Age at onset of rheumatoid arthritis: association with polymorphisms in the vascular endothelial growth factor A (*VEGFA*) gene and an intergenic locus between matrix metalloproteinase (*MMP*) 1 and 3 genes

Y. Chen, D.L. Mattey

Haywood Rheumatology Centre, Haywood Hospital, Stoke-on-Trent, Staffordshire, and Institute of Science and Technology in Medicine, Keele University, Staffordshire, United Kingdom.

Abstract

Objective

*The present paper aims to investigate whether polymorphisms in the vascular endothelial growth factor A (*VEGFA*) gene and the loci of matrix metalloproteinase (*MMP*) 1 and 3 genes are associated with age at onset of RA.*

Methods

A sample of 413 hospital-based RA patients of Caucasian origin was studied. Common single-nucleotide polymorphisms (SNP) with likely importance were typed, including rs699947, rs833061, rs2010963 and rs3025039 in VEGFA*, and rs1799750 in the* MMP1 *gene, rs3025058, rs679620 in the* MMP3 *gene and rs495366 located within the region between the* MMP1 *and* MMP3 *genes. Age at onset of RA was obtained on each patient. Demographic variables, smoking information, and a core set of clinical characteristics measured at recruitment were recorded. Hazard ratios (HR) that measured the effect size of genetic risk on age at RA onset were computed using Cox regression models.*

Results

The T allele at rs3025039 was associated with an increased risk of early onset (HR=1.25 [95% CI 1.0–1.58] for the risk over time; HR=1.84 [95% CI 1.20–2.83] for the risk of onset <40 years old). The AA genotype at rs495366 was also associated with an increased risk (HR=1.92 [95% CI 1.27–2.89] over time; HR=2.54 [95% CI 1.30–4.95] for onset <40 years old). These associations were independent of other risk factors such as sex, smoking and anti-CCP status.

Conclusion

Polymorphisms in the VEGFA *gene and the* MMP1-3 *intergenic locus may influence age at onset of RA.*

Key words

rheumatoid arthritis, *VEGFA*, *MMP*, polymorphism, age at onset

Ying Chen, PhD
Derek L. Mattey, PhD

Please address correspondence to:
Derek L. Mattey, PhD,
Haywood Rheumatology Centre,
High Lane,
Burslem,
ST6 7AG Stoke-on-Trent,
United Kingdom.
E-mail: d.l.mattey@keele.ac.uk

Received on November 28, 2011; accepted in revised form on January 24, 2012.

Funding: this work was supported by the Medical Research Council and the Haywood Rheumatism Research and Development Foundation.

Competing interests: none declared.

Introduction

Age at onset of rheumatoid arthritis (RA) has been reported to be influenced by genetic factors. These include polymorphisms in several cytokine genes (1-5), and the major genetic risk factors for RA, the *HLA-DRB1* shared epitope (SE), and the R620W allele in the protein tyrosine phosphatase, non-receptor type 22 (*PTPN22*) gene (5-7).

The initial development of RA is associated with active tissue neovascularisation and remodelling. Genes, important in these physiological processes, may therefore have an impact, possibly on the alteration (advance/delay) of disease onset. Vascular endothelial growth factor-A (VEGF-A) plays a fundamental role in the regulation of angiogenesis. It induces angiogenesis by acting on vascular endothelial cells, and promoting cell mitogenesis, cell migration and lumen formation. Matrix metalloproteinases (MMPs), a group of zinc- and calcium-dependent proteases, contribute to tissue remodelling and angiogenesis by degrading the components of the extracellular matrix (ECM), which serves as a scaffold for tissue formation. MMP-1 (collagenase-1) and MMP-3 (stromelysin-1) are among the most important members, since MMP-1 is specifically responsible for the degradation of the interstitial collagens (types I, II and III), while MMP-3 has a very broad substrate and is able to proteolytically activate other MMPs (*e.g.* MMP-1, -2, -9, -13).

The objective of this study was to investigate the effects of common single-nucleotide polymorphisms (SNP) in the *VEGFA* gene and the loci of *MMP1* and *3* genes on age at onset of RA. Polymorphisms in these genes have been shown to be associated with susceptibility and/or disease activity in RA (8, 9), so they were considered good candidates for an effect on age of disease onset.

Materials and methods

Patients

This study was carried out on a population (n=413) of RA patients (1987 ACR criteria) of Caucasian background, resident in North Staffordshire and attending the Clinical Rheumatology Unit at the Haywood Hospital. Written informed consent was provided by each patient according to the Declaration of Helsinki. The study was approved by the North Staffordshire local research ethics committee.

Age at onset of RA, based on age of diagnosis, was obtained on each patient. Demographic and clinical variables (Table I) were recorded, and peripheral bloods (for DNA isolation and serum preparation) were collected at recruitment.

SNP typing

Four SNPs in the *VEGFA* gene (rs699947, rs833061, rs2010963 and rs3025039), one SNP in the *MMP1* gene (rs1799750), two SNPs in the *MMP3* gene (rs3025058, rs679620) and one SNP located within the region between *MMP1* and *MMP3* genes (rs495366) were selected (Table II). All genotypes were determined by the PCR-RFLP method as described previously (9, 10).

Statistical analysis

Hardy-Weinberg equilibrium (HWE) for the genotypic distribution was tested by the Chi-Square goodness-of-fit test. The relationship between categorical and quantitative variables was assessed using *t*-test or one-way ANOVA. Cox regression models were applied to analyse the genetic hazard in relation to age at onset of RA, *i.e.* the disease free survival time. No censor was used for analysis of the risk over time, while for analysis of risk for developing early RA, a censor (age at onset <40) was used. Difference in disease free survival curves was compared by the Gehan-Breslow-Wilcoxon test using GraphPad Prism (version 5.0). Multivariate multiple regression analysis was used to assess the independence of the associations found, and to adjust for potential confounding factors. Unless specified otherwise, analyses were all carried out using the Number Cruncher Statistical System for Windows (version NCSS 2000). Results with p-values <0.05 (after Bonferroni correction when necessary) were considered significant.

Results

Characteristics of patients

Table I shows the selected demographic and clinical variables of 413 subjects

Table I. Demographic and clinical characteristics of patients with rheumatoid arthritis.

Characteristics	
Female gender	278/413 (67.3)
Age at onset	49.9 ± 13.0
Early onset (<40 years old)	86/413 (20.8)
Ever smoking*	273/410 (66.6)
Clinical variables at the time of recruitment:	
Disease duration, year	9.0 (3.0–17.0)
Rheumatoid factor (RF) status	232/410 (56.6)
Anti-cyclic citrullinated peptide (anti-CCP) status	302/400 (75.5)
Erythrocyte sedimentation rate (ESR), mm/hour	19.0 (10.0–34.0)
C-reactive protein (CRP) (≥10 mg/l)	220/412 (53.4)
Disease activity score 28 (DAS28)	4.18 ± 1.37
Health assessment questionnaire (HAQ) score	1.625 (1.0–2.0)
Erosive disease	300/407 (73.7)
Nodular disease	53/413 (12.8)

Values are number (%), or mean ± (standard deviation) when quantitative variables fit normality or median (interquartile range) when quantitative variables do not fit normality. *Patients classified as ever smokers were those who had smoked at least 1 cigarette per day for at least 1 year.

from a hospital based population of RA patients. At recruitment, the median age of these patients was 62.0 (interquartile range [IQR] 54.5–69.0).

Distributions of SNP

Genotypes of all SNP were distributed in accordance with a close fit to HWE. The genotypic frequency of each SNP is shown in Table II. Linkage disequilibrium (LD) was detected across the 3 *VEGFA* SNP located in the 5'-flanking region, and across the 4 *MMP1-3* SNP, as described previously based on the same population (9, 10).

Association of VEGFA SNP with age at onset

The T allele (CT + TT) at rs3025039 was associated with earlier onset of RA, compared to the CC genotype (Table II). No association was found with other *VEGFA* SNP or with any haplotype. Survival curves for rs3025039 (T allele *vs*. CC), where age at onset of RA is the disease free survival time, are shown in Figure 1a. The hazard ratio (HR) that compared the relative risk of RA onset in patients with the T allele *versus* patients of equal age without the T allele was 1.25 (95% CI 1.0–1.58). Thus, at any particular age, the risk of RA onset for a patient with the T allele was higher than that for a patient carrying the CC genotype. In terms of the risk of onset of early disease (age at onset <40), the T allele was associated with higher risk (HR 1.84, 95% CI 1.20–2.83), compared with the CC genotype at any particular age before 40.

Association of MMP1 *and* MMP3 *SNP with age at onset*

The AA genotype at rs495366 was associated with earlier onset of RA, compared with the AG and GG genotypes combined (Table II). No other *MMP1-3* SNP or any haplotype was associated. Disease free survival curves for rs495366 (AA *vs*. G allele) are shown in Figure 1b. The hazards of RA onset over time and before the age of 40 for a patient carrying the AA genotype were higher (HR 1.92, 95% CI 1.27–2.89 and HR 2.54, 95% CI 1.30–4.95, respectively), compared to those for a patient carrying the G allele.

Multivariate associations with age at onset

Analysis in multivariate multiple regression indicated that the genetic associations were significant after adjustment for sex, disease duration, ever smoking and anti-CCP status (measured at recruitment), and were independent of each other (Table III).

Discussion

The present study has demonstrated that variation in the *VEGFA* gene and an intergenic locus of the *MMP1-3* gene region may influence age at onset of RA in a UK Caucasian population. As far as we are aware this is the first study

Table II. Age at onset of rheumatoid arthritis stratified by the genotypes of VEGFA or MMP1-3 polymorphism.

Common SNP in the vascular endothelial growth factor A (VEGFA) gene

rs699947(A/C)	Age at onset	*p*-value	rs833061(C/T)	Age at onset	*p*-value	rs2010963(G/C)	Age at onset	*p*-value	rs3025039(C/T)	Age at onset	*p*-value
AA (110, 26.6%)	50.0 ± 13.4		CC (103, 24.9%)	49.0 ± 13.4		GG (196, 47.5%)	50.3 ± 13.0		CC (296, 71.7%)	51.1 ± 12.1	
AC (199, 48.2%)	50.3 ± 12.8		CT (208, 50.4%)	50.7± 12.6		GC (177, 42.9%)	50.0 ± 13.0		CT (103, 24.9%)	46.8 ± 14.8	
CC (104, 25.2%)	49.2 ± 13.0	0.80*	TT (102, 24.7%)	49.1 ± 13.2	0.42*	CC (40, 9.7%)	47.6 ± 12.8	0.49*	TT (14, 3.4%)	47.3 ± 11.8	0.011*
									CT + TT	46.9 ± 14.5	0.0026†,¶

Common SNP in the loci for the matrix metalloproteinase (MMP) 1 and 3 genes

rs3025058(6A/5A)	Age at onset	*p*-value	rs679620(G/A)	Age at onset	*p*-value	rs495366(A/G)	Age at onset	*p*-value	rs1799750(2G/1G)	Age at onset	*p*-value
6A6A (85, 20.6%)	47.8 ± 12.4		GG (86, 20.8%)	48.0 ± 12.5		AA (25, 6.1%)	42.2 ± 12.8		2G2G (99, 24.0%)	48.5 ± 14.1	
6A5A (220, 53.3%)	50.6 ± 12.6		GA (221, 53.5%	50.2 ± 12.6		AG (177, 42.9%)	51.4 ± 12.4		2G1G (191, 46.3%)	50.2 ± 12.3	
5A5A (108, 26.2%)	50.2 ± 14.0	0.24*	AA (106, 25.7%)	50.8 ± 14.1	0.30*	GG (211, 51.1%)	49.6 ± 13.2	0.0031*¶	1G1G (123, 29.8%)	50.6 ± 13.1	0.45*
						AG + GG	50.4 + 12.8	0.0020‡¶			

Selection of SNP was based on the likely importance of these variations in expression/protein level, protein function and/or association with a particular trait according to the current literature. Rs699947 (promoter), rs833061 (promoter), rs2010963 (5'-untranslated) and rs3025039 (3'-untranslated) are commonly known as VEGFA-2578, -460, +405 and +936, respectively. *MMP1* and *MMP3* are neighbouring genes, and *MMP3* is located upstream of *MMP1*. Rs3025058 (*MMP3* promoter), rs679620 (*MMP3* exon 2) and rs1799750 (*MMP1* promoter) are commonly known as *MMP3* 5A/6A, *MMP3* Lys45Glu and *MMP1* 1G/2G respectively, whereas there is no common name for the intergenic rs495366 (upstream region of *MMP1*). Values are mean ± standard deviation. *One-way ANOVA (any difference between 3 genotype groups); †CT + TT (T allele) *vs*. CC, Student's *t*-test; ‡AA *vs*. AG + GG (G allele), Student's *t*-test; ¶remained significant after the Bonferroni correction for multiple testing where the significant level was set at a *p*-value of 0.00625 (0.05/8).

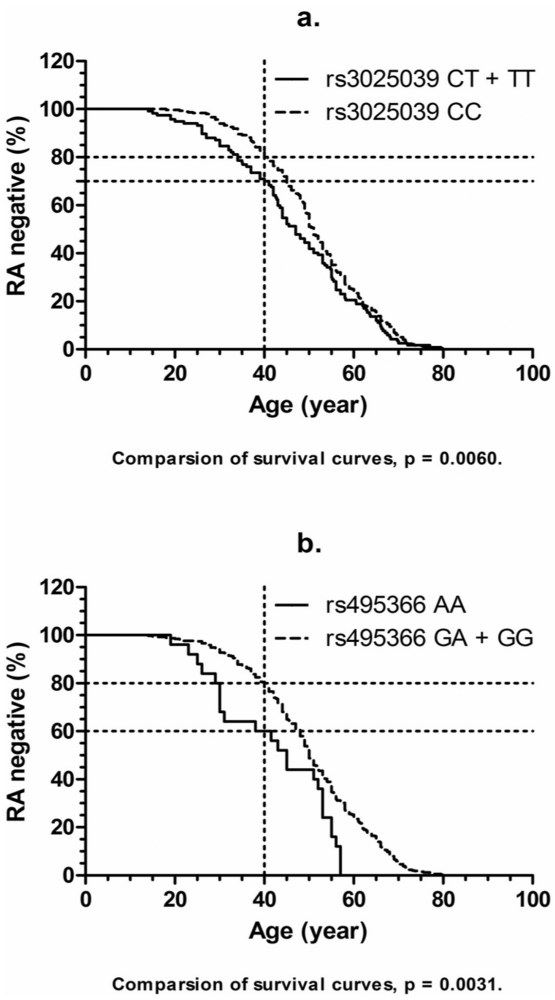

Fig. 1. Disease free survival curves stratified by the *VEGFA* rs3025039 (CT + TT *vs*. CC) or intergenic MMP1-3 rs495366 (AA *vs*. AG + GG) SNP according to age at onset of rheumatoid arthritis
p-values are obtained from the Gehan-Breslow-Wilcoxon test.

Table III. Multivariate multiple regression analysis of variables associated with age at onset of rheumatoid arthritis.

Dependent variable: age at onset of rheumatoid arthritis		
Independent variable	Regression coefficient (standard error)	*p*-value
Rs3025039 T allele	-3.256 (1.139)	0.0045
Rs495366 AA genotype	-4.988 (2.157)	0.021
Female	-3.445 (1.119)	0.0022
Ever smoking	2.049 (1.120)	0.068
Anti-CCP †	-2.303 (1.210)	0.058
Disease duration*,†	-0.765 (0.053)	<0.0001

The current model was based on 397 patients whose anti-CCP and smoking data were both available. Models based on 410 patients (exclusion of anti-CCP data), on 400 patients (exclusion of smoking data) or on 413 patients (exclusion of both variables) showed consistent results in terms of the associations presented. *Control for potential bias due to disease duration and related factors (*e.g.* premature death) during sample recruitment; †the significance level for the association of anti-CCP+ (measured at recruitment) with early onset was largely reduced (from p=0.0013 to 0.058) after adjustment for disease duration.

to show that polymorphisms in genes associated with angiogenesis and tissue remodelling may influence the age at onset of RA in a Caucasian population. This suggests that the onset of pathological changes in the development of RA may be determined by particular polymorphisms involved in these closely related physiological processes.

The association of the T allele at rs3025039 with earlier onset is in line with the observation of Han *et al*. who found a similar but non-significant trend in a Korean RA population (8). In their report, the T allele frequency was significantly higher in patients compared to that in controls, indicating this allele as a risk factor for RA susceptibility (8). In contrast, no polymorphism in the 5'-flanking region was found to be related to susceptibility or age at onset (8, 11).

The relationship between *VEGFA* polymorphism and protein level has been investigated. A study investigating disease associated with abnormal angiogenesis (epithelial ovarian cancer) showed that the 5'-flanking polymorphisms (rs699947, rs833061 and rs2010963) were associated with serum levels of VEGF-A, whereas the 3'-flanking polymorphism (rs3025039) was not (although an increased trend with the T allele was seen) (12). In this RA population, preliminary data on sera (collected at recruitment) demonstrated similar relationships, with only the 5'-flanking polymorphisms (rs699947 and rs833061) showing significant associations with VEGF-A levels (unpublished observations). However, data based on a healthy population indicated that rs3025039 was associated with altered serum levels, with the T allele associated with higher levels, whereas no association was seen with 5'-flanking SNP (13). It has been shown that there are elements sensitive to hypoxia in the upstream region of the *VEGFA* gene (14). Taking these data together, it is hypothesised that the 5'-flanking region may serve as a hypoxia-induced regulator and provide a predominant effect in diseases associated with pathological angiogenesis (*e.g.* cancer, RA), whereas in healthy or pre-disease conditions the regulatory effect from the 3'-flanking region may be more important.

Previous studies did not find an association between *MMP1* or *MMP3* polymorphisms and RA susceptibility, although we have reported an association of the *MMP3* polymorphisms (rs3025058 and rs679620) with disease activity (9). The intergenic SNP rs495366 is in strong LD (*D'*>0.93) with these *MMP3* polymorphisms, and in moderate LD (*D'*=0.47) with *MMP1* SNP rs1799750 (9). However, the correlation of the rs495366 SNP with the linked SNP is

only moderate or low ($r^2 \leq 0.38$), due to very different frequencies of alleles within the same haplotype (9). This may explain why the association with early disease onset was found only with the relatively rare rs495366 AA genotype and no other *MMP* SNP.

Previous observations in a genome-wide association study showed rs495366 as the strongest genetic marker associated with serum levels of *MMP-1* in healthy subjects, with the AA genotype associated with the lowest levels (15). In our RA population, rs495366 was also shown to be associated with lower serum *MMP-1* levels (9). However, other *MMP1* and *3* polymorphisms were also independently associated with *MMP-1* levels (9). It therefore remains unclear whether the association of rs495366 with age at onset is attributable to *MMP-1* levels. Due to the cross sectional nature of this study it was not possible to determine whether there was any relationship between *MMP* SNP and serum *MMP* level prior to disease onset. Further studies, preferably on a population-based prospective cohort will be needed to confirm these results.

Acknowledgements

We wish to thank Mrs Michelle Kirwan, Mrs Ann Barcroft (research nurses) and Mrs Janet Turner (research assistant) for their help with data collection and entry.

References

1. VINASCO J, BERAÚN Y, NIETO A *et al*.: Polymorphism at the TNF loci in rheumatoid arthritis. *Tissue Antigens* 1997; 49: 74-8.
2. YAMADA R, TANAKA T, UNOKI M *et al*.: Association between a single-nucleotide polymorphism in the promoter of the human interleukin-3 gene and rheumatoid arthritis in Japanese patients, and maximum-likelihood estimation of combinatorial effect that two genetic loci have on susceptibility to the disease. *Am J Hum Genet* 2001; 68: 674-85.
3. PASCUAL M, NIETO A, MATARÁN L, BALSA A, PASCUAL-SALCEDO, MARTÍN J: IL-6 promoter polymorphisms in rheumatoid arthritis. *Genes Immun* 2000; 1: 338-40.
4. LO SF, HUANG CM, LIN HC, CHEN WC, TSAI CH, TSAI FJ: Cytokine (IL-6) and chemokine (IL-8) gene polymorphisms among rheumatoid arthritis patients in Taiwan. *Clin Exp Rheumatol* 2008; 26: 632-7.
5. WU H, KHANNA D, PARK G *et al*.: Interaction between RANKL and HLA-DRB1 genotypes may contribute to younger age at onset of seropositive rheumatoid arthritis in an inception cohort. *Arthritis Rheum* 2004; 50: 3093-103.
6. MACGREGOR A, OLLIER W, THOMSOM W, JAWAHEER D, SILMAN A: HLA-DRB1-*0401/0404 genotype and rheumatoid arthritis: increased association in men, young age at onset, and disease severity. *J Rheumatol* 1995; 22: 1032-6.
7. STEER S, LAD B, GRUMLEY JA, KINGSLEY GH, FISHER SA: Association of R602W in a protein tyrosine phosphatase gene with a high risk of rheumatoid arthritis in a British population: evidence for an early onset/disease severity effect. *Arthritis Rheum* 2005; 52: 358-60.
8. HAN SW, KIM GW, SEO JS *et al*.: VEGF gene polymorphisms and susceptibility to rheumatoid arthritis. *Rheumatology* (Oxford) 2004; 43: 1173-7.
9. CHEN Y, NIXON NB, DAWES PT, MATTEY DL: Influence of variations across the MMP-1 and -3 genes on the serum levels of MMP-1 and -3 and disease activity in rheumatoid arthritis. *Genes Immun* 2012; 13: 29-37.
10. CHEN Y, DAWES PT, PACKHAM JC, MATTEY DL: Interaction between smoking and polymorphism in the promoter region of the VEGFA gene is associated with ischemic heart disease and myocardial infarction in rheumatoid arthritis. *J Rheumatol* 2011; 38: 802-9.
11. RUEDA B, GONZÁLEZ-GAY MA, LÓPEZ-NEVOT MA *et al*.: Analysis of vascular endothelial growth factor (VEGF) functional variants in rheumatoid arthritis. *Hum Immunol* 2005; 66: 864-8.
12. STEFFENSEN KD, WALDSTRØM M, BRANDSLUND I, JAKOBSEN A: The relationship of VEGF polymorphisms with serum VEGF levels and progression-free survival in patients with epithelial ovarian cancer. *Gynecol Oncol* 2010; 117: 109-16.
13. AL-HABBOUBI HH, SATER MS, ALMAWI AW, AL-KHATEEB GM, ALMAWI WY: Contribution of VEGF polymorphisms to variation in VEGF serum levels in a healthy population. *Eur Cytokine Netw* 2011; 22: 154-8.
14. LIU Y, COX SR, MORITA T, KOUREMBANAS S: Hypoxia regulates vascular endothelial growth factor gene expression in endothelial cells. Identification of a 5' enhancer, *Circ Res* 1995; 77: 638-43.
15. CHENG YC, KAO WHL, MITCHELL BD *et al*.: Genome-wide association scan identifies variants near matrix metalloproteinase (MMP) genes on chromosome 11q21-22 strongly associated with serum MMP-1 levels. *Circ Cardiovasc Genet* 2009; 2: 329-37.