^2)	24.56 (23.45 ± 25.68)	26.43 (25.70 ± 27.16)	30.40 (29.20 ± 31.61)	30.00 (29.32 ± 30.69)
3.	Diagnosis Age (Years)	33.06 (30.55 ± 35.56)	31.56 (29.97 ± 33.15)	46.95 (44.84 ± 49.06)	42.73 (40.61 ± 44.85)
4.	Disease Duration (Years)	5.83 (4.436 ± 7.223)	7.27 (6.464 ± 8.075)	7.06 (5.690 ± 8.436)	5.64 (4.985 ± 6.295)
5.	Positive Family History (%)	85.92	80.35	69.00	68.98

n = Number of Samples; BMI = Body Mass Index; Kg/m = Kilogram/meter; %= Percentage

Table 3
Significance of PADI-4 gene polymorphic sites at Allelic level among RA, OA and controls.

SNP's	Allele	RA Frequency (P/C)	X^2-value	p-value	OA Frequency (P/C)	X^2-value	p-value
rs874881	G C	0.787/0.038 0.213/0.962	693.234	0.002*	0.756/0.038 0.244/0.962	657.608	0.002*
rs11203366	G A	0.615/0.053 0.385/0.947	425.353	0.012	0.494/0.053 0.385/0.947	296.537	0.001*
rs11203367	C T	0.647/0.107 0.353/0.893	372.590	0.002*	0.460/0.107 0.540/0.893	261.381	0.001*
rs2240336	T C	0.553/0.127 0.447/0.873	243.375	0.002*	0.595/0.127 0.405/0.873	290.562	0.001*
rs2240337	C	1.000/1.000	–	–	1.000/1.000	–	–
rs2240339	C A	0.492/0.057 0.508/0.943	285.265	0.003*	0.573/0.057 0.427/0.943	278.625	0.002*
rs1748033	T C	0.443/0.083 0.557/0.917	200.423	0.001*	0.502/0.083 0.498/0.917	257.434	0.001*
SCV000807675	G T	1.000/0.000 0.000/1.000	408.001	0.005*	1.000/0.000 0.000/1.000	842.000	0.002*
rs2240340	C T	0.572/0.077 0.428/0.923	335.524	0.001*	0.459/0.077 0.541/0.923	226.673	0.002*
SCV000804840	C T	0.500/0.020 0.500/0.980	119.751	0.003*	0.500/0.020 0.500/0.980	119.751	0.003*

X^2: Chi Square; p-value: Significance at 0.01 level; P: Patients; C: Controls

and rs1748033 were significant risk factors in the onset of RA (Shamsian et al., 2016; Hashemi et al., 2015). In south Mexican population rs874881, rs11203366 and rs11203367 were reported as significant risk alleles of RA development (Banos-Hernmíndez et al., 2017). In the same pipeline, rs2240340 was reported as significant disease factor among Nordics (Plenge et al., 2005). Association of the studied gene was first time reported by (Suzuki et al., 2003), with RA in Japanese population. Later this linkage was evaluated in different populations and found significant in Korean (rs1748033), Chinese rs2240340 and rs1748033), German (rs874881, rs11203366 and rs2240340) and North American population (rs2240340) (Bang et al., 2010; Fan et al., 2008; Hoppe et al., 2006; Kang et al., 2006; Plenge et al., 2005).

In contrast to current findings no association of *PADI-4* gene and RA development was reported in the British, French, Spanish, and Indian populations (Burr et al., 2010; Caponi et al.,

Table 4
Significance of PADI-4 gene polymorphic sites at genotypic level among RA, OA and controls.

SNP's	Genotype	RA Frequency (P/C)	X^2-value	p-value	OA Frequency (P/C)	X^2-value	p-value
rs874881	GG GC CC	0.593/0.000 0.387/0.077 0.020/0.923	499.732	0.002*	0.535/0.000 0.443/0.077 0.022/0.923	509.599	0.001*
rs11203366	GG GA AA	0.337/0.040 0.557/0.027 0.107/0.933	411.688	0.001*	0.364/0.040 0.259/0.027 0.377/0.933	209.070	0.003*
rs11203367	CC CT TT	0.267/0.107 0.173/0.000 0.560/0.893	246.043	0.010	0.142/0.107 0.794/0.000 0.063/0.893	466.649	0.002*
rs2240336	TT TC CC	0.203/0.083 0.700/0.087 0.097/0.830	332.628	0.010	0.244/0.083 0.703/0.087 0.054/0.830	363.602	0.010
rs2240337	CC	1.000/1.000	–	–	1.000/1.000	–	–
rs2240339	CC CA AA	0.080/0.000 0.823/0.113 0.097/0.887	375.859	0.001*	0.145/0.000 0.855/0.113 0.000/0.887	291.299	0.003*
rs1748033	TT TC CC	0.000/0.037 0.887/0.093 0.113/0.870	378.341	0.010	0.146/0.037 0.712/0.093 0.142/0.870	327.162	0.010
Scv000807675	GG TT	1.000/0.000 0.000/1.000	421.000	0.002*	1.000/0.000 0.000/1.000	421.000	0.002*
rs2240340	CC CT TT	0.233/0.017 0.677/0.120 0.090/0.863	243.709	0.020	0.155/0.017 0.608/0.120 0.237/0.863	243.709	0.020
SCV000804840	CT TT	1.000/0.040 0.000/0.960	184.615	0.001*	1.000/0.040 0.000/0.960	184.615	0.001*

X^2: Chi Square; p-value: Significance at 0.01 level; P: Patients; C: Controls

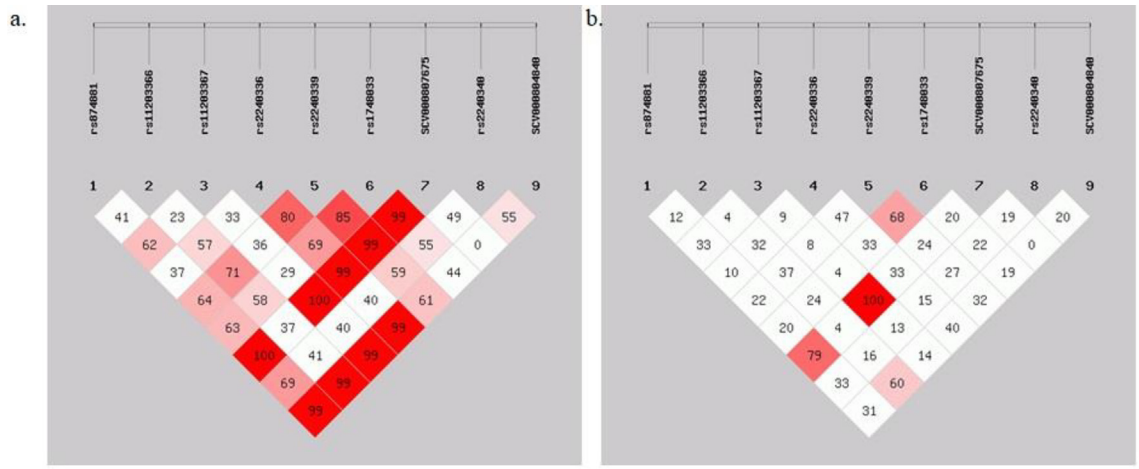

Fig. 2. (a, b): Linkage Disequlibrium chart of PADI-4 gene polymorphic sites in RA patients and controls. With increase in color intensity indicate the increase in risk of disease onset.

2005; Martinez et al., 2005; Panati et al., 2012). However the actual reasons for the ethnic discrepancy are not clear, but it is important to study differences in the fundamental genetic background and social factors between different populations. Ethnically various subjects may have unique cultures and lifestyles that can contribute to different genetic characteristics and susceptibility to specific diseases. It was reported that linkage between rs1748033 polymorphism and RA might be susceptible to differences in ethnicity because we found only one study from Spain, Hungary, and Iran, respectively (Lee and Bea et al., 2016; Hashemi et al., 2015). In a study comprised of meta-analysis, stratified by ethnicity, a significant association between rs1748033 polymorphism and RA risk was detected in China and Japan (Jian and Pu, 2020). A result of the meta-analysis by (Burr et al., 2010), reported that rs2240340 was more significant in the Asian population as compared to the European population. The variations in results in different populations may be because of a complicated environment as well as population heterogeneity.

It was observed in the present study that rs874881, rs11203366, rs11203367, rs2240336, rs2240339, rs1748033, rs2240340 were significant risk factors in the onset of OA at the allelic level whereas except rs2240336, rs1748033, and rs2240340 rest were associated with disease onset at the genotypic level in the studied population. This was the first-ever study revealing the association of OA and the *PADI-4* gene. However, the exact role of the *PADI-4* gene in the OA onset is still not fully understood and needs to be studied.

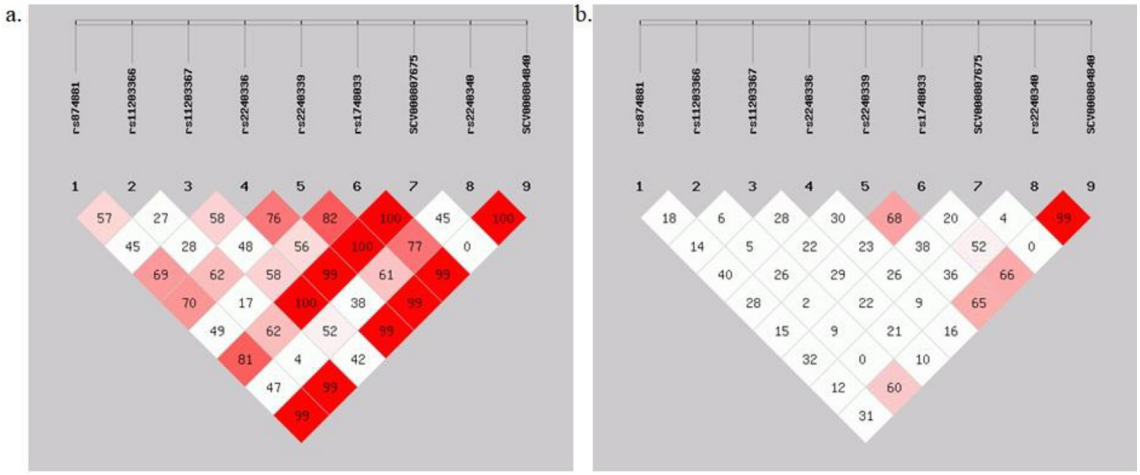

Fig. 3. (a, b): Linkage Disequlibrium chart of PADI-4 gene polymorphic sites in OA patients and controls. With increase in color intensity indicate the increase in risk of disease onset.

Table 5
Haplotype analysis among PADI-4 gene polymorphic sites.

Haplotype: rs874881; rs11203366; rs11203367; rs2240336; rs2240337; rs2240339; rs1748033; *SCV000807675*, rs2240340; *SCV000804840*					
Haplotype	**RA Frequency (P/C)**	**p-value**	**Haplotype**	**OA Frequency (P/C)**	**p-value**
GACCCACGCC	0.085/0.000	0.008*	GATTCACGCC	0.037/0.000	0.002*
GACCCACGCT	0.085/0.000	0.006*	GATTCACGCT	0.037/0.000	0.002*
GACCCCTGCC	0.027/0.000	0.001*	GATTCACGTC	0.037/0.000	0.001*
GACCCCTGCT	0.027/0.000	0.006*	GATTCACGTT	0.037/0.000	0.003*
GACCCCTGTC	0.049/0.000	0.002*	GATTCATGCC	0.037/0.000	0.005*
GACCCCTGTT	0.049/0.000	0.002*	GATTCATGTC	0.037/0.000	0.002*
GACTCACGCC	0.085/0.000	0.001*	GATTCATGTT	0.037/0.000	0.002*
GACTCACGCT	0.085/0.000	0.002*	GATTCCCGCC	0.037/0.000	0.001*
GACTCCTGCC	0.027/0.000	0.003*	GATTCCCGCT	0.037/0.000	0.001*
GACTCCTGCT	0.027/0.000	0.008*	GATTCCCGTC	0.037/0.000	0.002*
GACTCCTGTT	0.049/0.000	0.006*	GATTCCCGTT	0.037/0.000	0.001*
GGCCCACGCC	0.027/0.000	0.001*	GATTCCTGCC	0.037/0.000	0.006*
GGCCCACGCT	0.027/0.000	0.001*	GATTCCTGTT	0.037/0.000	0.008*
GGCCCACGTC	0.036/0.000	0.002*	GGCTCACGCC	0.014/0.000	0.005*
GGCCCACGTT	0.036/0.000	0.006*	GGCTCACGCT	0.014/0.000	0.001*
GGCCCCCGCC	0.020/0.000	0.005*	GGCTCACGTC	0.014/0.000	0.001*
GGCCCCCGCT	0.020/0.000	0.002*	GGCTCACGTT	0.014/0.000	0.002*
GGCTCACGCC	0.027/0.000	0.001*	GGCTCATGCC	0.014/0.000	0.002*
GGCTCACGCT	0.027/0.000	0.002*	GGCTCATGCT	0.014/0.000	0.001*
GGCTCACGTC	0.036/0.000	0.003*	GGCTCATGTC	0.014/0.000	0.003*
GGCTCACGTT	0.036/0.000	0.001*	GGCTCATGTT	0.014/0.000	0.004*
GGCTCCCGCC	0.020/0.000	0.001*	GGCTCCCGCC	0.014/0.000	0.001*
GGCTCCCGCT	0.020/0.000	0.002*	GGCTCCCGCT	0.014/0.000	0.009*
			GGCTCCCGTC	0.014/0.000	0.008*
			GGCTCCCGTT	0.014/0.000	0.002*
			GGCTCCTGCC	0.014/0.000	0.002*
			GGCTCCTGCT	0.014/0.000	0.002*
			GGCTCCTGTC	0.014/0.000	0.001*
			GGCTCCTGTT	0.014/0.000	0.001*

p-value: Significance at 0.01 level; P: Patients; C: Controls

Table 6
Amino acid alignment through Mega 6 software of functional polymorphic sites.

SNP ID	**Wild Amino Acid**	**Mutant Amino Acid**
Rs874881	Alanine	Lysine
Rs11203366	Serine	Lysine
Rs11203367	Valine	Alanine
Rs1748033	Leucine	Leucine
SCV000807675	Glysine	Cystein
SCV000804840	Histamine	Tyrosine

Two functional novel SNPs were also identified in the Pakistani population SCV000807675 and SCV000804840 in both RA and OA subjects. SCV000807675 lead to the change of glycine into cysteine whereas SCV000804840 polymorphism replaced histamine from tyrosine and ultimately cause dysfunctioning of the protein encoded by the *PADI-4* gene by altering its structure.

Haplotype analysis comprising of 10 ten SNPs was performed to determine their role in the pathogenesis of the disease. It was revealed that all the haplotypes played a major role in the pathogenesis of diseases. Similarly, it was reported that SNPs on the coding region were associated with the stability of the mRNA (Yang et al., 2003; Zhou et al., 2002). Susceptible haplotypes may increase the stability of the *PADI-4* protein in monocytes, synovial tissues, and neutrophils and may also elevate citrullinated peptides production (Suzuki et al., 2003). Similar results were also reported in southern as well as Mexican population (Banos-Hernmíndez et al., 2017).

Overall it was demonstrated that a strong association exists between *PADI-4* gene polymorphism and the onset of not only RA as well as OA in the Pakistani population. Thus these observations need to be analyzed in RA as well as OA subjects on a larger scale.

5. Conclusion

In conclusion, our case-control study indicated that the *PADI-4* gene is a vital genetic risk factor in the onset of RA as well as OA. The study should be extended up to a larger population size as well as should be replicated with the OA subjects of different cohorts.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We would like to acknowledge this study to the Vice-Chancellor, University of the Punjab Lahore for ethically allow us to conduct this study as well as Higher Education Commission Pakistan for providing funds for the conduction of the study.

References

Bang, S.-Y., Han, T.-U., Choi, C.-B., Sung, Y.-K., Bae, S.-C., Kang, C., 2010. Peptidyl arginine deiminase type IV (*PADI-4*) haplotypes interact with shared epitope regardless of anti-cyclic citrullinated peptide antibody or erosive joint status in rheumatoid arthritis: a case control study. Arthritis Res. Therapy 12 (3), R115. https://doi.org/10.1186/ar3051.

Baños-Hernández, C.J., Navarro-Zarza, J.E., Parra-Rojas, I., Vázquez-Villamar, M., Ramón Padilla-Gutiérrez, J., Valle, Y., Reyes-Castillo, Z., Magdalena Torres-Carrillo, N., García-Arellano, S., Brennan-Bourdon, L.M., Muñoz-Valle, J.F., 2017. *PADI-4* polymorphisms and the functional haplotype are associated with increased rheumatoid arthritis susceptibility: A replication study in a Southern Mexican population. Hum. Immunol. 78 (9), 553–558.

Barton, A., Bowes, J., Eyre, S., Spreckley, K., Hinks, A., John, S., Worthington, J., 2004. A functional haplotype of the *PADI-4* gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population. Arthritis Rheumatism: Off. J. Am. College Rheumatol. 50 (4), 1117–1121.

Burr, M.L., Naseem, H., Hinks, A., Eyre, S., Gibbons, L.J., Bowes, J., Wilson, A.G., Maxwell, J., Morgan, A.W., Emery, P., Steer, S., Hocking, L., Reid, D.M., Wordsworth, P., Harrison, P., Thomson, W., Worthington, J., Barton, A., 2010. *PADI-4* genotype is not associated with rheumatoid arthritis in a large UK Caucasian population. Ann. Rheum. Dis. 69 (4), 666–670.

Caponi, L., Petit-Teixeira, E., Sebbag, M., Bongiorni, F., Moscato, S., Pratesi, F., et al., 2005. A family based study shows no association between rheumatoid arthritis and the *PADI-4* gene in a white French population. Ann. Rheum. Dis. 64, 587–593.

Cha, S., Choi, C., Han, T., Kang, C.P., Kang, C., Bae, S., 2007. Association of anti□Ç ôcyclic citrullinated peptide antibody levels with *PADI-4* haplotypes in early rheumatoid arthritis and with shared epitope alleles in very late rheumatoid arthritis. Arthritis Rheum. 56, 1454–1463.

Chen, R., Wei, Y., Cai, Q., Duan, S., Ren, D., Shen, J., He, D., Fang, M., Lv, K.e., Cheng, N., Sun, S., 2011. The *PADI-4* gene does not contribute to genetic susceptibility to rheumatoid arthritis in Chinese Han population. Rheumatol. Int. 31 (12), 1631–1634.

Cheng, J., Zhang, H., Zhuang, C., Liu, R., 2012. Peptidylarginine deiminase type 4 and methyl-CpG binding domain 4 polymorphisms in Chinese patients with rheumatoid arthritis. J. Rheumatol. 39 (6), 1159–1165.

El-Gabalawy, H.S., Robinson, D.B., Daha, N.A., Oen, K.G., Smolik, I., Elias, B., Hart, D., Bernstein, C.N., Sun, Y., Lu, Y., Houwing-Duistermaat, J.J., Siminovitch, K.A., 2011. Non-HLA genes modulate the risk of rheumatoid arthritis associated with HLA-DRB1 in a susceptible North American Native population. Genes Immun. 12 (7), 568–574.

Fan, L.Y., Wang, W.J., Wang, Q., Zong, M., Yang, L., Zhang, H., et al., 2008. A functional haplotype and expression of the PADI-4 gene associated with increased rheumatoid arthritis susceptibility in Chinese. Tissue Antigens 72, 469–473.

Goulielmos, G.N., Zervou, M.I., Myrthianou, E., Burska, A., Niewold, T.B., Ponchel, F., 2016. Genetic data: the new challenge of personalized medicine, insights for rheumatoid arthritis patients. Gene 583 (2), 90–101.

Hashemi, M., Zakeri, Z., Taheri, H., Bahari, G., Taheri, M., 2015. Association between peptidylarginine deiminase type 4 rs1748033 polymorphism and susceptibility to rheumatoid arthritis in Zahedan, Southeast Iran. Iranian J. Allergy, Asthma Immunol. 14, 255–260.

Hassine, H.B., Zemni, R., Bouagina, E., Zaglaoui, H., Ben Fradj, H., Slama, F., Boukadida, J., Sghiri, R., 2012. Lack of association between *PADI-4* polymorphisms and rheumatoid arthritis in the Tunisian population. Joint Bone Spine 79 (3), 329–330.

Hoppe, B., Heupl, T., Gruber, R., Kiesewetter, H., Burmester, G.R., Salama, A., et al., 2006. Detailed analysis of the variability of peptidylarginine deiminase type 4 in German patients with rheumatoid arthritis: a case–control study. Arthritis Res. Therapy 8, R34.

Hough, A.J., 2001. Pathology of osteoarthritis. Arthritis Allied Conditions 2, 2180.

Jian, C., Pu, J., 2020. The Relationship Between PADI-4-94G/A Polymorphism and Rheumatoid Arthritis: A Meta-analysis in Multi Ethnic Groups.

Kang, C.P., Lee, H.-S., Ju, H., Cho, H., Kang, C., Bae, S.-C., 2006. A functional haplotype of the *PADI-4* gene associated with increased rheumatoid arthritis susceptibility in Koreans. Arthritis Rheumatism: Off. J. Am. College Rheumatol. 54 (1), 90–96.

Lee, Y.H., Bae, S.C., 2016. Association between susceptibility to rheumatoid arthritis and *PADI-4* polymorphisms: a meta-analysis. Clin. Rheumatol. 35 (4), 961–971.

Li, Q., Lin, K.Q., Wang, J., Yu, J., Yu, L., Yi, W., Yang, Z.Q., 2013. Association of polymorphisms of PTPN22 and *PADI-4* genes with rheumatoid arthritis in Yunnan. *Zhonghua yi xue yi chuan xue za zhi= Zhonghua yixue yichuanxue zazhi = Chinese J. Med. Genet.* 30 (1), 111–115.

MacGregor, A.J., Snieder, H., Rigby, A.S., Koskenvuo, M., Kaprio, J., Aho, K., Silman, A. J., 2000. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheumatism: Off. J. Am. College Rheumatol. 43 (1), 30–37.

Martinez, A., Valdivia, A., Pascual-Salcedo, D., Lamas, J.R., Fernandez-Arquero, M., Balsa, A., et al., 2005. PADI-4 polymorphisms are not associated with rheumatoid arthritis in the Spanish population. Rheumatology 44, 1263–1266.

Mirza, R.R., 2016. Types of Anaemia and its Correlation with Disease Activity in Patients with Rheumatoid Arthritis among Kurdish Population of Iraq.

Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S., Graham, R.R., Manoharan, A., Ortmann, W., Bhangale, T., Denny, J. C., Carroll, R.J., Eyler, A.E., Greenberg, J.D., Kremer, J.M., Pappas, D.A., Jiang, L., Yin, J., Ye, L., Su, D.-F., Yang, J., Xie, G., Keystone, E., Westra, H.-J., Esko, T., Metspalu, A., Zhou, X., Gupta, N., Mirel, D., Stahl, E.A., Diogo, D., Cui, J., Liao, K., Guo, M.H., Myouzen, K., Kawaguchi, T., Coenen, M.J.H., van Riel, P.L.C.M., van de Laar, M.A.F.J., Guchelaar, H.-J., Huizinga, T.W.J., Dieudé, P., Mariette, X., Louis Bridges Jr, S., Zhernakova, A., Toes, R.E.M., Tak, P.P., Miceli-Richard, C., Bang, S.-Y., Lee, H.-S., Martin, J., Gonzalez-Gay, M.A., Rodriguez-Rodriguez, L., Rantapää-Dahlqvist, S., Ärlestig, L., Choi, H.K., Kamatani, Y., Galan, P., Lathrop, M., Eyre, S., Bowes, J., Barton, A., de Vries, N., Moreland, L.W., Criswell, L.A., Karlson, E.W., Taniguchi, A., Yamada, R., Kubo, M., Liu, J.S., Bae, S.-C., Worthington, J., Padyukov, L., Klareskog, L., Gregersen, P.K., Raychaudhuri, S., Stranger, B.E., De Jager, P.L., Franke, L., Visscher, P.M., Brown, M.A., Yamanaka, H., Mimori, T., Takahashi, A., Xu, H., Behrens, T.W., Siminovitch, K.A., Momohara, S., Matsuda, F., Yamamoto, K., Plenge, R.M., 2014. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 506 (7488), 376–381.

Panati, K., Pal, S., Kv, R., Reddy, V.D., 2012. Association of single nucleotide polymorphisms (SNPs) of PADI-4 gene with rheumatoid arthritis (RA) in Indian population. Genes Genetic Syst. 87, 191–196.

Plenge, R.M., Padyukov, L., Remmers, E.F., Purcell, S., Lee, A.T., Karlson, E.W., et al., 2005. Replication of putative candidate-gene associations with rheumatoid arthritis in >4000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and *PADI-4*. Am. J. Human Genet. 77, 1044–1060.

Shamsian, E., Azarian, M., Akhlaghi, M., Samaei, M., Jamshidi, A.R., Assar, S., et al., 2016. *PADI-4* Polymorphisms in Iranian Patients with Rheumatoid Arthritis. Acta Reumatologica Portuguesa 41, 338–343.

Suzuki, A., Yamada, R., Chang, X., Tokuhiro, 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 *PADI-4*, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat. Genet. 34 (4), 395–402.

Yang, T., McNally, B.A., Ferrone, S., Liu, Y., Zheng, P., 2003. A Single-nucleotide Deletion Leads to Rapid Degradation ofTAP-1 mRNA in a Melanoma Cell Line. J. Biol. Chem. 278 (17), 15291–15296.

Zavala-Cerna, M.G., Gonzalez-Montoya, N.G., Nava, A., Gamez-Nava, J.I., Moran-Moguel, M.C., Rosales-Gomez, R.C., et al., 2013. PADI-4 haplotypes in association with RA Mexican patients, a new prospect for antigen modulation. Clin. Dev. Immunol. 2013.

Zhou, X., Tan, F.K., Reveille, J.D., Wallis, D., Milewicz, D.M., Ahn, C., Wang, A., Arnett, F.C., 2002. Association of novel polymorphisms with the expression of SPARC in normal fibroblasts and with susceptibility to scleroderma. Arthritis Rheumatism: Off. J. Am. College Rheumatol. 46 (11), 2990–2999.

Zhou, Z., Ménard, H.-A., 2002. Autoantigenic posttranslational modifications of proteins: does it apply to rheumatoid arthritis? Curr. Opin. Rheumatol. 14 (3), 250–253.