Genetic Associations of Rheumatoid Arthritis in Chinese

LI, Martin

A Thesis Submitted in Partial Fulfilment

of the Requirements for the Degree of

Doctor of Philosophy

in

Medical Sciences

The Chinese University of Hong Kong

September 2011

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Abstract

Genetic Associations of Rheumatoid Arthritis in Chinese

Submitted by Martin Li for thesis submitted for the degree of Doctor of Philosophy in Medical Sciences at The Chinese University of Hong Kong in April 2011

Rheumatoid arthritis is an autoimmune disease characterized by the swelling and inflammation of the joints. It causes irreversible damage to the synovium, cartilage and bone. The incidence of RA in women is three times higher than that in men. The prevalence of RA in Chinese is 0.35%.

We investigated the genetic predisposition from 24 SNPs in 600 RA patients and 900 controls. Seven SNPs were associated with increased risk for RA with polymorphisms in *HLA-DRB1* (p for trend = 1.8×10^{-4}), *HLA-DQA1* (p for trend = 2.0×10^{-16}), and *STAT4* (p for trend = 3.9×10^{-9}) being significant in both men and women. Four SNPs had gender-specific effects; *PADI-4-94* loci (p for trend = 0.005) affected men only, whereas *TNFA1P2* (p for trend = 0.0064), *MMEL1-TNFRSF14* (p for trend = 3.3×10^{-4}) and rs11761231 loci (p for trend = 1.7×10^{-4}) affected women only.

The odds ratios (OR) were presented in dominant and recessive genetic models. In the dominant model *STAT4* (OR 0.48, 95% CI: 0.38–0.61, p-value = 9.8×10^{-10}) and *HLA-DQA1* (OR 0.35, 95% CI: 0.25–0.48, p-value < 1.0×10^{-7}) were associated with the risk for RA in both genders. *PADI4*-94 (OR 2.16, 95% CI: 1.28–3.66,

p-value = 0.002) and *ANAPC4* (OR 0.59, 95% CI: 0.37–0.94, p-value = 0.02) were associated with RA in men only, whereas *TNFA1P2* (OR 1.36, 95% CI: 1.05–1.77, p-value = 0.016), *MMEL1-TNFRSF14* (OR 0.48, 95% CI: 0.34–0.67, p-value = $9.6x10^{-6}$), *SLC22A4* (OR 1.28, 95% CI: 1.00–1.64, p-value = 0.043) and rs11761231 (OR 0.52, 95% CI: 0.38–0.71, p-value = $2.4x10^{-5}$) were associated with RA in women only.

In the recessive model *HLA-DRB1* (OR 1.8, 95% CI: 1.34–2.44, p-value = $5.7x10^{-5}$), and *HLA-DQA1* (OR 2.17, 95% CI: 1.71-2.72, p-value<$1.0x10^{-7}$) were associated with risk for RA in both men and women. *STAT4* (OR 2.38, 95% CI: 1.28–4.41, p-value = $2.8x10^{-3}$), *ANAPC4* (OR 0.29, 95 % CI: 0.09–0.90, p-value = 0.017), and *SLC22A4* (OR 2.18, 95% CI: 1.04–4.58, p-value = 0.024) were associated with the risk for RA in men only, whereas *CTLA-4* (OR 2.01, 95 % CI: 1.03–3.97, p-value = 0.027), *TNFA1P2* (OR 0.67, 95% CI: 0.45–1.01, p-value = 0.044), *TNFA1P3* (OR 0.68, 95% CI: 0.49–0.94, p-value = 0.015) and *IL2RA* (OR 0.42, 95 % CI: 0.17–0.99, p-value = 0.030) were associated with the risk for RA in woman only.

There were five significant gene-gene interactions, *HLA-DRB1* and *HLA-DQA* (p = 0.075), *HLA-DRB1* and *PADI-4-92* (p = 0.0072), *HLA-DRB1* and *PADI-4-94* (p = 0.0094), *HLA-DRB1*and *IL2RB* (p = $4.22x10^{-5}$), and between *HLA-DRB1* and rs1162922 (p = 0.0072).

Bone erosion in RA patients was associated with a polymorphism in the *HLA-DRB1* loci (p for trend = 0.037) in females. No association was found with *HLA-DQA* loci (p for trend>0.05) and *KAZALDI* loci (p for trend>0.05).

Anti-CCP was found in 85% of the RA patients. It was associated with a polymorphism in the *HLA-DRB* loci (p for trend = 0.043) in males and *HLA-DQA*

loci (p for trend = $1.9x10^{-3}$) in females. No association was found with polymorphisms in *PADI-4* gene ($p > 0.05$) or smoking status ($p > 0.05$).

Serum rheumatoid factor (RF) was found in 65% of the RA patients. RF was associated with a polymorphism in *MMEL1-TNFRSF14* loci in women (p for trend = 0.001), and with *ANAPC4* (p for trend<0.05) and rs1937506 loci (p for trend<0.05) in men.

In conclusion, there was a gender-specific effect of genetic predisposition to RA. The genetic predisposition to RA in Chinese involves population-specific loci and also other susceptibility loci found in Asians and Caucasians.

摘要

類風濕關節炎(RA)是一種自體免疫系統疾病，它的病徵是關節腫脹和發炎。它會永久性地破壞滑膜，軟骨和骨結構。女性患者人數是男性的三倍，在中國人種中的患病率爲 0.35%。

我們在六百 RA 患者和九百名人群中研究了二十四個單核苷酸多態性相關基因。結果顯示，共有七個基因會增加患上 RA 的風險。其中三個均影響响男女患者，包括人類白細胞抗原 *DRB* 一型 (p for trend=1.8×10^{-4})或 *DQA* 一型 (p for trend=2.0×10^{-16}) 和傳感器的信號和轉錄激活四型 (*STAT4*) (p for trend=3.9×10^{-9})。肽酰基精氨酸脫亚氨酶四型 (*PADI-4-94*) (p for trend=0.005)則只影響男性。而腫瘤壞死因子 α 誘導蛋白二型 (*TNFA1P2*) (p for trend = 0.0064)、金屬內肽酶樣一型 - 腫瘤壞死因子受體超家族成員＋四 (*MMEL1-TNFRSF14*) (p for trend=3.3×10^{-4})和 rs11761231 (p for trend=1.7×10^{-4}) 只影響女性。

基因多態性在顯性遺傳模式中增加男女患上 RA 風險的比值比是 *STAT4* (OR=0.48，95%CI: 0.38-0.61，p=$9.8\text{x}10^{-10}$)和 *HLA-DQA1* (OR=0.35，95%CI: 0.25-0.48，p<1.0×10^{-7})。只影響男性的是 *PADI4 – 94* (OR=2.16，95%CI: 1.28-3.66，p=0.002) 和 *ANAPC4* (OR=0.59，95%CI: 0.37-0.94，p=0.02)。而女性則是 *TNFA1P2* (OR=1.36，95%CI: 1.05-1.77，p=0.016)，*MMEL1-TNFRSF14* (OR=0.48，95%CI: 0.34-0.67，p=9.6×10^{-6})， *SLC22A4* (OR=1.28，95%CI: 1.00-1.64，p=0.043)和 rs11761231 (OR=0.52，95%CI: 0.38-0.71，p=2.4×10^{-5})。

而基因多態性在隱性遺傳模式中增加男女患上 RA 的風險的比值比是 *HLA-DRB1* (OR=1.8，95%CI: 1.34 - 2.44，p=5.7×10^{-5})，和 *HLA-DQA1* (OR=2.17，

95%CI: 1.71 - 2.72，$p<1.0\times10^{-7}$)。只影響男性的是 *STAT4* (OR=2.38，95%CI: 1.28-4.41，$p=2.8\times10^{-3}$)，*ANAPC4* (OR=0.29，95%CI: 0.09-0.90，p=0.017) 和 *SLC22A4* (OR=2.18，95%CI: 1.04-4.58，p=0.024)，而女性則是抗原四型 (*CTLA-4*) (OR=2.01，95%CI: 1.03-3.97，p=0.027)，*TNFA1P2* (OR=0.67，95%CI: 0.45-1.01，p=0.044)，*TNFA1P3* (OR=0.68，95%CI: 0.49-0.94，p=0.015)和 *IL-2RA* (OR=0.42，95%CI: 0.17-0.99，p=0.030)。

這次研究發現了五個基因相互之間有影響作用。它們是 *HLA-DRB1* 和 *HLA-DQA* (p=0.075) 、*HLA-DRB1* 和 *PADI-4-92* (p=0.0072)、*HLA-DRB1* 和 *PADI-4-94* (p=0.0074) 、*HLA-DRB1* 和 *IL-2RB*（$p=4.22\times10^{-5}$）、*HLA-DRB1* 和 rs1162922（p=0.0072）。

HLA-DRB1 (p for trend=0.037)的基因多態性與 RA 女患者的骨侵蝕有關。85%的 RA 患者帶有抗-CCP 抗體。*HLA-DRB* (p for trend=0.043) 的基因多態性與男患者有關，而 *HLA-DQA* (p for trend=1.9×10^{-3}) 與女患者有關。抗-CCP 抗體與 *PADI-4* 基因多態性(p>0.05)或吸煙狀況(p>0.05)均沒有相關。65%的 RA 患者帶有類風濕因子。*MMEL1-TNFRSF14* (p for trend=0.001)與 *ANAPC4* (p for trend=0.05)的基因多態性與女患者有關；rs1937506 (p for trend=0.05)與男患者有關。

至此總結，這次研究中，我們發現了基因多態性引發 RA 的主因爲性別和種族。而種族當中亦分爲個別性和普遍性。

To my parents

David and Yuk Lin Li

Acknowledgements

I am immensely grateful to my supervisors, Professor Brian Tomlinson, Professor Edmund Li and Professor Nelson Tang, for their patient guidance, invaluable advice and inspirations. During my part-time study, I am very grateful to their understandings and continuous support. Professor Tomlinson and Professor Li have given me so much advice and not only on research, but also in many aspects of life. I am so thankful to Professor Tang for his trust in allowing me to work in his laboratory.

I must also thank Professor Lai Shan Tam for her support and in patient recruitment.

I would also like to express my sincere gratitude to Professor Jean Woo, Professor Timothy Kwok and Professor Ping Chung Leung for their kind consent to allow me to access the DNA in the health aging men and women study, so that my work could be further examined for validity.

I would like to express my sincere thanks to my colleagues at the Department of Medicine and Therapeutics, in particular Ms Tena Li, and Ms Lorence Chiang. I also want to express my thankfulness to my colleagues at the Department of Chemical Pathology, including Miss Kathy Kong, Miss Amy Wang, Miss Rachel Kwok, Miss Chen Di Liao, Mr. Harris Fan, Mr. Elvis Koon, and Miss Holly Chen, for their assistance and encouragements.

I would like to express my sincere thanks to Prof. Reinhard Renneberg, Dr. Cangel

Chan and Dr. Kirsten Ip of Department of Chemistry, The Hong Kong University of Science and Technology for carrying out the serum Anti-CCP level assay for me.

I must express my warmest thanks to my sons Arthur and Aidan, and to my wife Irie for their understanding and allowing me to pile up a vast collection of books and "papers" at home.

Finally, I dedicated this thesis to my parents for giving me a wonderful life.

Contributors

I declare that I am the main contributor to the works described in this thesis. Under Professor Brian Tomlinson, Professor Edmund Li and Professor Nelson Tang's supervision, I was responsible for the design of the study, sample processing, performing the assays, data analysis and interpretation, and the writing of the manuscripts unless otherwise stated below. Assays for anti-CCP described in chapter 3 were kindly performed by Prof. Reinhard Renneberg, Dr. Cangel Chan and Dr. Kirsten Ip of Department of Chemistry, The Hong Kong University of Science and Technology. The gene-gene interaction analysis described in chapter 3 was kindly performed by Dr. Fred YQ Qiu, Department of Chemical Pathology, The Chinese University of Hong Kong.

Publications

Conference Abstracts

M. Li, N Tang, LS. Tam, E. K. Li, T. K. Li, T. Zhu, E. Suen, R. Kwok, B. Tomlinson (2009). A case-control study to determine the role of ANAPC4 gene (rs3816587) in the Chinese population with rheumatoid arthritis. Annual European Congress of Rheumatology EULAR Copenhagen, Denmark, 10 - 13 June 2009.

M Li, N Tang, E K Li, B Tomlinson, A Wang , T Zhu, T Li, L Tseung, T Kwok and LS Tam (2009). Study of the 5q loci (*SLC22A4-SLC22A5-IL13*) predisposition in Chinese patients with rheumatoid arthritis. 4th Asian Congress on Autoimmunity, Singapore, 11 – 13 September 2009.

M Li, N Tang, B Tomlinson, LS Tam, A Wang, T Li, T Zhu1and E K. Li (2009). Association of SNPs rs11761231 and rs1937506 polymorphism in Chinese patients with rheumatoid arthritis. International Conference on Personalized Medicine, Hong Kong, 19 – 20 September 2009.

M Li, N Tang, B Tomlinson, LS Tam, A Wang, T Li, T Zhu1and E K. Li (2009). Association of IL2 receptor alpha (rs2104286) and beta chains (rs743777) polymorphism in Chinese patients with rheumatoid arthritis. International Conference on Personalized Medicine, 19 – 20 September 2009.

M Li, N Tang, B Tomlinson, LS Tam, E Li (2010). Study of sex-specific association of TNFA1P2 in Chinese rheumatoid arthritis patients. International Conference of Genetic and Genomic Medicine, Taipei, Taiwan, 2 - 5 May 2010.

M. Li, N Tang, B Tomlinson, LS Tam, EK Li (2010) Strong association of MHC-HLA-DRB region in Chinese patients with rheumatoid arthritis. Annual European Congress of Rheumatology EULAR Rome, Italy, 16 – 19 June 2010.

M. Li, N Tang, B Tomlinson, LS Tam, EK Li. (2010) Strong association of STAT4 (rs7574865) and TNFA1P2 (rs2771369) region in Chinese patients with rheumatoid arthritis. Annual European Congress of Rheumatology EULAR Rome, Italy, 16 – 19 June 2010.

M Li, N Tang, EK Li, B Tomlinson, R Renneberg, C Chan, K Ip, A Wang ,T Li, T Kwok LS Tam (2010). Study of the *PADI4* loci predisposition in Chinese patients with rheumatoid arthritis. Hong Kong Society of Rheumatology, Hong Kong. 28 November 2010.

M Li, N Tang, F Qiu, LS Tam, B Tomlinson, A Wang, T Li, T Kwok, EK Li (2011). Gene - gene interactions in Chinese patients with rheumatoid arthritis. Annual European Congress of Rheumatology EULAR, London, England.25 – 28 May 2011.

List of Abbreviations

95% CI	95% confidence interval
ACP	Antigen presenting cells
ANAPC4	Anaphase-promoting complex 4
anti-CCP	Antibodies against cyclic citrullinated peptides
anti-BiP	Immunoglobulin heavy chain binding protein
AS	Ankylosing spondylitis
BMD	Bone mineral density
CCR6	Chemokine (C-C motif) receptor 6
CD	Cluster of differentiation
CEP-1	Citrullinated alpa enolase peptide 1
CEU	Caucasian with European ancestry
CHB	Chinese Han Beijing
CRP	C-reactive protein
CSK	C-terminal src kinase
CTLA-4	Cytotoxic T lymphatic antigen
DAS28	Disease activity score in 28 joints
EMS	Early morning stiffness
EIRA	Epidemological Invegistion of Rheumatoid Arthritis
ESR	Erythrocyte sedimentation rate
GWAS	Genome wide association study
GZEB	Granzymes
HapMap	The International HapMap Project
HLA-DRB1 SE	Class II subunits shared epitope
HnRNP-A2	Heterogenous nuclear ribonucleoprotein A
HRM-PCR	High resolution melting PCR
hsCRP	High sensitivity C-reactive protein
HWE	Hardy-Weinberg equilibrium
IgG	Immunoglobulin G
IL2	Interleukin-2
IL2R	Interleukin-2 receptor

IL6	Interleukin-6
LD	Linkage disequilibrium
LYP	Lymphoid tyrosine phosphatase
JRA	Juvenile RA
KAZALD1	Kazal type serine protease inhibitor domain
Kd	kilodalton
NARAC	North American Rheumatoid Arthritis Consortium
MCTD	Mixed connective tissue disease
MHC-HLA	Major histocompatibility complex, class II, DQ alpha 1
MMEL1-TNFRSF14	Metallo-endopeptidase-like 1 gene - Tumour Necrosis Factor Receptor Superfamily, Member 14
OA	Osteoarthritis
OCTN	Organic cation transporter 1
OR	Odds ratio
PCR	Polymerase chain reaction
pQCT	Peripheral quantitative computed tomography
SLC22A4 -IL13	Solute carrier family 22 – Interleukin 13
T1D	Diabetes mellitus type 1
TRA-alpha	T cell receptor alpha
Tm	Melting temperature
TNF	Tumour necrosis factor
TNFAI3-OLIG3	Tumour necrosis factor, alpha induced protein 3 and oligodendrocyte transcription factor 3
TRAF1-C5	Tumour necrosis factor receptor associated factor 1 and complement component 5
RA	Rheumatoid arthritis
PADI-4	Peptidylarginine deiminase 4
RF	Rheumatoid factor
RFLP- PCR	Restriction fragment length polymorphism
PRKCQ	Protein kinase C, theta
PTPN22	Protein tyrosine phosphatase, non-receptor type 22
SLE	Systemic lupus erythematosis
SNP	Single nucleotide polymorphism

STAT4	The signal transducer and activator of transcription 4
WTCCC	Wellcome Trust Case Control Consortium

List of tables

List of Figures

Table of Contents

Chapter 1. Rheumatoid arthritis: Public health and clinical aspects

1.1. Background

Rheumatoid arthritis (RA) is well known for its adverse destructive effects and irreversible damages on bone and joints (Alwan et al.1988; Harris et al.1975; Herring1984). The goal of RA treatment is to make the diagnosis early on and start therapy as soon as possible so that the disease status can be maintained and further damage can be prevented.

For over a century, the diagnosis of RA was based on the opinion of objective experts. Over the last few decades, details of the disease, its history, its development and its prognosis are becoming clear. With the discovery of RA-associating serum clinical biomarkers, such as erythrocyte sedimentation rate (ESR) (Fischel1956; Kalliomaki1954), high sensitivity C-reactive protein (hsCRP) (Balci et al.1989), rheumatoid factor (RF) (Waaler1939; Zlabinger et al.1990) and antibodies against cyclic citrullinated peptides (anti-CCP), early diagnosis of the disease is now possible. However, these biomarkers are only found in 65 to 85% of all the RA cases (Mackenzie1988; Zlabinger et al.1990).

Diagnostic imaging with X-ray, ultrasonography (Wiell et al.2007), magnetic resonance imaging (MRI) (Ostergaard et al.2004; Ostergaard and Szkudlarek2003) and peripheral quantitative computed tomography (pQCT) are tools that can confirm the diagnosis, but in most cases, a confirmed report means irreversible damage to the joint.

The urge for an efficient, sensitive, specific, and cost effective test for RA is the goal of all rheumatologists alike. Recently, the mode of treatment for RA has been

shifted to the more costly, monoclonal antibodies targeted to specific mediators such as tumour necrosis factor (TNF) (Lipsky et al.2000), or CD20 positive B-cells found in different bodily pathways. These powerful biological agents bring in new hopes, however, moderate to serious side effects have been reported, and sometimes these can be fatal.

Study of the role of heritability in RA started in 1978, the year when RA was first genetically associated with the major histocompatibility-human leukocyte antigen Class II subunits shared epitope (HLA-DRB) (Roitt et al.1978). For the past 30 years, vast volumes of information on the genetics of RA were obtained through surveys, case-control studies, affected twin pair studies and clinical drug studies. They have been replicated in different patient cohorts from various ethnic groups. However, these results are either incomplete or inconclusive.

This year marks the 20^{th} year on of the Human Genome project (Hapmap) and the completion of the second phase of the Human Genome project, from which 3.1 million loci and the haplotypes of 270 individuals from four distinctive global regions are made freely to the general public. Along with this project, new technology has been developed to study multiple diseases on thousands of subjects. In the fall of 2007, two of the landmark RA genome-wide association studies (GWAS), Wellcome Trust Case Control Consortium (WTCCC) (WTCCC2007) and North American Rheumatoid Arthritis Consortium (NARAC) were published. The WTCCC study showed that 49 single nucleotide polymorphisms (SNPs) with p-values in the range of 0.1 to 1×10^{-7} are associated with RA, nine of them being newly found to be associated with RA. Other important SNPs are the protein tyrosine phosphatase, non-receptor type 22 (*PTPN22*) and the major histocompatibility-human leukocyte antigen (*MHC-HLA*) chromosome at 6q23

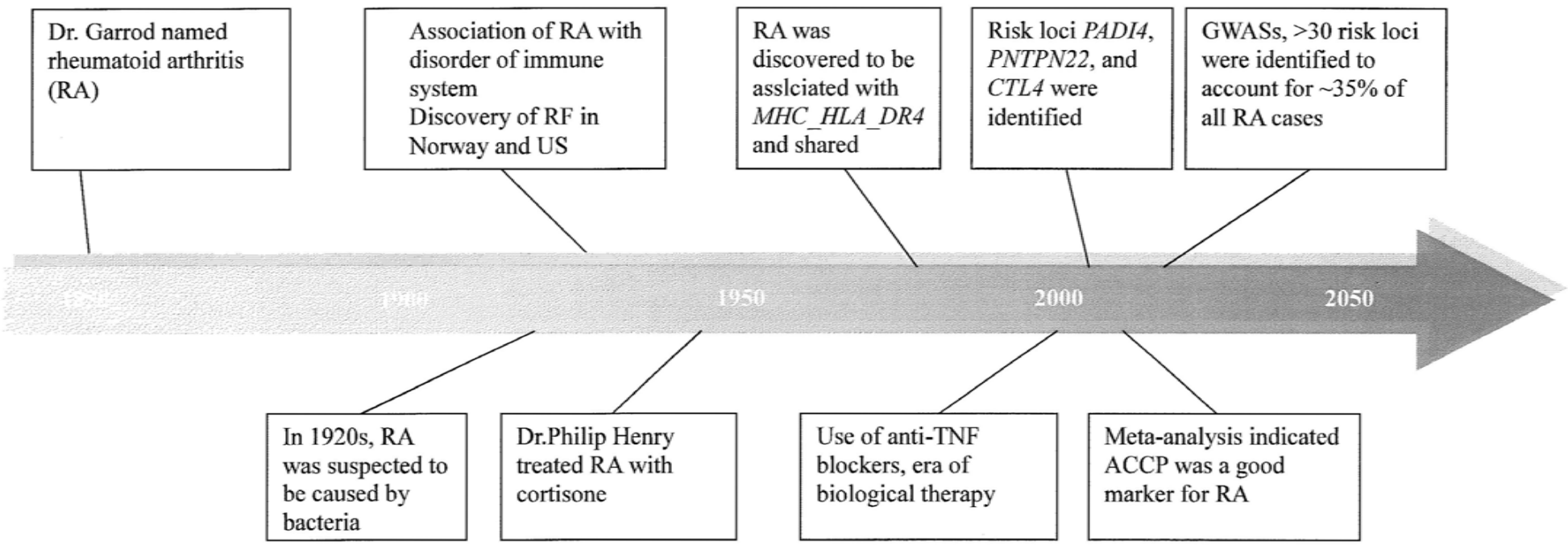
region. The NARAC study indicated that the signal transducer and activator of transcription 4 (*STAT4)* on chromosome 2 is also a risk SNP for RA. This SNP has now been shown as the cause of many autoimmune diseases.

The genetic associations between RA and 25 of these SNPs will be studied in Chinese RA patients. Chinese men and women with RA will be investigated separately. The genotypic effect on biomarkers will also be examined.

1.2. Historical account of rheumatoid arthritis

Rheumatoid arthritis (RA) is an immune-mediated, chronic, inflammatory syndrome. It affects peripheral synovial joints in a symmetrical manner. In 1858, an English clinician Dr. Alfred Baring Garrod named the condition 'rheumatoid arthritis' to distinguish it from gout and rheumatic fever (Figure 1-1).

Figure 1-1. A chronological account of rheumatoid arthritis (Data extracted from Chapters 1 and 3)



1.3. Prevalence and incidence

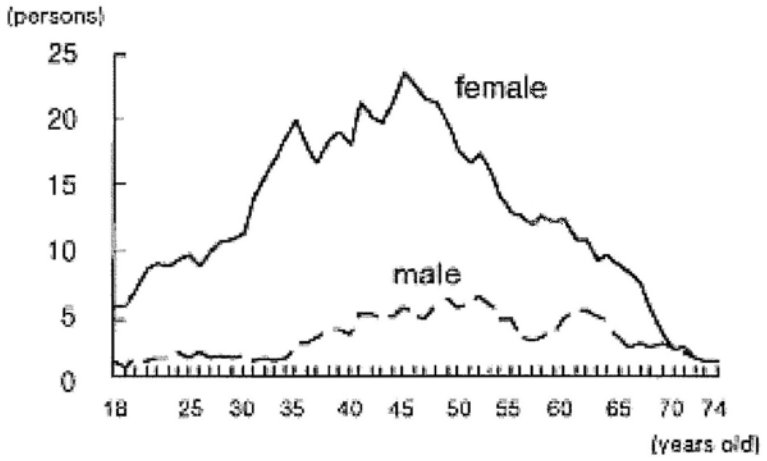
The prevalence of RA worldwide stands at 0.2-1.0% in Caucasians aged 18 or over. It increases to 2% and 5% in men and women over age 65 years. The incidence of RA peaks at 40-60 years. The prevalence of RA is 0.2-0.93% in China (Zeng et al.2008), 0.75% in India (Malaviya et al.1993), 0.2% Japan (Shichikawa et al.1999), 0.8% in England (Macgregor et al.1994), 0.0026 % in South Africa tribe and 5-8% in native South American tribes (Del Puente et al.1989; Harvey et al.1981). In Hong Kong Chinese it is reported as 0.35% (Lau E1993). RA is absent from some African populations (Brighton et al.1988) (Table 1-1.). However, in British black-Caribbeans it is 0.29%, which may suggest that RA is affected both by genetic and environmental factors.

Table 1-1 Prevalence of RA in some selected countries

Population	Prevalence (%)	Reference
Indians, US	5.3-6.0	(Harvey, Lotze et al. 1981)
China	0.2-0.93	(Zeng et al.2008)
England	0.80	(Macgregor et al.1994)
India	0.75	(Malaviya et al.1993)
Greece	0.58	(Anagnostopoulos et al.2010)
Sweden	0.51	(Simonsson et al.1999)
France	0.50	(Guillemin et al.2005)
Spain	0.50	(Carmona et al.2002)
Hong Kong	0.35	(Lau et al.1993)
Japan	0.2	(Shichikawa et al.1999)
South African tribe	0.0026	(Brighton et al.1988)
UK black-Caribbeans	0.29	(Macgregor et al.1994)

There is a gender difference of 1:3-4 ratio between men and women for developing RA (Carmona et al.2002; Papadopoulos et al.2003; Pritchard1992), while the onset age is 35 to 45 years in women and 40-50 years in men (Yukioka et al.1998) (Figure 1-2). However, juvenile RA (JRA) may occur in adolescent girls (Burgos-Vargas and Vazquez-Mellado1995).

Figure 1-2 The distribution of age at disease onset for men and women with rheumatoid arthritis in Japan (Yukioka et al.1998)



The prevalence of RA is higher in developed countries than developing countries. The prevalence may be underestimated in the developing countries (Kalla and Tikly2003), because the inhabitants do not live long enough to develop the disease (Adebajo1990). Global studies from Denmark and the US suggested that the incidence rate is around 24-35 cases per 100,000 person year (Dugowson et al.1991; Pedersen et al.2009). Retrospective studies from Japan, and UK showed that reduced incidence rate was recorded in women with the use of contraceptives pills during the 60s and 80s (Brennan et al.1997; Shichikawa et al.1999; Symmons2002). In one study, an increase in age of onset of RA was established and this implied environmental factors may play a role in the onset of the disease (Imanaka et al.1997). In a 10 year longitudinal study carried out from 1994 to 2004 in Norway, the investigators showed that the health, the pain and the global status of the patients improved during the study period, and the authors concluded that it was attributed to better medication (Uhlig et al.2008)

RA affects morbidity and mortality, reducing life expectancy by an average of 7 years in men and 3 years in women (Vandenbroucke et al.1984). Other studies established RA increased mortality (Heliovaara et al.1995) and reduced life expectancy by 15-20% from the date of RA onset (Myllykangas-Luosujarvi et al.1995). There were three major causes of death in RA, around 30-40% of the deaths were due to cardiovascular diseases, 30% were due to infections and 15% were due to renal disease (Mutru et al.1989; Myllykangas-Luosujarvi et al.1995). Other risk factors contributing to death including smoking (Oliver and Silman2006), aging, being male, greater functional impairment, test for RF being positive, numbers of swollen joint, education level (Pincus, T and Callahan1985), and co-morbidity.

1.4. Diagnosis of rheumatoid arthitis at early stage

The onset of RA usually takes weeks to several months, but 15–20 % of the patients showed a rapid onset over days to week, and 8-15 % showed acute onset of symptoms over days.

It usually starts off with symmetrical small joint pain and swelling in the hands, feet and wrist bones. As the disease progresses, large size joints such as elbows, knees and hips become involved. In established RA, it is characterized by a deforming symmetrical polyarthritis of varying severity. It may involve the synovium of joints and tendon sheaths, articular cartilage loss, erosion of the bone and the presence of IgM RF in the blood of 70% of the patients. These changes are irreversible. Therefore, the aim is to diagnosis the disease as early as possible.

The clinical challenge to early diagnosis of RA is that there is no definitive laboratory test or confirmatory physical finding in early disease. At the early stage of the disease. The arthritis symptoms may be similar to systemic lupus erythematosis

(SLE), psoriasis, ankylosing spondylitis (AS), osteoarthritis (OA), gout and chronic infection. In order to distinguish rheumatoid arthritis from other form of arthritis, the location and the form of arthritis are used. Arthritis may be grouped into three types (Table1-2).

Table 1-2 Types of arthritis (Li, EK2002)

Type of arthritis	**Joints (n)**	**Disease example**
Acute monoarthritis	1	Infection, gout, pseudogout
Chronic monoarthitis	1	Psoriasis, RA, AS, OA
Chronic asymmetrical oligoarthritis	≤ 4	Reactive arthritis, psoriasis, AS
Chronic symmetrical polyarthritis	≥ 5	RA, SLE, psoriatic arthritis

The American College of Rheumatology has developed a seven criteria list for the classification (Table 1-3) (Arnett et al.1988). In order to define RA, four or more of the criteria are needed at presentation. The sensitivity and specificity of the list are about 40-60% and 90% in early and active RA, respectively (Table 1-4). The list consists of five clinical, one serological and one radiological feature. The reason for reduced sensitivity and specificity in early stage RA is because formation of rheumatoid nodules is rare in the early stage. In the early stage, RA often presents as a mono or oligoarticular disease. RF is frequently absent in the early stage. Bone erosion is rarely found in the early stage, but joint erosions occur in 50-70% of RA patients.

Table 1-3 Criteria for the classification of rheumatoid arthritis (Arnett et al.1988)

Criteria	Description
1	> 1 hour morning stiffness in and around joints
2	3 or more joint soft tissue swellings observed by a physician
3	Swelling (arthritis) of proximal interphalangeal, metacarpophalangeal or wrist joints
4	Symmetrical swelling of joints
5	Subcutaneous rheumatoid nodules
6	Presence of Ig M RF in abnormal amounts
7	Radiographic erosions and / or peri-articular osteopenia in hand and / or wrist joints

Table 1-4 Criteria for active rheumatoid arthritis

Criteria	Description
1	More than 6 tender and 3 or more swollen joints
2	ESR> 30 min / hour
3	Early morning stiffness (EMS), more than 45 minutes
4	DAS more than 5.2

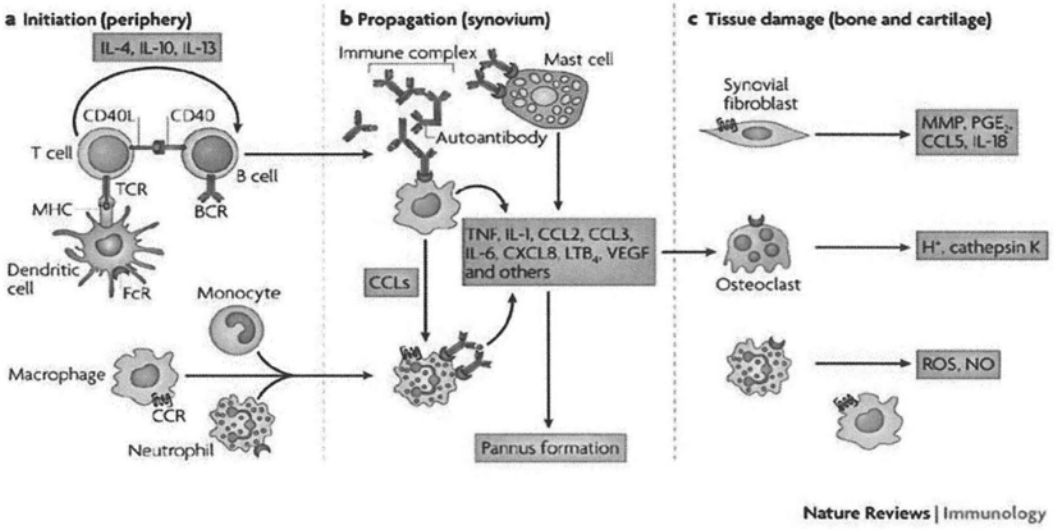
Because of these limitations, the diagnosis of early RA before irreversible joint damage occurs and the use of appropriate medications in the early stage of the disease may be difficult. Thus, additional tools may be needed, including antibodies to citrullinated peptides, medical imaging, genetics and the use of biological agents in treatment.

1.5. The etiology of RA

The etiology of RA is not fully understood. It is currently believed to consist of two components, environment and genetics. The progression of RA may be classified

into three phases: the initiation phase; the propagation phase; and the tissue damage phase (Figure 1-3).

Figure 1-3 Schematic diagram to show the progression of rheumatoid arthritis (Rommel et al.2007)



Rommel *et al. Nature Reviews Immunology* **7**, 191–201 (March 2007) | doi:10.1038/nri2036

The initial phase is noted by the presence of autoantibodies to the Fc portion of immunoglobulin (IgG) (Corper et al.1997) prior to the onset of the disease, it is also called RF. The initial phase and the propagation phase are associated with both environmental and genetic factors. The destructive phase is accomplished with cytokines and osteoclast activity.

Extra-articular manifestations are also common in RA patients, which are related to vasculitis such as rheumatoid nodules, episclerlitis, peripheral neuropathy and palpable purpura.

1.6. Auto-antibodies in autoimmune diseases

Auto-antibodies are usually found in patients with autoimmune diseases. They usually have no pathological role, but are good surrogate markers for diagnosis and prognosis of the disease. Some of these specific markers are listed in Table 1-5 along with their corresponding diseases.(Maddison and Huey2004)

Table 1-5 Specific markers in autoimmune diseases

Disease	Marker
SLE	Anti-double-stranded (ds) DNA
Scleroderma	Anti-topoisomerase I
Poly-dermatomyositis	Anti-histidyl-tRNA synthetase
RA	Anti-CCP

1.7. Auto-antibodies in RA

Auto-antibodies in RA may be put into three groups according to their target antigens, RF, anti-CCP, and nuclear proteins (Table 1-6).

Table 1-6 Biological markers used in the diagnosis of RA

Antibody	Antigen	Sensitivity (%)	Specificity (%)	Reference.
RF (IgM)	IgG	59-79	80-84	(Ates et al.2007)
anti-CCP family				
Anti-CCP	Citrullin	70-78	88-96	(Lee and Schur 2003)
Anti-perinuclear	Citrullin	49-91	73-99	(Nienhuis and Mandema1964)
Anti-keratin	Citrullin	36-59	88-89	(Niewold et al.2007)
Nuclear proteins				
Anti-RA33	HnRNP-A2	30	90	(Hassfeld W et al.1989)
Anti-p68	Stress protein BiP	63-71	73-96	(Purcell Aw2003)
RF + anti-CCP		30-39	98-100	(Schellekens et al.2000; Vittecoq et al.2004)

Table partly adapted from:(Steiner. G and Js2006)

1.7.1. Rheumatoid factor

RF is identified 70 years ago (Waaler1939). It is found in 60-80% of RA patients and it is also found in the sera with other auto-immune diseases such as SLE (Witte et al.2000), scleroderma (Barland and Lipstein1996), polymyositis (Zhao, Jinxia et al.2008), and Sjogren syndrome (Civilibal et al.2007; Ramos-Casals et al.2010). Reactive arthritis (Laasila et al.2003) and osteoarthritis patients (Fonseca et al.2000) and healthy controls also showed low RF titers (Mori, S et al.2009).

RF appears late in the progress of RA, whereas at the early state of RA, serum RF levels are lower and fluctuate. Test sensitivity depends on laboratory cut-offs, since it is present in 15% of healthy controls.

RFs that are found in the synovial space are predominantly IgM isotypes, but IgG and Ig A are also present. In both humans and mice, the production of those immunoglobulins is in plasma B-cells and the process is antigen–driven. Rat models show that the production of RFs correlates with arthritis prognosis (Lawrence1967; Wernhoff2003). It is higher in female than in male rats. RF predominantly recognizes epitopes located in the Fc region of the IgG molecule, especially the CH2 and CH3 chains (Corper et al.1997; Hay1988). It is postulated that RFs and IgG molecules formed a stable RF-IgG complex during antigen capture, presentation of antigens and also for IgG clearance.

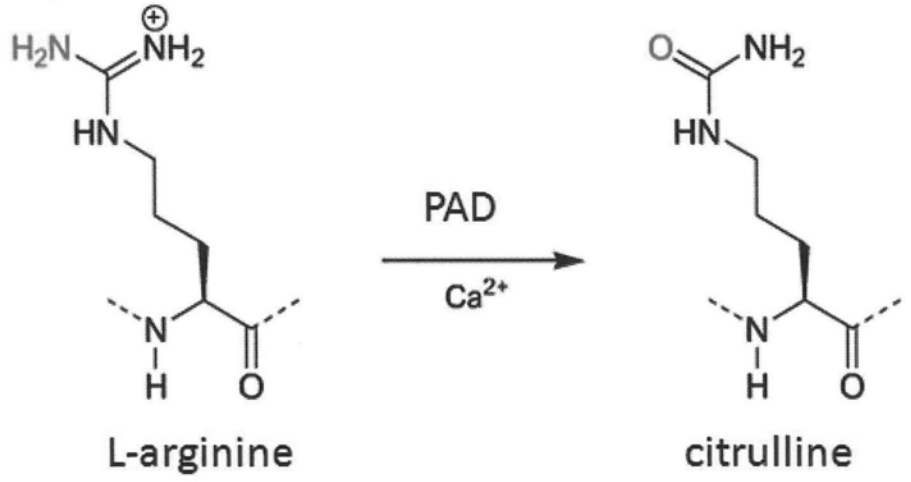
A meta-analysis shows an association between RF status and a polymorphism in the tumour necrosis factor receptor associated factor 1 and complement component 5 (*TRAF1-C5*) (Patsopoulos and Ioannidis2010). RF status in RA patients is associated with polymorphisms in rs3761847 (OR = 1.32, 95 % CI: 1.21-1.45, p-value= 8.5×10^{-11}) and rs6920220 (OR=1.27, 95% CI: 1.21 -1 .34, p-value=3.3×10^{-10}) in the 6q23 loci. The presence of RF increases the risk for RA

(Patsopoulos and Ioannidis2010).

1.7.2. Anti-citrulline-binding autoantibodies

The discoveries of anti-CCP spans over a period of 40 years. This group of auto-antibodies consisted of three family members, the anti-perinuclear (Nienhuis and Mandema1964), anti-keratin (Niewold et al.2007) and anti citrulline-binding auto-antibodies. They share high specificity for RA and are shown to be good indicators for the detection of early RA. Overall anti-CCP shows the highest values in terms of both sensitivity and specificity (Niewold et al.2007). However, the serum levels of the former two markers fluctuate greatly. It is only in 1993, with the use of antibody to citrulline, that it is realized that all three auto-antibodies are derived from the anti-CCP pathway in which arginine is transformed to citrulline (Simon et al.1993), anti-perinuclear and anti-keratin antibodies are both intermediate compounds. Currently, the more sensitive and specific anti-CCP-2 kits are used to determine serum anti-CCP levels, as they are based on selected highly-expressed citrullated epitopes.

Figure 1.4 Conversion of arginine to citrulline by peptidyl-arginine deminase (PAD)



Anti-CCP antibodies can be detected years in advance of the development of RA. In the historical cohort study, the Nurses' Health Study (NHS), it was suggested that anti-CCP antibodies can be detected from sera 5.6 ± 3.5 years (range: 0.3-2 years) in advance of RA onset (Chibnik et al.2009). Anti-CCP antibodies can also be used to follow the prognosis of the disease activity. In a three years prospective cohort study of RA patients, van Gaaler et al. established that in 93% of anti-CCP antibodies positive undifferentiated arthritis patients they developed RA within 1-3 years (Van Gaalen et al.2004). The positive association between serum anti-CCP antibody level and RA status was found in different populations including Caucasians (Bongi et al.2004; Chibnik et al.2009), Asian (Liao et al.2011; Zhao, J. et al.2010) and African (Hodkinson et al.2010). Moreover, disease activity is positively correlated with serum anti-CCP antibody levels (Chibnik et al.2009). In the three year Swedish early intervention to rheumatoid arthritis study (TIRA), Kastbom et al. reported that serum anti-CCP antibodies level reduced with positive responses to anti-rheumatic treatment , but this does not alter the patient's anti-CCP antibody positive status (Kastbom et al.2004).

Table 1-7 Parameters shown to be correlated with positive anti-CCP status in rheumatoid arthritis patients

Parameter	Correlation direction (+ /-)
ESR	+
CRP	+
Swollen joint count	+
Worse physician global	+
RF	+
Ever smoker	+
DAS	+
MHC-HLA-DRB1	+

Several genetic association studies suggested that serum anti-CCP antibody

level is associated with MHC-Class II alleles (Ding et al.2009; Lee, HS et al.2008). In a blood bank sera study, the presence of both anti-CCP antibody and the *HLA-DRB1SE* allele predicts the RA onset (OR = 66.8) (Berglin et al.2004). Other studies indicated that the anti-CCP level was associated with *PTPN22* haplotype (Johansson et al.2006). In a Japanese RA cohort, the haplotype of the peptidylarginine deiminase 4, the enzyme that converts arginine to citrulline, was associated with RA risk (Barton, A. et al.2004; Suzuki et al.2003), however, this result could not be replicated in Caucasian studies (Barton, A. et al.2005; Burr et al.2010). Recently, it was suggested that anti-CCP positive and anti-CCP negative RA were two genetically distinct diseases (Ding et al.2009; Ohmura et al.2010).

A meta-analysis showed that anti-CCP status associated with polymorphisms in the tumour necrosis factor receptor associated factor 1 and complement component 5 (*TRAF1-C5*) (Patsopoulos and Ioannidis2010). Anti-CCP status in RA patients was associated with the polymorphisms rs3761847 (OR = 1.25, 95% CI: 1.13 - 1.37, p-value = 6.8×10^{-6}) and rs6920220 (OR = 1.29, 95% CI: 1.21 - 1.39, p-value = 1.6×10^{-13}) in 6q23 loci.

1.7.3. Anti-nuclear proteins

Anti-A2/RA33 auto-antibodies (Hassfeld W et al.1989) target a nuclear protein, heterogeneous nuclear ribonucleoprotein A (hnRNP-A2) (Hassfeld W et al.1995), which is an enzyme that splices and transports mRNA within the endoplasmic reticulum. Auto-antibody to hnRNP-A2 is found in 30% of RA (Trembleau et al.2010), 20 - 30% of SLE (Fritsch-Stork et al.2006; Sun et al.2003) and 40% of the patients with mixed connective tissue disease (MCTD) (Skriner et al.1994). It is also found in early RA patients but it is absent in osteoarthritis, reactive arthritis or psoriatic arthropathy tissues. Its specificity for RA increased to 96% when it is used

with a negative selection panel for SLE and MCTD (Steiner G et al.1996). It is over-expressed in the synovial membrane of RA patients and may form the target of B and T cell auto-reactivity (Panayi and Corrigall2006). In a mouse model, the auto-antibody caused severe erosive arthritis similar to RA in humans (Tighe et al.1990).

Anti-BiP (immunoglobulin heavy chain binding protein) is also known as gp78. It is a member of the heat shock proteins (Purcell Aw2003). Found in 60% of RA sera (Bodman-Smith et al.2004) and also found in other rheumatic diseases, it is highly expressed in synovial tissue similar to hnRNP-A2,. Its function and role in RA pathology are not fully elucidated. Contradictive reports have been published that it may modulate T–cell reactivity and it may also co-mediate auto-immune responses.

1.8. Classical laboratory investigations

1.8.1. C-reactive protein

C-reactive protein (CRP) is an inflammation marker, which is synthesized in the liver in response to pro-inflammatory cytokines, such as interleukin 6 (IL-6). Elevated levels of CRP are found in patients within the first 24 to 74 hours following tissue damage, infection and systemic inflammation. CRP activates a numbers of cell types in the immune pathways, such as neutrophils, monocytes and activates platelet proliferation. The CRP gene is located on chromosome 1q21. CRP is a pentamer with five 23 kd sub units. The monomer binds to IgG immune complexes and it also binds to the Fc receptors of granulocytes which in turn facilitate phagocytosis. Moreover, CRP may have a role in gene regulation, as it has been shown to bind to chromatin, histones and small nuclear ribonucleoprotein (snRNPs). The role of CRP in auto-immune diseases is not fully understood. Significant levels of CRP have been

associated with myocardial infarction, stroke, vascular disease and RA. It has been suggested that inflammation is critical in the pathogenesis of these diseases.

1.8.2. The erythrocyte sedimentation rate

This is a one-hundred year old non-specific test for autoimmune diseases, in which red cells aggregate together to from champs and increase the sedimentation rate. In RA patients it can reach over 100 mm / hour.

In summary, the above biomarkers provided tools for RA detection. In established RA, RF is a good indicator as it is readily measured in all hospitals. Anti-CCP is a good marker for early RA detection, but it is not freely available in all clinical laboratories due to its cost. Auto-antibody to hnRNP-A2 is the most specific of all markers, but it is most costly, due to a set of negative selection control marker panels for SLE and MCTD being also needed for each test. CRP and ESR provided additional information. Diagnostically, anti-BiP auto-antibodies have no significant value, but this may be an area for drug development.

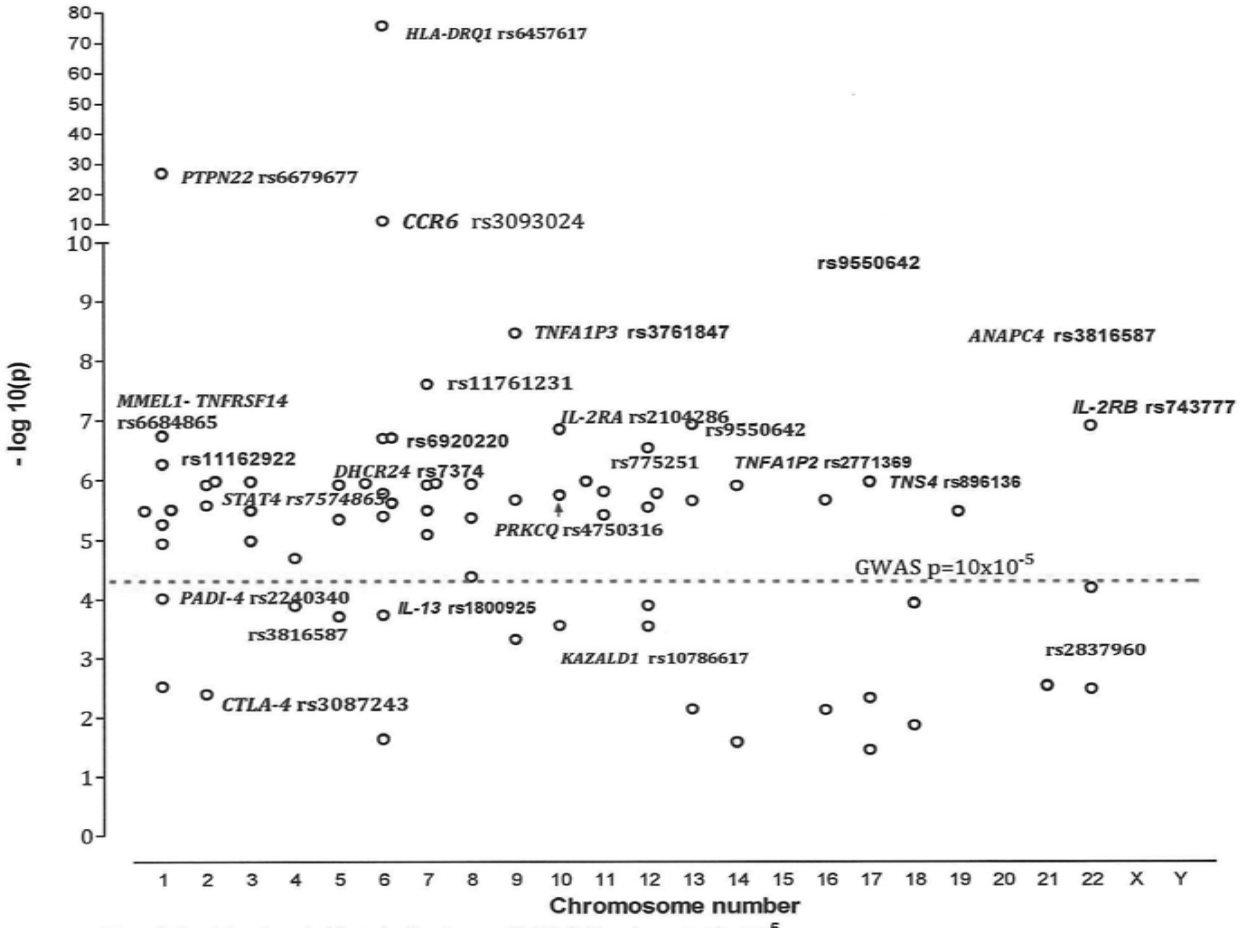
Chapter 2 Genetic susceptibility to rheumatoid arthritis

2.1. Overview

Prior to genome wide association studies (GWAS), the study of genetic susceptibility of RA was mainly examined by the candidate gene approach. Results from early single SNP studies and twin studies suggested that the major histocompatability complex human leukocyte antigen (*MHC-HLA*) alleles accounted for 30% of all known genetic contribution (Deighton et al.1989; Rotter and Landaw1984) and the non-MHC loci made up to the rest of the genetic contribution. These non-MHC loci included *PTPN22, STAT4, PADI4, CTLA-4, TNFAI3-OLIG3, TRAF, IL2-alpa, IL2-beta, OCTN-IL13* loci (Chen, R et al.2010; Li, Y et al.2008; Patsopoulos and Ioannidis2010; Remmers et al.2007; WTCCC2007).

In 2007, two GWAS studies reported on RA: the WTCCC (WTCCC2007) and the North American Rheumatoid Arthritis Consortium NARAC study (Remmers et al.2007). These are the very first GWAS studies to address the RA predisposition SNPs. Over 330,000 SNPs were genotyped, and around 49 were found to be associated with RA, ten of them being new to RA. The NARAC study used a different Caucasian cohort and established that protein tyrosine phosphatase (*PTPN22*) is involved in RA predisposition (Remmers et al.2007).

Figure 2-1 Established candidate SNPs associated with risk of rheumatoid arthritis



The dished horizontal line indicates a GWAS P value of $10x10^{-5}$

2.2. Major histocompatibility Complex susceptibility alleles

The MHC is also called the HLA locus, which is located on chromosome 6p21.3 and is well known as the primary locus consistently and repeatedly associated with RA in many population cohorts (Taneja et al.1996). In 1978, Statney established that the DW4 of the class II HLA proteins were more commonly found in patients with RA when compared with controls (Colbaugh and Stastny1978; Stastny, P.1978; Stastny, P.1978). Gregersen et al. located the DW4 subunits to a shared conservative amino acid sequence R-A-A (arginine, alanine and alanine) at positions 72-74 in the third hypervariable region (HVR3) of the HLA-DRB1 protein chain (Gregersen et al.1986; Gregersen et al.1987). The HLA-DRB1 SE is associated with RA (Gregersen et al.1987; Meyer et al.1999), its severity(Meyer et al.1999), bone erosion (Gorman et al.2004) and anti-CCP antibodies (Kapitany et al.2008; Linn-Rasker et al.2006). In 2005, du Montcel et al. (Du Montcel et al.2005) established the RAA motif is modulated by amino acid at positions 70 and 71. At amino acid position 70, amino acid residues in descending order of risk for RA are glutamine (Q), arginine (R), and aspartic acid (D), while at position 71 are lysine (K), arginine (R), alanine (A) and glutamic acid (E). The alleles are classified as *S2*, *S3P* and *L* (*S1*, *S3D* and *X*) according to the risk for RA and the type of amino acid at positions 70 and 71. The S and X groups indicated the presence and the absence, respectively, of the RAA sequence at positions 72-74. The L group indicated a low risk for RA (Table 2.1)

The class II subunits of the *HLA-DRB1* locus coded for a shared epitope (*HLA-DRB1 SE*) region which includes *DRB1*01 DRB1*04* and *DRB1*05* loci (Gregersen et al.1987; Kim, HY et al.1995). A meta-analysis showed an association

of bone erosion with the *DRB1*04* loci (Gorman et al.2004). In Caucasians, severe RA is associated with *HLA-DRB1*04,* but this locus is rare in Asian, Kim et al. established that the *HLA-DRB1*05* is the most common form of the *HLA-DRB1* locus in Koreans (Kim, HY et al.1995).

The *HLA-DRB1*SE allele is associated with being anti-CCP positive in 70% of all RA patients (OR = 3.0). The presence of anti-CCP antibodies is associated with the severity and the onset of RA. This may imply that anti-CCP antibody positive RA patients and anti-CCP antibody negative RA patients have different pathologies.

Table 2-1 Classification of *HLA-DRB1* alleles (Gyetvai et al.2010; Michou et al.2006)

Allele (%)	Positions 70-74[a]	*HLA-DRB1*
S2 (9%)	Q-K-RAA	*0401
	D-K-RAA	*1303
S3P (21%)	Q-R-RAA	*0101,*0102,*0404,*0405, *0408
	R-R-RAA	*1001
S1 (24%)	D-E-RAA	*0103, *0402, *1102, *1103, *1301, *1304, *1323
	Q-Q-RAA	*15
S3D (10%)	D-R-RAA	*1101, *1104,*12, *1305, *1306, *1325, *1422, *16
X (36%)	Q-K-RGR	*03
	Q-R-RAE	*0403, *0407, *0411
	D-R-RGQ	*07
	D-R-RAL	*08
	R-R-RAE	*0901, *1401, *1404

[a]Animo acid codes: Alanine (A), aspartate (D), glutamate (E), glycine (G), lysine (K), leucine (L), arginine (R),

2.2.1. Major histocompatibility complex, class II, DQ alpha 1 (*HLA-DQA1)*

The *HLA-DQA1* class II gene encodes for the T-cell receptor (TRAb - V). It was reported that it enhances the risk of arthritis in JRA in Caucasian RA patients from north Europe (Haas et al.1995; Maksymowych et al.1992). A 43kd nuclear binding protein DEK (Murray et al.1997) and its autoantibody - anti-DEK were frequently found in 57% of the patients with JRA compared to 3% of controls ($p < 0.0001$) (Adams et al.2003; Murray et al.1997). DEK is a DNA binding protein, whose target is the highly conserved TC-rich Y-box regulatory element (5'GATTG3') in MHC-HLA class II gene promoters. The Y box penta-oligo motif is highly conserved in DQA1, DRA, DQB and DRB genes.

This DEK is an oncoprotein as well as an auto-antigen over-expressed in inflammatory cells in synovial fluid with chemo-attractant activity (Mor-Vaknin et al.2006). These observations suggested that DEK may play a role in the regulation of *HLA-DQA* antigens. In 2007, the WTCCC GWAS study reported that SNP rs6457617 (WTCCC2007) is highly significant in RA patients ($p = 10 \times 10^{-76}$), and in HapMap data, it is mapped to the chromosome 6 region *HLA DQ1* region.

2.3. Non *MHC-HLA* SNPs

2.3.1. Protein tyrosine phosphatase, non-receptor type 22

Protein tyrosine phosphatase non-receptor type 22 (*PTPN22*) is associated with RA pathogenesis in Caucasians (Begovich et al.2004; Lee, AT et al.2005; Plenge et al.2005). It is located on chromosome 1(1p13), and carriers of the T-allele in (rs2476601, 1858C>T) increases the relative risk of RA by 1.6 fold (OR:1.6, 95% CI: 1.4-1.8, $p=4.8 \times 10^{-12}$) (Plenge et al.2005). The T allele variant results in an amino acid change from arginine to tryptophan (R620W) (Rodriguez-Rodriguez et al.2011)

in one of the several proline rich motifs at the non catalytic C-terminal end. This motif binds with the negative regulatory kinase C- terminal src kinase (CSK). This LYP-CSK complex inhibits the T cell antigen receptor (TCR) signaling pathway by down-regulating the tyrosine kinase, lymphocyte specific protein tyrosine kinase (LSK) (Tanaka et al.2005) and reduces the T cell activation (Tanaka et al.2005).

The risk allele T in rs2476601 is found in all population groups studied from UK (Vinkemeier et al.1998), Swedish (Plenge et al.2005), Dutch (Zhernakova et al.2005) or North American descendents (Begovich et al.2004; Kokkonen et al.2007; Potter et al.2009). From the human genome project, Caucasian with European ancestry (CEU) data showed that rs2476601 is tagged to rs6679677. However, in both Asian and African populations, no association is found between *PTPN22* and RA. Hapmap data showed that rs6679677 is not polymorphic in Chinese and the MAF is 2% for rs2476601

The role of *PTPN22* in autoimmune diseases is seen in SLE (Carlton et al.2005), juvenile idiopathic arthritis (JIA) (Ota et al.2004; Vinkemeier et al.1998), Grave's disease (Criswell et al.2005) and vitiligo (Arora et al.2003; Chen, X et al.2003; Yoshimoto et al.2003). However, no association is found with multiple sclerosis (Mathur et al.2007; Vinkemeier et al.1998), Crohn's disease (Watford et al.2004), psoriasis (Tyler et al.2007) or psoriatic arthritis (Vinkemeier et al.1998).

PTPN22 is the most replicable non *MHC-HLA* SNP often found to be associated with RA in Caucasians (Costenbader et al.2008; Goeb et al.2008; Rodriguez-Rodriguez et al.2011), and along with *MHC-HLA-DRB*, the two genes explain 50% of all RA cases (Tyler et al.2007; Ungureanu and Silvennoinen2005). Similar to *MHC-HLA-DRB1,* the *PTPN22* is strongly associated with being anti-CCP positive (OR = 1.49, $p = 2 \times 10^{-5}$ but not anti-CCP negative) (Plenge et al.2005).

These findings further support that anti-CCP positive and anti-CCP negative RA have different pathologies. In addition, T cells that express the T-allele produce less interleukin-2 (*IL2*) in response to T cell signaling. Therefore, *PTPN22* C>T induces weaker signaling to effectively remove auto-reactive T cells.

Similar to the *HLA SLE Se* allele, the SNP *PTPN22* is associated with a 4.2 years earlier onset of RA than those without (Karlson et al.2008). In Caucasians, no association was found between the *PTPN22* genotype and bone erosion (Karlson et al.2008).

2.3.2. Signal Transducer and Activator of Transcription 4

The signal transducer and activator of transcription 4 (*STAT4*) is an 86kb protein, with gene located on chromosome 2 (2q32.2-q32.3) (Remmers et al.2007). A SNP was indentified at position 274 (rs7574865 G>A) within the third intron of the *STAT4* gene . STAT4 is a member of the STAT family of transcription factors (Remmers et al.2007). It is induced by cytokines, such as interleukin12 (*IL-12*) (Nishikomori et al.2002; Thierfelder et al.1996), *IL-23* (Watford et al.2004) and type 1 interferons (*IFNs*). They are phosphorylated by the receptor-associated kinase in the cytosol and form homo or heterodimers that are translocated to the cell nucleus, where they induce interferon-gamma (*IFN-gamma*) transcription and promote type 1 helper (Th1) differentiation (Watford et al.2004). STAT4 also plays a role in Th17 cell differentiation through *IL-23* activation (Mathur et al.2007). STAT4 is expressed in natural killer (NK) cells (Huang, Y et al.2011; Miyagi et al.2007), T cells (Mathur et al.2007; Thierfelder et al.1996), and follicular lymphomas (Zhang, SS et al.2001).

The T variant (rs7574865) of *STAT4* is in linkage with three other SNPs (rs11889341, rs8179673 and rs10181656) and rs7574865 shows the strongest

association with RA (OR=1.37, 95% CI: 1.17, 1.6, p=8.29 x 10^{-5}) (Lee, HS et al.2007). The association of the rs7574865 T-allele with RA was reported in patient groups of both Caucasians (Barton, Anne et al.2008) and Asians (Lee, HS et al.2007). The prevalence of the T allele in rs7574865 is 26% in Caucasians (Remmers et al.2007) and 32% in Asians (Lee, HS et al.2007).

The association of rs7574865 with RA and SLE was first reported in 2007 by Remmers et al. in the NARAC case-control series study (Remmers et al.2007). It was then replicated in a Korean population (Lee, HS et al.2007). Moreover, *STAT4* has been established to be associated with SLE in Caucasian (Namjou et al.2009), Korean, Japanese (Tsuchiya et al.2010) and Chinese populations (Li, P et al.2010; Luan et al.2011).

2.3.3. Cytotoxic T Lymphatic Antigen 4

Cytotoxic T lymphatic antigen 4 (CTLA-4) is also known as CD152 (cluster of differentiation 152). The *CTLA-4* gene on chromosome 2q33 encodes immunoglobulin CTLA-4, which is expressed on the surface of helper T cells and regulatory T cells (Tregs). The role of CTLA-4 in the T cell activation pathway is inhibitory, as it down-regulates T cell activation. T cell activation has three stages. The first stage involves the recognition of an antigen-presenting cell such as B cells, macrophage or dendritic cells bound to an antigen that is complementary to its HLA epiotope. The CD80/CD86 (also as known as B7) receptors on the APC surface then bind to CTLA-4 and its co-stimulator CD28 on the T cell surface. The function of CD28 is to up-regulate T cell activation. This cell-to-cell recognition mechanism is further strengthened by the application of soluble fusion protein abatacept which consists of a *CTLA-4* linked to modified Fc protein of human immuoglobulin G 1,

(IgG1). This fusion protein reduces T cell activation by binding to CD80/CD86 and blocking the interaction with CD28. Abatacept has been shown to be effective in the management of RA.

In 2007, two large Caucasian cohorts, the North American Rheumatoid Arthritis Consortium (NARAC) and the Swedish Epidemiological Investigation of Rheumatoid Arthritis (EIRA) showed that rs3087243 of the *CTLA-4* gene is associated with RA development (OR=1.1, 95 % CI: 1.0-1.2, p=0.004) (Plenge et al.2005). Similar to *HLA-DRB1*and *PTPN22*, the association is limited to anti-CCP positive cases.

2.3.4. Peptidylarginine Deiminase 4

The peptidylarginine deiminase 4 (*PADI-4*) locus is located on chromosome 1 (1p36) and is associated with RA in Asian populations. Meta-analysis showed a SNP rs2240340 at intron 3 (A>T) is associated with Asian RA patients (OR = 1.3, 95% CI: 1.2-1.4, p=0.0001). Inconsistent results for the association between *PADI-4* and RA have been found in different Caucasian study populations. Positive associations were found in North American and German populations but it could not be replicated among Spanish, Swedish and UK populations (Burr et al.2010). In Asian populations both Korean and Japanese studies showed *PADI-4* is associated with RA. However, recently, Chen et al. showed 5 SNPs (rs11203366, rs11203367, rs874881, rs2240340 and rs1748033) of the *PADI-4* gene do not contribute to genetic susceptibility to RA in Chinese (Chen, R et al.2010).

The function of *PADI-4* involves the post-translational modification of arginine to citrulline, and during this process citrullinted peptide antibodies (anti-CCP) are formed. Other evidence for the involvement of *PADI-4* in RA includes the presence

of *PADI-4* mRNA in haematological cells, pathological synovial tissues and blood of RA patients (Chang, X et al.2009). In humans, there are five types of *PADIs* of which *PADI2* and *PADI-4* are found in RA synovial membrane. Elevated PAD activity induced influx of calcium to the cell membrane and lead to cell death. Citrullinated proteins have been detected in the synovial membrane of patients with various forms of arthritis and in the inflamed tissue, which may imply that citrullination is associated with inflammation. In RA patients, anti-CCP autoantibody producing plasma cells have been found in the synovial membrane and higher anti-CCP autoantibody levels are found in the RA joint compared to serum. Auto-antibodies to anti-CCP are also highly expressed in synovial fluid. This auto-antibody is mapped to citrullinated alpa enolase peptide 1 *(CEP-1)* (Fisher et al.2011). It is reported that both anti-CCP and anti-CEP1 levels are higher in RA patients when compared with osteoarthritis or spondylarthritides (mean alpa enolase is 6.4 ng / L, 4.3 ng / L and < 0.9 ng / L in RA, SpA and OA controls, respectively) (Kinloch et al.2008). Sixty percentage of RA synovial fluid contained anti-CEP1 compared to 5 % in OA and none for SpA samples (Kinloch et al.2008). This implies that citrullinated protein expression is a result of inflammation and a site of up-regulated alpa–enolase. In one study of 373 RA patients, the association of anti-CCP with disease severity was affected by *PADI-4*-94 rs2240340, C>T (CC, OR = 0.93, CT, OR =2.92, TT, OR = 15.3) (Hoppe et al.2009).

2.3.5. The Tumour Necrosis Factor pathway

The WTCCC, the NARAC and a meta-analysis study identified three SNPs in the tumour necrosis factor (TNF) pathway that are associated with RA in Caucasians. They are *TNFA1P2* (rs2771369, $p = 7.57 \times 10^{-5}$) (WTCCC2007), *TNFA1P3*

(rs3761847, p = 2.8 x 10^{-8}, OR = 1.4, 95 % CI: 1.2-1.5) (Bowes and Barton2008; Patsopoulos and Ioannidis2010; Plenge et al.2007) and *TRAF1/C5* (rs6920220, p = 1.58 x 10^{-5}, OR = 1.33, 95 % CI: 1.33-2.22) .(Patsopoulos and Ioannidis2010) They map to chromosomes 14q32, 9q33 and 6q23, respectively. TNFA1P2 is also called the B94 protein 3. It was originally identified as a TNF–alpha inducible primary response gene in endothelial cells.

It is involved in haematopoietic tissue development and expressed in mononuclear progenitor cells in bone marrow, monocytes and in granulocytes.

The frequency for these SNP variants is different between Caucasians and Chinese. Hapmap data showed that the rs2771369 A-allele frequency is 0.144 in Caucasians and 0.378 in Asians; for rs3761847 the A-allele frequency is 0.483 in Caucasians and 0.511 in Asians; rs6920220 is not in polymorphic among Chinese.

2.3.5.1. Tumour Necrosis Factor, alpha induced protein 3 and oligodendrocyte transcription factor 3 locus

A GWAS study showed two SNPs, rs6920220 (p-value = 1 x 10^{-4}) and rs10499194 (p-value= 1 x 10^{-19}), on chromosome 6q23 are associated with RA in Caucasians (Plenge et al.2007). However, in one meta-analysis based on five GWAS studies it confirmed the association of rs69020220 with RA but rs104991944 could not be replicated (Patsopoulos and Ioannidis2010). These two SNPs are not polymorphic among Chinese populations. They are located in a 63 kbp linkage disequilibrium (LD) block. From correlation analysis, they appear to be two independent risk variants *TNFAIP3 / OLIG3* (WTCCC2007). They are commonly found in Caucasians (7 to 21%). They have a modest influence on the risk of RA (OR = 1.33, 95 % CI: 1.33-2.22, p = 1.58 x 10^{-5} and 1.33 for rs6920220 and rs10499194,

respectively). TNFAIP3 is a cytoplasmic *TNF* alpha inducible Zinc finger protein that functions as a ubiquitin editing enzyme modulating both TNF and Toll-like receptor induced NF-K beta activation. Mice without this enzyme develop severe inflammation in multiple organs including the joints, whereas mice lacking olig3, showed abnormal neuronal development. Factor analysis suggests rs10499194 is in LD with another SNP rs17066662 in Caucasians and it is in polymorphic in Chinese. Therefore, this tag SNP rs17066662 was included in this study.

2.3.5.2. Tumour Necrosis Factor Receptor-associated factor 1 - Complement Component 5

Tumour necrosis factor receptor-associated factor 1 - complement component 5 (*TRAF1-C5*) has a variant of rs3761847 which mapped to 9q33-34 (OR = 1.4, 95 % CI: 1.2-1.5; $p = 2.8 \times 10^{-8}$) which was associated with RA (WTCCC2007). This has been replicated in a large Caucasian cohort (Plenge et al.2007). The functions of TRAF1–C5 are not fully understood, but the function of *TRAF1* appears to be down-regulating signals mediated from TNF receptors, T-lymphocyte receptors and C5 is a member of the complement pathways.

2.3.5.3. Membrane Metallo-Endopeptidase-Like 1 Gene - Tumour Necrosis Factor Receptor Superfamily, Member 14

The SNP rs6684865 in membrane metallo-endopeptidase-like 1 gene- tumour necrosis factor receptor superfamily, member 14 (*MMEL1 / TNFRSF14*) is located in chromosome 1p36.32. The risk GG genotype was found to be associated with RA with modest GWAS p-value (OR = 1.54, 95 % CI: 1.25-1.90, $p = 3.14 \times 10^{-5}$) (Plenge et al.2007; WTCCC2007). The implication of the *MMEL1* gene in RA

required a very large sample size, but the nearby *TNFRSF14* gene is a TNF receptor super-family member. It interacts with TRAF proteins and activates the transcription factors AP-1 and induces the inflammation transcriptional regulator nuclear factor-κB (NF-κB) (Marsters et al.1997). Moreover, Perdigones et al. recently established that there was a gene-gene interaction between *TNFRSF14* and *TNFRSF6B* (OR=1.49) in Spanish patients with RA (Perdigones et al.2010).

2.3.6. Solute carrier family 22 and interleukin-13

The chromosome 5q31 region has been reported to be associated with multiple autoimmune diseases including Crohn's disease (Martinez et al.2006), SLE and RA. In 2003, a Japanese study reported that solute carrier family 22, member 4 (*SLC22A4*) which encodes the organic cation transporter 1 (*OCTN1*) and runt-related transcription factor 1 (*Runx1*) was associated with RA (Tokuhiro, S et al.2003). A variant of *SLC22A4* (rs2073838) with modified Runx1 binding site reduced its expression. *RUNX1* is a hematopoietic transcription factor and was found to be associated with leukemia, SLE and PSA. However, these associations have not been replicated in other Japanese study populations (Kuwahara et al.2005), where no associations were observed between *SLC2F1* and *RUNX1* in RA patients (OR = 1.24, 95 % CI: 0.92-1.67, p-value = 0.15 and OR = 0.85, 95 % CI: 0.64-1.05, p-value = 0.11,) or in Caucasians (Martinez et al.2006). In 2008, a GWAS study (He et al.2008) studied this 5q31 region of 725kb by including 99 SNPs that covered 14 genes, and the result showed only two SNPs rs11568506 (OR = 0.68, 95 % CI: 0.47-0.99, p = 0.043) and rs1800925 (OR = 0.77, 95 % CI: 0.67-0.88, p-value = 1.5×10^{-4}) were associated with autoimmune disease, psoriasis and Crohn disease (Li, Y et al.2008). The rs11568506 is located in the intronic region of *SLC22A4* (*OCTN1*) and the

rs1800925 is located at the 5' of the Interleukin-13 (*IL-13*) gene. These reports implied that the 5q31 cytokine cluster harbored multiple SNPs associated with psoriasis as well as other autoimmune diseases (Chang, M et al.2008).

2.3.7. Interleukin-2 Receptor Alpha and Interleukin-2 Receptor Beta

Previous reported studies from candidate gene markers and microsatellite genotyping often yielded only weak evidence for linkage of cytokine genes to RA even in a large sibling pair study (John et al.1998). In 2007, the WTCCC RA GWAS study showed SNP rs2104286 in the *IL2RA* gene mapped to chromosome 10p15 (OR = 1.68 95% CI: 1.31-2.14, p-value = 2.52×10^{-5}) and rs743777 in the *IL2RB* gene mapped to chromosome 22q13 (OR = 1.72, 95% CI: 1.40-2.11, p-value = 1.15×10^{-6}) were associated with RA in Caucasians (WTCCC2007) and these associations were replicated in another Dutch Caucasian population (Kurreeman et al.2009). Both WTCCC and the Dutch studies demonstrated that *IL2RA* and *IL2RB* are important in the development of RA. Lately, the *IL2RA* region was also shown to be associated with other autoimmune diseases such as type 1 diabetes (Vella2005) and multiple sclerosis (Matesanz et al.2007).

2.3.7.1. The structure and function of the interleukin 2 receptor

The *IL2* receptor is composed of three subunits: alpha (CD25), beta (CD122) and gamma (CD132) (Minami et al.1993). Only CD122 and CD132 have large intracellular domains and transduce signalling across the cell membrane. There are two forms of IL2 receptors: the low affinity dimer of CD122 and CD132, and the high affinity CD25, CD122 and CD132 trimer. The low affinity IL-2 receptor is found at low levels in $CD4^+$ T cells and high levels in $CD8^+$ and NK cells. The high

affinity IL-2 receptor is found in forkhead box P3 (FoxP3) positive $CD4^+CD25^+$ regulatory T (Treg) and activated T cells.

2.3.7.2. Interleukin-2 and interleukin-2 receptor

IL-2 is a 15kd cytokine first described as a T cell growth factor in 1976. In the 1990s, results from IL-2 and IL-2R knockout mice suggested that IL-2 is linked to severe autoimmunity. Subsequent studies linked this abnormality to IL-2-dependent $FoxP3^+CD4^+CD25^+$ Treg cells whose function is to maintain immune tolerance. IL-2 exhibits bipolar functions, where it promotes proliferation and differentiation of $CD4^+$ and $CD8^+$ T cells. It also induces apoptosis in $CD4^+$ cells maintaining the balance between protection from pathogens and suppression of self-reactive T cells.

2.3.8. Granzyme

Synovial tissue destruction has been attributed mainly to serine proteinases and matrix metallo-proteinases. They are produced by fibroblast-like synoviocytes, marcophages, chondrocytes and polymorphonuclear cells. There are granzyme A and granzyme B.

Granzyme A is a trypsin-like proteinase. In fibroblasts and epithelial cells, it induces IL-6 and IL-8 production, and in monocytes it also induces TNF alpha production.

Granzyme B is a serine protease that is expressed in cytotoxic T cells (CTL), natural killer (NK) cells, and human plasmacytoid dendritic cells (pDCs) (Jahrsdorfer et al.2010). Granzyme is stored in the granules of activated CTL, NK cells and pDCs.

During CTL and NK cells adherence to virus-infected or malignant cells,

granzyme containing granules move to the site of cell contact and dispose their cytotoxins to the target cell. Granzyme initiates apoptosis by digesting caspase and fragmentising target cell DNA. A number of studies showed that granzyme is involved in inflammation and destruction of RA joints. *In vitro* experiments showed that granzyme can digest cartilage into its proteoglycan subunits (Ronday et al.2001). A large number of granzyme positive NK cells are found in RA synovial tissue and synovial fluid when compared to their corresponding plasma samples. Thus, it is suggested that granzyme is produced locally at the inflamed site and directly digests the cartilage matrix in RA joints. Furthermore, free oxygen radicals, nitric oxide and destructive enzymes are located abundantly in this cartilage-pannus junction, which is associated with severe cartilage invasion. The number of granzyme B+ cytotoxic cells, T cells and fibroblast-like synoviocytes are related to the severity of joint damage.

Immuno-histochemical analyses, *in situ* hybridization and RT-PCR revealed that granzyme and pore-forming protein perforin (PFN) levels are raised in chondrocytes in RA cartilage. The levels of granzyme and perforin expression are similar to that of apoptotic cells (Horiuchi et al.2003). Recently, it was suggested that granzyme B might have roles other than the cytotoxic effect. In human plasmacytoid dendritic cells, granzyme B is secreted directly to T cells, it inhibits T cell proliferation by cell to cell contact and it is free from perforin co-stimulation (Jahrsdorfer et al.2010).

2.3.9. Chemokine (C-C motif) Receptor 6

CCR6 is the gene encoding for chemokine (C-C motif) receptor 6. It is a surface marker for the helper T cells (Th17 cells) that produce IL-17, and it is also expressed

on inactivated memory T-cells and dendrite cells. It is down regulated in activated T cells. Mouse animal models suggested Th17 cells play an important role in RA. Kochi el al (Kochi et al.2010) demonstrated that in a large Japanese GWAS, replication case control study (case, n=7069, control, n=20,729) a SNP rs3093024 at the *CCR6* region in 6q27 increased the risk for RA (OR = 1.27, 95% CI:1.18-1.37, p-value = 4.5×10^{-10}), where the risk allele was T. Furthermore, Kochi el al also demonstrated the expression level of CCR6a transcripts was increased with the number of the risk T-alleles. The role of CCR6 in the development of RA was further supported from another European GWAS meta-analysis (Stahl et al.2010), which consisted of 5,539 RA patients and 20,169 normal controls (OR = 1.13, 95 % CI: 1.08-1.18, p-value = 3.6×10^{-7}). However, these latest findings involved very large study population sizes and the effects of *CCR6* for RA were small.

2.3.10. Association of T-cell Receptor Alpha

The first study on the association of T-cell receptor alpha with autoimmune disease came from a study with narcolepsy in 2009 (Hallmayer et al.2009). Narcolepsy or dyssomnia is a chronic sleep disorder characterized by excessive daytime sleepiness. This study indicated that narcolepsy in man is associated with a single HLA allele rs1154155, DQB*0602 among all ethnic groups (OR = 1.69, 95 % CI: 1.52-1.88, $p < 1 \times 10^{-21}$). *TRA alpha* gene encodes for the TRA alpha-chain of the TRA alpha beta dimer. This heterodimer is expressed by T cells, and interacts with both MHC-HLA Class I antigens in CD8 cytotoxic cells and the MHC_class II antigens in CD4 T helper cells. The alpha and beta chains of the TRA receptor genes, like the immunoglobulin heavy and light chain loci, undergo somatic cell recombination. Specific TRA and TRB chains with specific functional variables and

J segments are formed. This mechanism ensures distinct TRA alpha beta idiotype-bearing T cells can be recognized and presented by MHC-HLA class I and II cells. At present, no other autoimmune diseases have been found to be associated with TRA.

2.3.11. Association of Protein Kinase C, theta

The SNP rs4750316 is in *PRKCQ* on chromosome 10 which encodes for the PRKCQ protein, and it was indicated to increase the risk of RA (p for trend = 5.55 x 10^{-5}).

2.3.12. Association of Kazal Type Serine Protease Inhibitor Domain

The SNP rs10786617 was one of the nine SNPs reported by the WTCCC study that had increased risk for RA with a weak GWAS p value (p for trend = 3.56×10^{-3} (WTCCC2007). In humans, it is located on chromosome 10. Studies of the mouse model suggested that the gene product of *KAZALD1* may play a role in bone development and bone healing after injury (Shibata et al.2004). This SNP may be involved in the prognosis of bone erosions in RA.

2.3.13. Association of Sex differentiated SNPs in RA

The WTCCC GWAS study demonstrated that the GG genotype of SNP rs11761231 in chromosome 7 increased the risk for RA (p for trend = 1.74×10^{-6}). Subgroup analysis showed that it posed no effect in male patients but in female patients, it increased the risk of RA (OR = 1.64, 95 % CI: 1.35 - 1.99, p for trend = 3.91 x 10^{-7}). This SNP appeared as a non-functional SNP and showed no linkage to any other SNPs. However, the physical position of this SNP (7q23, 131020579) is in

a recombination hotspot.

2.3.14. The Wellcome Trust Case Control Consortium GWAS 2007

The WTCCC GWAS study was published in 2007 and it was the first GWAS case-control study consisting of 14,000 cases and 3000 shared controls. It studied seven common diseases including rheumatoid arthritis. The study found 10 new SNPs with modest p for trend values range from ($3.45 \times 10^{-3} - 7.92 \times 10^{-6}$), it also reconfirmed genes in the *PTNP22*, and *MHC-HLA* regions were involved in RA risk. It also showed RA is a sex-differentiated disease this was reflected by the clinical data that there was a 4 to 1 ratio of female to male RA patient ratio. The study also showed that multiple diseases could be the result of a single SNP change. Furthermore, there were 47 SNPs with p-values between 0.1 to 1×10^{-5} involved in RA risk (WTCCC2007). Some of these SNPs were mapped to DNA regions with undetermined function, such as rs9550642 and rs1937506. These two SNPS are also included in the present study.

2.4. Gene to gene interaction in rheumatoid arthitis

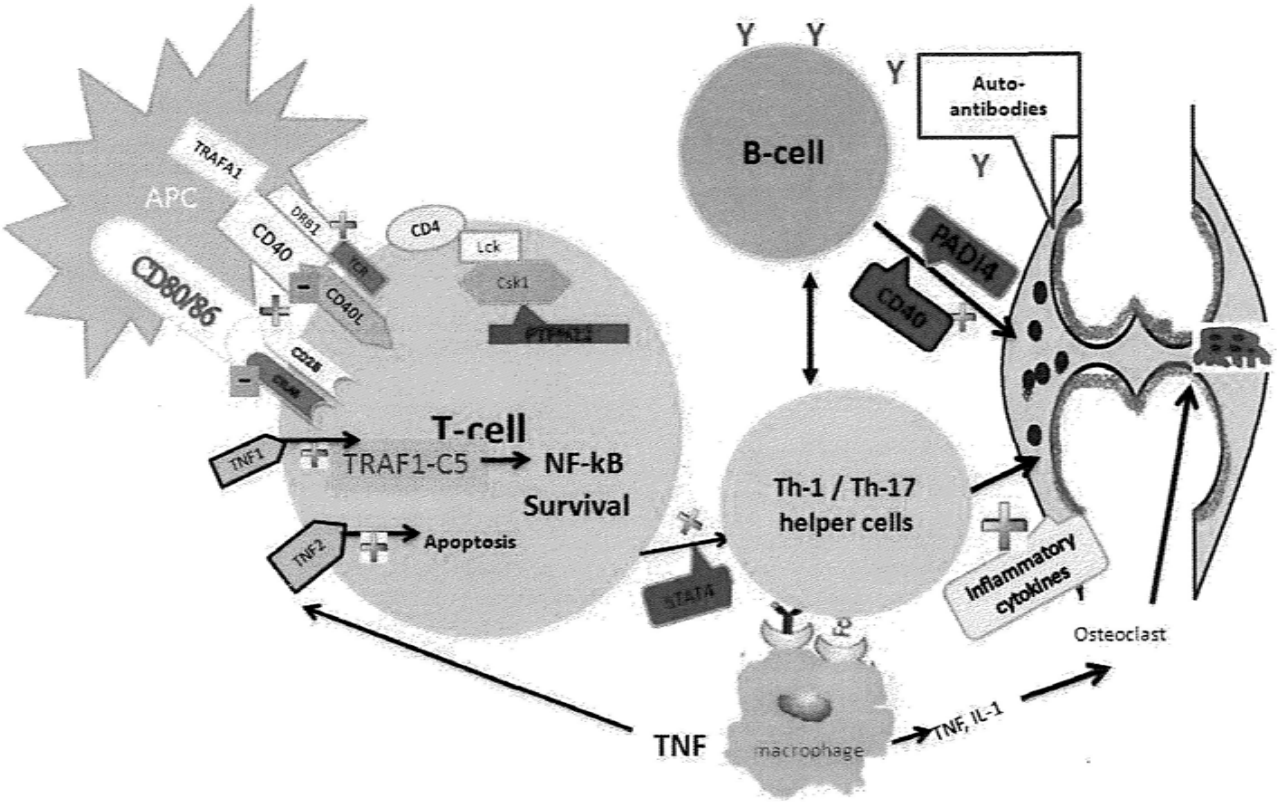
Epistasis is used to describe those phenotypes that only arise when specific alleles of two or more genes are present in the same individual. It is the result of a variety of underling mechanisms: proteins within a protein complex or the activity of a protein encoded by an allele of one gene are altered due to its interaction with anther protein encoded by a second gene. The effective size of such interaction depends upon the specific alleles encoding each of the induct proteins.

For example, A-allele of gene 1 alters the activity of the Y-allele of gene 2, but B-allele of gene 1 has no effect on the activity of the Y-allele of gene 2. The two

locus epistatic interaction can be either synergistic (additive effect) or antagonistic (reducing the activity).

Figure 2-2 A schematic diagram to show some of the loci studied in this study (Data extracted from chapter 3)

Cells	Partners	Cellular functions
APC	DRB1/TCR	⇧ T-cell proliferations
	CD40/CD40L	⇩ T-cell proliferations
	CD80/86 /CD28	⇧ T-cell proliferations
	CD80/86/ CTLA-4	⇩ T-cell proliferations
T-Cell	CD4 / PTPN22	⇧ T-cell proliferations
	STAT4	⇧ Th1 / Th-17
Macrophage	TNF-Th1/Th17	⇧ Inflammatory cytokines
	TNF-TNF1	⇨ Cell Survival
	TNF-TNF2	⇨ Apotosis
	TNF-Ostroclast	⇧ Bone resorption
	TNF-Chondrocyte	⇩ Cartilage formation
B-Cell	CD40-auto Abs	⇧ Bone resorption?
	PADI-4-	Anti-CCP ?



Chapter 3. Methodology

3.1. Objectives

Hypothesis: Genetic polymorphisms are associated with rheumatoid arthritis

Objectives:

1. Replicate study of association found in the rheumatoid arthritis predisposition loci identified from the WTCCC and NARAC genome wide association studies (Plenge et al.2007) in Chinese patients, in order to show whether these associations exist among the Chinese population.

2. To develop the significant findings in (1) into clinically useful prognostic indicators after controlling for other confounding factors and other prognostic factors.

3.2. Subjects and methods

3.2.1. RA patients -disease definition and selection criteria

Patients who fulfilled the following criteria were recruited:

a) Diagnosed by a qualified physician to have rheumatoid arthritis (RA) and fulfilling the American College of Rheumatology criteria for a diagnosis (possessing at least 4 of 7 criteria listed in table 1-3 and criteria 1 through 4 must have been present for at least 6 weeks) (Arnett et al.1988).

b) Ethnicity is Chinese.

c) Aged 18 years or over.

These ACR criteria had 91-94% sensitivity and 89% specificity for RA when compared with non-RA rheumatic disease control subjects.

Patients were recruited from the Rheumatology unit, Department of Medicine and Therapeutics, The Chinese University of Hong Kong at the Prince of Wales Hospital (Shatin, Hong Kong). All eligible patients who were registered with the clinic were invited to join the study. The study was conducted from January 2005 to December 2009.

All participants gave written informed consent before the interview and laboratory measurements. This study has been approved by The Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee.

3.2.2. Exclusion (cases)

Inability to give informed consent.

3.2.3. Normal Control subjects

Control subjects who fulfilled the following criteria were recruited:-

a) Ethnic Adult Chinese

b) No medical history of any joint disease

c) Have never had a history of metabolic bone disease

d) Have never been on oral or inhaled corticosteroid therapy for 3 months or more.

The controls subjects were recruited from an ongoing study of osteoporosis healthy community elderly male and female: Hong Kong Mr Os (Woo et al.2008) and Hong Kong Ms Os Study (Woo et al.2009; Woo et al.2008). The control cohort included healthy men and women of age 60-70 years, recruited from the community for examination of bone mineral density. A detailed questionnaire on disease risk factors, past medical history and overall health and medication was recorded. Subjects were questioned about the medical and surgical conditions that resulted in hospital admission, as well as by going through a check list of 17 conditions

previously found to be associated with bone mineral density (BMD). Subjects were asked to bring in all medications for identification.

Only subjects with absolutely no history of joint and bone symptoms were included in this study.

3.3. Sample size determination for genetic association study

To determine the sample size required for this study

Given: The frequency of a high risk genotype in control group = 20% (minor allele frequency is 0.2)

- Expected allelic odds ratio ~1.75 (average odd ratios defined in recent genome wide association studies)
- Power of 80% at a 2 sided α value of 0.001 (p value = 0.001 is used for Bonferroni correction)

The required number of subjects in a 1 case: 2 control design will be 346 cases and 692 controls (Gauderman2002).

3.4. Survey methods - Questionnaire

A standardized structured questionnaire was used to measure the clinical severity assessment including: extra-articular manifestations (for example, Sjögren syndrome, leg ulcers, and evidence of vasculitis such as nail fold infarcts, splinter haemorrhages, motor neuropathy, rheumatoid nodules, pulmonary disease and Felty syndrome), active joint count (number of tender or swollen joints) (Young2000), functional grade (Steinbrocker et al.1949), bone erosion of the hands and feet by

X-ray (Dixey et al.2004), rheumatoid factor (RF), number of joints irreversibly damaged (fixed deformity or surgical replacement) (Orces2002), the socioeconomic status and the patient's score on the modified health assessment questionnaire (M-HAQ) (Pincus, TS, C. Wolfe, F.1999).

We recorded treatment by patient review and chart review. Because treatment is often intermittent or at varying dosages, we tabulated medication use as "never," "former," or "current." Disease activity was assessed based on the 28 joint disease activity score (DAS 28), a validated composite score incorporating tender and swollen joint count, erythrocyte sedimentation rate (ESR), and a patient global assessment of disease activity (100 mm visual analogue scale). A DAS 28 $\leqq$1.6 indicates remission, whereas a value $\geqq$4.3 suggests active disease (Van Der Heijde et al.1992). Disease-related damage among RA patients was assessed by an index of irreversible joint damage. Bone erosion on X-ray data were obtained from patient medical records within plus or minus one month of the date when the blood sample was given for DNA extraction. Other data including smoking history, where current and ex- smoker were defined as 'ever smokers' and those who never smoked were defined as 'never smokers'. (James2004), Current and ex-alocohol drinkers were defined as 'drinkers', whereas non-drinkers were defined as 'never drinks'. Medical history, drug treatment, the disability index of the Health Assessment Questionnaire(HAQ) (Fries et al.1982), and family history of RA were noted. A DAS28 score of 2.4 indicates active RA.

3.5. Genomic DNA extraction

Genomic DNA was extracted from 10 mL peripheral blood samples using phenol-chloroform extraction (Yaich et al.1992). In brief, the blood samples were

centrifuged at 1,500 g for 10 minutes at 4℃ and plasma was removed. The leucocytes in the buffy coat were lysed by 100 mM ammonium solution. The DNA bound protein was digested with protease K overnight at 65 ℃. After that the DNA was extracted with phenol followed by chloroform. Finally the DNA was then allowed to dissolve in Milli-Q water at 4℃ for two days.

The concentrations of DNA samples extracted were measured by NanoDrop Spectrophotometer (Thermo Fisher Scientific, USA) and adjusted to 100 ng/mL. The quality of extracted DNA was checked with reference to the ratios of 260 nm /280 nm. An optical density value greater than 1.8 represented a good DNA quality with low levels of contamination. The DNA samples were further diluted to 25 ng/100 μL and stored in 96-well plate. Each plate contained 80 samples from column 1 to 10. Colour dye and control DNA samples were added to column 11 for PCR quantity control as well as plate identification tag. The extracted DNA samples were stored at -80 ℃ before genotyping.

3.6. SNP selection

The WTCCC study revealed a list of loci potentially associated with RA, however, they may not be the underlying functional SNPs causing the disease susceptibility. Therefore, the WTCCC results are considered as a means of high resolution fine-mapping to locate the predisposition variant and it provided the clue of the location but did not directly point out the exact functional SNP or the gene involved.

In the present study, based on the fine positional information from WTCCC I identified the potential regions containing the predisposition gene (Table 3-2). With

the information from HapMap, this set of candidate SNPs was defined and plotted to show their position in the autosomes (Table 3-3).

Chromosomes 1, 2 and 6 were indicated to be the most important regions, since these were the genetic regions where genes *PTPN22* (Plenge et al.2005), *STAT4* (Remmers et al.2007), and the highly versatile *MHC-HLA* are found. New regions were identified such as SNPs found in chromosome 10 and 14, which were mainly identified through the WTCCC GWAS study. However, previously reported regions such as the *SLC22A4-SLC22A5-IL13* regions of the organic cation transporter-interleukin 13 on chromosome 5 were also included in the SNP selection (Okada et al.2008; Tokuhiro, SY, R. Chang, X. Suzuki, A. Kochi, Y. Sawada, T. Suzuki, M. Nagasaki, M. Ohtsuki, M. Ono, M. Furukawa, H. Nagashima, M. Yoshino, S. Mabuchi, A. Sekine, A. Saito, S. Takahashi, A. Tsunoda, T. Nakamura, Y. Yamamoto, K.2003).

We replicated 25 SNPs from 17 loci and determined if there were gene-gene interactions in the predisposition to RA. The loci studied were major histocompatibility complex (*HLA-DRB1* rs615672 and *HLA-DQA1* rs6457617) (WTCCC2007), organic cation transporter–interleukin 13 (*SLC22A4-SLC22A5-IL13* rs2073838, rs1800925) (Li, Y et al.2008; Tokuhiro, SY, R. Chang, X. Suzuki, A. Kochi, Y. Sawada, T. Suzuki, M. Nagasaki, M. Ohtsuki, M. Ono, M. Furukawa, H. Nagashima, M. Yoshino, S. Mabuchi, A. Sekine, A. Saito, S. Takahashi, A. Tsunoda, T. Nakamura, Y. Yamamoto, K.2003), tumour necrosis factors (*TNFA1P2* rs2771369, *TNFA1P3* rs3761847, *MMELI-TNFRSF14* rs6684865 and *TNFA1P2* tagged SNP rs17066662) (Plenge et al.2007; WTCCC2007), interleukin-2 receptor (*IL-2RA* rs2104286, and *IL-2RB* rs743777) (WTCCC2007), cytotoxic T lymphocyte 4 (*CTLA-4* rs3087243 and rs10497873) (Costenbader et al.2008; Karlson et al.2008),

granzymes (*GZEB* rs854350) (WTCCC2007), signal transducer and activator of transcription 4 (*STAT4* rs7574865), (Remmers et al.2007) the peptidylarginine deiminase 4 (*PADI-4* rs874881 and rs2240340), anaphase-promoting complex 4 (*ANAPC4* rs3816587) (WTCCC2007) , protein kinase C, theta (*PRKCQ* rs4750316) (WTCCC2007), Kazal type serine protease inhibitor domain (*KAZALD* rs10786617) (WTCCC2007), chemokine (C-C motif) receptor 6 (*CCR6* rs3093024) , T cell receptor alpha (*TRA*-alpha rs1154155) (Hallmayer et al.2009), protein tyrosine phosphatase, non receptor type 22 (*PTPN22* rs2476601) (WTCCC2007) and 3 other SNPS rs11162922, rs1937506 and rs11761231 (WTCCC2007).

Table 3-1 List of candidate loci studied in the present study

Genes / Loci	SNP	Physical location[a]
Major histocompatibility complex		
HLA-DRB1	rs615672	6-34125905
HLA-DQA1	rs6457617	6-32771829
Organic cation transporter-interleukin 13		
SLC22A4-SLC22A5- IL13	rs2073838	5-13167721
	rs1800925	5-132020708
Tumour necrosis factor family		
TNFA1P2	rs2771369	14-10272331
TNFA1P3	rs3761847	9-122730060
MMEL1- TNFRSF14	rs6684865	1-2536089
rs17066662	rs17066662	6-138002706
Interleukin-2 alpha / beta		
IL-2RA	rs2104286	10-6139051
IL-2RB	rs743777	22-35881553
Cytotoxic T lymphocyte 4		
CTLA-4	rs3087243	2-204447164
CTLA-4	rs10497873	2-204470572
Granzymes		
GZEB	rs854350	14-24317403
Signal transducer and activator of transcription 4		
STAT4	rs7574865	2-191672878
The peptidylarginine deiminase 4		
PADI-4-92	rs874881	1-17533086
PADI-4-94	rs2240340	1-17535226
Aanaphase-promoting complex 4		
ANAPC4	rs3816587	4-25026342
Protein kinase C, theta		
PRKCQ	rs4750316	10-6433266
Kazal type serine protease inhibitor domain		
KAZALD	rs10786617	10-102825244

Table 3-1 List of candidate genes studied in the present study (*continued*)

Genes / Loci	SNP	Physical location
Chemokine (C-C motif) receptor 6		
CCR6	rs3093024	6-167452783
TRA-alpha	rs1154155	14-22072524
Protein tyrosine phosphatase, non-receptor type 22		
PTPN22	rs2476601	1-17535226
Other GWAS SNPs		
	rs11162922	1-80344646
	rs1937506	13-66933372
	rs11761231	10-6139051

[a] The physical positions of the SNPs were obtained from Genome Build 36.3, the reference assembly used in this study.

Table 3-2 Stem and leaf graph to show the candidate genes on chromosomes studied in this study

chromosome no.					No. of SNPs
1	*MMELI-TNFRSF14*	*PADI-4*	rs11162922	*PTPN22*	5
2	*STAT4*	*CTLA-4*	*CTLA-4*		3
3					
4	*ANAPC4*				1
5	*SLC22A4*	*IL-13*			2
6	*HLA-DRB1*	*HLA-DQA1*	*CCR6*	*rs17066662*	4
7					
8					
9	*TNFA1P3*				1
10	*IL-2RA*	*PRKCQ*	*KAZALD*	rs11761231	4
11					
12					
13	rs1937506				1
14	*TNFA1P2*	*TRA Alpha*	*GZEB*		3
15					
16					
17					
18					
19					
20					
21					
22	*IL-2RB*				1
				Total	25

SNPs denoted in red had been shown to have strong genetic association with RA in Caucasians

3.7. The International HapMap project

The human genome project (Consortium2005) was set up to determine the haplotypes of the human genome across the human races. It had 90 Caucasians from Utah, USA, 45 Chinese from Beijing, 45 Japanese from Tokyo and 90 Yoruba from Nigerian (Table 3-4).

Table 3-3 The composition of subjects of Phase II Human Genome project

Abbreviation	Population description	N=270
CEU	CEPH (Utah residents with ancestry from northern and western Europe)	90
CHB	Han Chinese in Beijing, China	45
JPT	Japanese in Tokyo, Japan	45
YRI	Yoruba in Ibadan, Nigeria	90

Over 310,000 SNPs (Consortium2007) had been typed and the genotypes of all of the 270 individuals could be viewed at the HapMap website. In this study we used the data of the Chinese and Japanese panels from the HapMap Phase II release 24. (HapMap Data Rel24/phaseII Nov08, on NCBI B36 assembly, dbSNP b126). An example for the SNP rs6578617 was included in table 3.5

http://hapmap.ncbi.nlm.nih.gov/cgi-perl/gbrowse/hapmap24_B36/

3.8. HapMap Lymphoblastoid Cell Lines

DNA obtained from the HapMap lymphoblastoid cell lines were used in this study as the reference genotypes (Table 3-5).

Table 3-4 The genotype of SNP rs615672 of the Chinese and Japanese HapMap subjects

CHB-HapMap number	rs615672	JPT-HapMap number	rs615672
NA18524	CC	NA18943	CC
NA18526	CC	NA18945	CC
NA18529	CC	NA18952	CC
NA18540	CC	NA18965	CC
NA18542	CC	NA18967	CC
NA18552	CC	NA18971	CC
NA18562	CC	NA18973	CC
NA18566	CC	NA18976	CC
NA18572	CC	NA18980	CC
NA18576	CC	NA18987	CC
NA18577	CC	NA18990	CC
NA18582	CC	NA18995	CC
NA18594	CC	NA18999	CC
NA18608	CC	NA19003	CC
NA18609	CC	NA19007	CC
NA18620	CC	NA18940	CG
NA18621	CC	NA18942	CG
NA18632	CC	NA18944	CG
NA18635	CC	NA18949	CG
NA18537	CG	NA18951	CG
NA18550	CG	NA18953	CG
NA18555	CG	NA18956	CG
NA18558	CG	NA18960	CG
NA18561	CG	NA18961	CG
NA18563	CG	NA18968	CG
NA18564	CG	NA18969	CG
NA18570	CG	NA18970	CG
NA18571	CG	NA18974	CG
NA18573	CG	NA18981	CG
NA18579	CG	NA18992	CG
NA18592	CG	NA18994	CG
NA18593	CG	NA18997	CG
NA18603	CG	NA19005	CG
NA18605	CG	NA18947	GG
NA18611	CG	NA18948	GG
NA18612	CG	NA18959	GG
NA18622	CG	NA18964	GG
NA18633	CG	NA18966	GG
NA18636	CG	NA18972	GG
NA18637	CG	NA18975	GG
NA18532	GG	NA18978	GG
NA18545	GG	NA18991	GG
NA18547	GG	NA18998	GG
NA18623	GG	NA19000	GG
NA18624	GG	NA19012	GG

3.9. Genotyping of candidate genes

All PCRs were prepared by a Biomek-2000 (Beckman Coulter, Fullerton, CA, USA) robotic system and carried out under standard conditions. Genotyping for SNPs in the candidate genes was carried out by restriction fragment length polymorphism (RFLP) PCR or allele specific high resolution melting (HRM) PCR in the laboratory of the Department of Chemical Pathology.

Tag SNPs that were identified among the candidate genes were genotyped in the whole sample set for the genetic association study. The total number of genotypes was 25 SNPs for the candidate loci.

3.9.1. SNP genotyping with restriction fragment length polymorphism

RFLP-PCR was performed with 25 ng genomic DNA, 25 pM sense and anti sense primers. PCR products were amplified by initial denature at 95℃ for 10 minutes, followed by 35 cycles of 94℃ for 30 s, 60℃ for 45 s, 72℃ for 40 s, and a final extension at 72°C for 10 min (Bio-Rad, i-Cycler, Hercules, CA). The amplified fragments were digested with specific endonuclease (New England BioLabs, Beverly, US) overnight at 37 ℃ (table 3.1). The fragments were resolved by electrophoresis in 3% agarose G-10 (Biowest, Spain) and stained with ethidium bromide. Genotypes were visualized and stored in gel documentation system (Syngene, Cambridge, England).

3.9.2. Primer design

The software Primer Version 0.5 developed by the Whitehead Institute for Biomedical Research (Cambridge, MA, USA) was used to design all primers. The DNA sequence for each SNP was downloaded from the NCBI data base

(http://www.ncbi.nlm.nih.gov/pubmed/) (Figure 3.1). DNA sequence from 120 bases upstream from the SNP and 120 bases downstream from the SNP was used for primer design. Primers generated from the program with annealing temperature around 58 to 62 ℃ and PCR product length of 200 bases pairs were selected for further process. A 14 base-pair 5' GC leading sequence (5'-gCgggCAgggCggC-3') with an additional restriction site for the restriction enzyme was added to the 5' end of the forward primer. The purpose of this tail was for quantity control during enzyme incubation. The PCR fragments sizes were deduced as show in Figure 3-1. Likewise, mismatch primers were also determined.

Figure 3-1 Schematic diagram to show the product sizes of rs2104286 cut by restriction enzyme *Nde*1

1. Part of the rs2104286 sequence extracted from NCBI database (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=2104286)

ATAGGAAATGTAAAAA

(A/G)

GTG

GTCCACTGTGCTCCTGCCCTGGACAGGGAGGGCGGGGTC

2. Targeted PCR product range to cover the restriction site CATATG

(A/G)

3. Addition of 14bp GC tail and an additional Nde1 restriction site CATATG

14bp Oligotaq Nde1 Forward primer

5'-gCgggCAgggCggCCATATGTGAGCGACATAGTCCTGATCCTTAAGAGAGT
CATAAGTTGGTGAGGAGGAGAAAGGCATAGATATAGTCATGGTAACACA
AGTCATATGTGGTAAGATCTACTGAGCATGGGGCTATCAGAGAAAGCTTC
ATGAGGGTGACATGGAAAGGGAGCCCTGAGGGACTGGTAAATTTCCATC
AGTGACATTCCAGGGAGAGAGGCCCCAAGACACAGGGGCAGCAG-3'

PCR product size = 243bp
Complete digestion by *Nde1* at 37 ℃ would yield three fragments and a 16 bp tag

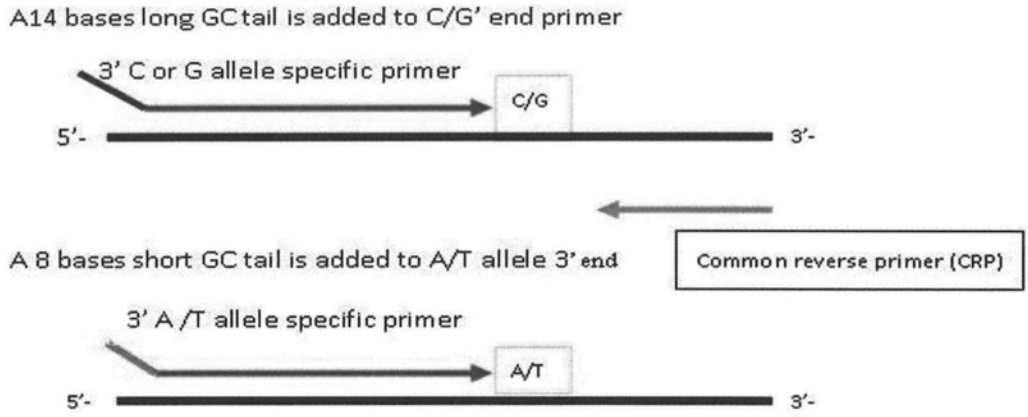
Genotype: GG=227
CG=89,137,227
CC=89,137

For incomplete digestion the 243 bp would appear on the agarose gel

3.9.3. PCR amplification and high resolution melting

A 3 prime allele-specific high resolution melting (temperature shift, Tm) PCR was used (Figure 3.2.). Each PCR reaction contained two 3' allele specific forward primers and a shared common reverse primer (Germer and Higuchi1999; Wang et al.2005). A 14 bp 5' GC leading sequence (5'-gCgggCAgggCggC-3') was added to one of the allele-specific primers, usually to the C or G, whereas a short 8 bp 5'GC sequence was added to the 5' leading end of the other primer, usually to the A, T or C-allele. The 6 bp differences between the two leading sequence would provide a 3 – 4°C difference in Tm between the allele-specific PCR products (Wang et al.2005). A common 22-30 reverse primer located at around 20 bp downstream from the SNP was used. A typical PCR product was around 70 bp with a melting temperature between 68-80°C (Figure 3-2).

Figure 3-2 Diagram to show the principle of high resolution melting

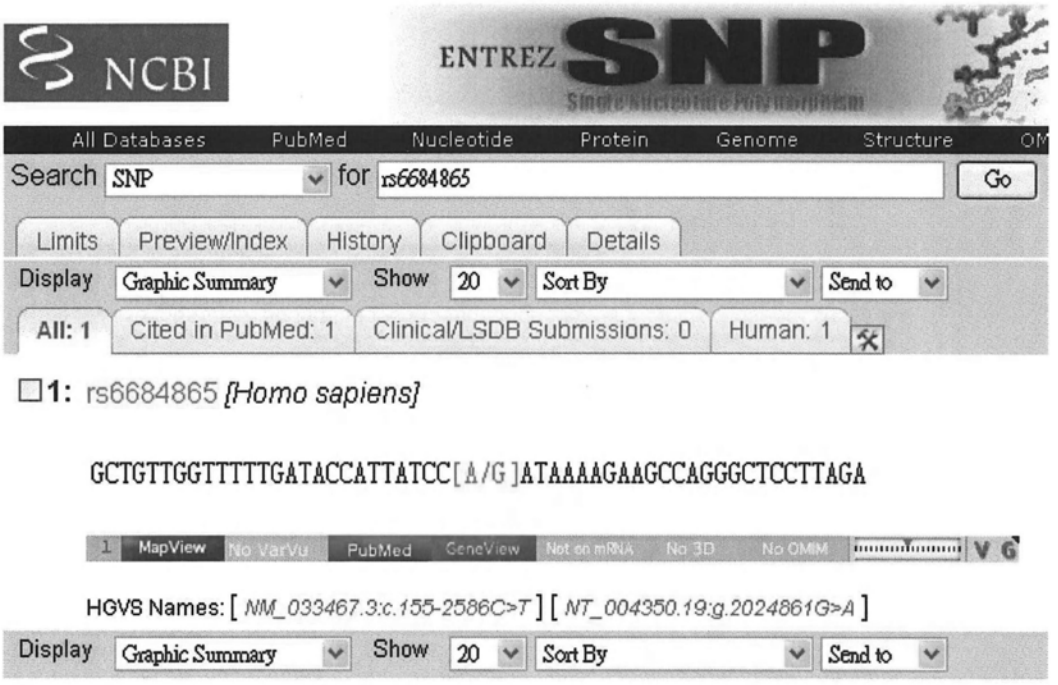


Each PCR has 2 forward and 1 common reverse primers

- CC or GG genotypes would be amplified with the long 14 GC-bp primer and only gave a single high temperature peak
- AA or TT genotypes would be amplified with the short 8 GC bp primer and only gave a single low temperature peak
- Two peaks would be formed when the genotype was heterozygous

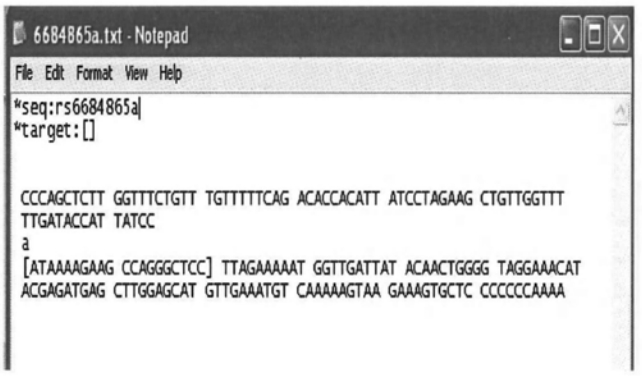
Figure 3-3 A schematic diagram to show the primer design for high resolution melting

1. SNP information was obtained from NCBI



2. DNA sequence was copied and saved as a txt file

ACATTGGAAT GGAAAAAGTA AGTGGATAGC AGGTGGCTGG GCTTATTGCT GCTGCAGTGA
GAAGTTACAG ACAAGGAAAG GAGGAAGTCA GAATGGTCCA TGTGGTGGTG GAGTAGAGTT
GGAGACAACA GTGTGAACTC ATTTTTAGCT TAATATAGAT GCAGATGGTT GCATATATAA
ATATTTATAG ATATGAGTAT CTAGATGGGT TCATGTATAC ATACATATTC CTTTGCTCTG
CCAGGTGAGC TAGACTAGAA GCAACACCAA CCAAGGAGTA GCAATAAGCA AAGAATAGGA
CCCAGCTCTT GGTTTCTGTT TGTTTTTCAG ACACCACATT ATCCTAGAAG CTGTTGGTTT
TTGATACCAT TATCC
R
ATAAAAGAAG CCAGGGCTCC TTAGAAAAAT GGTTGATTAT ACAACTGGGG TAGGAAACAT
ACGAGATGAG CTTGGAGCAT GTTGAAATGT CAAAAAGTAA GAAAGTGCTC CCCCCCAAAA
AACTCACCCC AAATTCACGA TAATGGGGAT ACATCAGAGG GTCATAGAAG CAACTGACAA
CTCCCAATGG CCAGAGTTAG AACAATTTGA GCAATGCAAT ATAGCAGTGG ATTATAATGC
AAAGTGTAAA ATAAATATCT ACAAATCATT GAGTAAATAA ATAAATGAGG GAGAATAGAT



3. Primer Version 0.5 was used to generate the primer sequence.

4. Primer sequences were saved as the following file.

```
*****************************************************************************
* Output from:                                                              *
*                                                                           *
*                              PRIMER                                       *
*                           (Version 0.5)                                   *
*                                                                           *
*     Copyright 1991, Whitehead Institute for Biomedical Research           *
*****************************************************************************


                                          |Self   |--L1-- -THE - -ALU -|
? base primer 5'->3'              Tm   GC% |any 3' |any 3' any 3' any 3'|
>  113 GTTGGTTTTTGATACCATTATCC   56.6  35% | 5  4  |12 10  10 10   6  5 | ok
<  167 CATTTTTCTAAGGAGCCCTGGCTTCT 66.7  46% | 8  8  |11 11  10  9   7  6 | ok

PCR product length:   55, GC = 36%, pairwise match (any 3' 3'): 3 1 3
```

5. GC-tails were added to the forward primers.

All primers were ordered from a local company TechDragon Ltd (Shatin, Hong Kong)

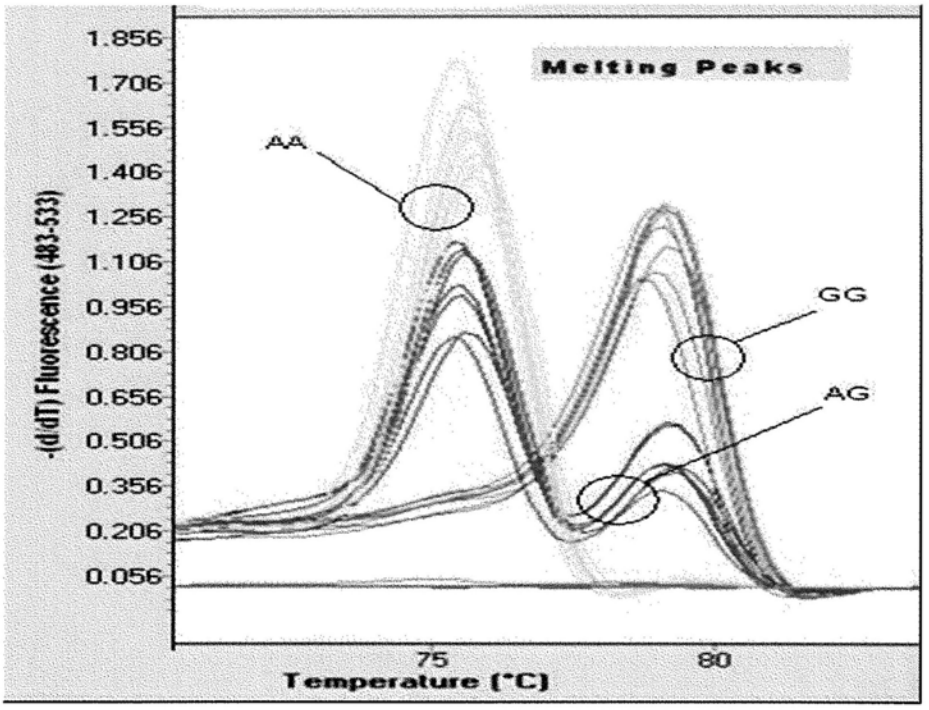
3.9.4. SNP genotyping with high resolution melting

All PCRs were set up using the BIOMEK@2000 robotic system (Beckman Coulter, Fullerton, USA). For each 12.5 μL PCR reaction mixture contained 25 ng genomic DNA, 1 x TaqGold buffer, 0.4 nM of each dNTP, 0.2 X SyBR Green I (Invitrogen, Carlsbad, CA, US), 10% dimethyl sulfoxide (DMSO) (Sigma-Aldrich, Dorset, England), 0.6 unit of Ampli Taqgold polymerase (Applied Biosystem, CA, USA) and 1 to 2.5 mM magnesium chloride ($MgCl_2$) in 1-2 pM forward or reverse primers (Jiang et al.2010; Tang et al.2010).

PCR products were amplified by initial polymerase activation at 95°C for 15 min, followed by 35 cycles at 95°C for 20 sec, 60°C for 1 min, 72°C for 30 sec and a final extension at 72°C for 10 min in a BioRad PCR machine with 384 block (Bio-Rad, i-Cycler, Hercules, CA). Melting curve analysis was performed and analyzed with software V1.5 on LightCycler 480 Real-time PCR system (Roche Applied Science, Indianapolis, US). The melting curve data were collected by detecting the fluorescence intensity of the PCR products in a denaturation ramp from 60°C–90°C at 0.5°C/s temperature increment.

Primers for the HRM were designed by the program Primer 0.5 with allele specific 3'-end. The primers sequences, primer ratios, PCR conditions, PCR product sizes, and the genotype frequencies obtained from HapMap were presented in tables 3-6 to 3-18. A typical melting profile of genotype was included in Figure 3-4.

Figure 3-4 Melting profiles of genotypes for SNP rs2073838, using Roche LightCycler® 480



The yellow peaks represent homozygous alleles (AA) with low melting temperature (Tm), the blue peaks represent homozygous alleles with high Tm (GG) and the red double peaks represent heterozygous alleles.

3.10. Statistical analysis for the association between genotype and rheumatoid arthritis

In order to conceal the identity of subjects, data were coded and analyzed with Statistical Package for Social Science (SPSS) for windows, version 13.0 (Somers, New York, US). The genotypes were compared between RA and control groups by Cochran-Armitage test for trend with Microsoft Excel or Fisher exact test (Ahn et al.2007). The association between RA patients – controls and the three genotypes within each SNP were checked for Hardy-Weinberg equilibrium. A p-value greater than 0.05 was considered insignificant (Armitage1955; Cochran1954). Predisposing alleles were found at higher frequencies in the case group. OR, 95% CI, and Chi-square in the dominant allele genetic model and the recessive genotype model (Lettre et al.2007) were calculated with Statcalc, EpiInfo Version 6 (Dean et al.1994).

Multivariate analysis was performed with categorical forward conditional binary logistic regression to analyze the predictive power of age, sex and each of the 24 SNPs in the development of RA. A p-value smaller then 0.05 was considered significant.

Haplotypes were generated from genotypes of SNPs that are located within a genetic locus by an EM algorithm, Haploview v4.1. Haplotype specific risks were determined. Type I error due to multiple statistical testing were corrected by spectral decomposition (Nyholt2004).

3.11. Gene to gene interaction

In this study, a log-regression model was used to determine the gene-gene interactions between these loci.

Gene-gene interactions were examined by whole-genome association analysis toolset PLINK ver1.07 (Purcell et al.2007).

http://pngu.mgh.harvard.edu/~purcell/plink/index.shtml

3.12. Anti-CCP antibodies determination

Anti-CCP antibodies were determined with second generation enzyme-linked immunosorbent assay (ELISA), from Immunoscan CCPlus (Fernandez-Suarez et al.2005; Liao et al.2011) (Euro-Diagnostica AB, Malmö, Sweden).

Table 3-5 *IL-2RA*, rs11761231 and *ANAPC4* primer sequences, conditions used in restriction fragment length polymorphism

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size bp
				Mg2+	restriction enzymes	
IL-2RA	10p15	Chromosome position: 6139051	AA=A, A/G			
rs2104286			A/G(0.86/0.14)			Uncut:243
rs2104286F		gCgggCAgggCggCCATATGTgAgCgACATAgTCCTgATCC	AA:AG:GG	1.5mM	*NdeI*	GG=227
rs2104286R		CTgCTgCCCCTgTgTCTT	0.73:0.24:0.02		37℃	AG=89,137,227
						AA=89,137
rs11761231	7q23	Chromosome position:131020579	AA=C, C/T			Uncut=179
rs11761231F		gCgggCAgggCggCggCCTgTgACTACATTTggAgATAAgTAggc	T/C(0.72/028)			TT=163
rs11761231R		TCTATAgCTCTTCTgATCCTCACCC	TT:CT:CC	1.5mM	*HaeIII*	CT=28,135,163
			0.53:0.38:0.09		37℃	CC=28,135
ANAPC4	4	Chromosome position:25026342	AA=C, C/T			Uncut:194
rs3816587F		gCgggCAgggCggCgTACGTTTAGGAATAGGATCAGAC	C/T(0.66/0.34)		*Rsa*	TT=180
rs3816587R		CTGCAGGGCAGTAAGGTTCT	CC:CT:TT	1.5mM	37℃	CT=28,152,180
			0.44:0.44:0.12			CC=28,152

AA=Ancestral Allele

Table 3-6 ***MHC-HLA*** **primer sequences, conditions used in Tm shift allele specific PCR reaction**

Primer	loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size (bp)
				Mg2+	Primer ratio Long:Short	
MHC-HLA-DRB1						
rs615672	6q23	Chromosome position: 34125905	AA=G, G>C			
rs615672g		gCgggCAgggCggCTTCAgTgTTTgTggTTAggAAAAG	G/C (0.65/0.34)	2.0mM	1:2	74(GG)
rs615672c		gATTACCg TTCAgTgTTTgTggTTAggAAAAC	CC:CG:GG			68,74(CG)
rs615672CRP		gAcggtgcATTcTgCTgTTgt	0.42:0.47:0.11			68(CC)
MHC-HLA-DQAI						
rs6457617	6q23	Chromosome position: 32771829	AA=C, C/T			
rs6457617c		gCgggCAgggCggCCATATgCACAgATCTTTgTTAgTCAC	C/T (0.478/0.52)	1.5mM	1:2	84(CC)
rs6457617t		gATTACCgCATATgCACAgATCTTTgTTAgTCAT	TT:TC:CC			78,84(CT)
rs6457617crp		AgTggACTCAATTACTgTTTgTTgAgTCC	0.31:0.42:0.27			78(TT)

AA=Ancestral Allele

Table 3-7 ***SLC22A4-SLC22A5- IL-13*** **primer sequences, conditions used in Tm shift allele specific PCR reaction**

Primer	loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size (bp)
				Mg2+	Primer ratio Long:Short	
SLC22A4						
rs2073838	5	Chromosome position: 13167721	AA=G, G>A			
rs2073838g		gCgggCAgggCggCCAgAgCCCTgAAAAAACACTAAg	G/A(0.79/0.21)			64(GG)
rs2073838a		gATTACCg CAgAgCCCTgAAAAAACACTAAA	GG:GA:AA	2.0mM	1:2	58,64(CG)
rs2073838CRP		TgCAAgCCCTgAgTCAggCA	0.62:0.33:0.04			58(AA)
IL-13						
rs1800925	5	Chromosome position: 132020708	AA=T, T>C			
rs1800925c		gCgggCAgggCggCGTTTCTGGAGGACTTCTAGGAAAAC	C/T(0.844/0.156)			79(CC)
rs1800925t		gATTACCgGTTTCTGGAGGACTTCTAGGAAAAT	CC:CT:TT	1.5mM	1:2	73,79(CT)
rs1800925crp		CCCTGCAGCCATGTCGCCT	0.71:0.27:0.02			73(TT)

AA=Ancestral Allele

Table 3-8The tumour nerosis factor (TNF) pathway primer sequences, conditions used in Tm shift allele specific PCR reaction

Primer	loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size (bp)
				Mg2+	Primer ratio Long:Short	
TNFA1P2						
rs2771369	14	Chromosome position: 10272331	AA=A, A>G			
rs2771369g		gCgggCAgggCggCgTTgggACCggCTTAACg	G/A(0.622/0.378)			70(GG)
rs2771369a		gATTACCgTTgggACCggCTTAACA	GG:GA:AA	2.0mM	1:2	64,70(CG)
rs2771369crp		gCCTAgCCTCTgAgCATAACTTCC	0.36:0.53:0.11			64(CC)
TNFA1P3	9	Chromosome position: 122730060	AA=G, G>A			
rs3761847g		gCgggCAgggCggCCCCTTCTCTCCCCTCCg	A/G(0.511/0.489)			67(GG)
rs3761847a		gATTACCgCCCTTCTCTCCCCTCCa	AA:AG:GG	2.0mM	1:2	61,67(AG)
rs3761847crp		ggAATgAgCAggTAgAgAgggTg	0.24:0.53:0.23			61(AA)
rs17066662	6	Chromosome position: 138002706	AA=T, T>C			
rs17066662c		gCgggCAgggCggCCAACATTCTATTTCAATTCCCAGAC	C/T(0.922/0.078)			75(CC)
rs17066662t		gATTACCgCAACATTCTATTTCAATTCCCAGAT	CC:CT:TT	2.0mM	1:3	69,75(CT)
rs17066662crp		GGTGGCGCTAATGTATTTTAGCTGG	0.84:0.16:0			69(TT)

AA=Ancestral Allele

Table 3-8 ***(continued)*** **The tumour nerosis factor (TNF) pathway primer sequences, conditions used in Tm shift allele specific PCR reaction**

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size (bp)
				Mg2+	Primer ratio Long:Short	
MMEL1-TNFRSF14						
rs6684865	1	Chromosome position: 2536089	AA=G G/A			
rs6684865g		gCgggCAgggCggCGTTGGTTTTTGATACCATTATCCG	A/G(0.520/0.480)			73(GG)
rs6684865a		gATTACCgGTTGGTTTTTGATACCATTATCCA	AA:AG:GG	2mM	1:3	67,73(AG)
rs6684865crp		CATTTTTCTAAGGAGCCCTGGCTTCT	0.24:0.56:0.20			67(AA)

AA=Ancestral Allele

Table 3-9*IL-2RB* and *GZEB* primer sequences, conditions used in Tm shift allele specific PCR reaction

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size (bp)
				Mg2+	Primer ratio Long:Short	
IL-2RB	22q13	Chromosome position: 35881553	AA=G, G>A			
rs743777g		gCgggAgggCggCCATTTTACAGGCCAGGAAACCG	A/G (0.830/0.170)			77(GG)
rs743777a		gATTACCgCATTTTACAGGCCAGGAAACCA	AA:AG:GG	2mM	1:3	71,77(AG)
rs743777CRP		TCTGCCACATTCTGGTAGGATGAC	0.69:0.29:0.02			71(AA)
GZMB						
rs854350	14	Chromosome position: 24317403				
rs854350g		gCgggCAgggCggCCATTgCCCATTggTACCg	AA=NA, C>G			74(GG)
rs854350C		gATTACCgCATTgCCCATTggTACCC	C/G (0.660/0.340)	2mM	1:3	68,74(CG)
rs854350crp		TCACTAgTgAAACTAggAACACCAgAgg	CC:CG:GG			68(CC)
			0.38:0.56:0.07			

AA=Ancestral Allele

Table 3-10 Primer sequences of two SNPs in the *CLTA-4* genes, conditions used in Tm shift allele specific PCR reaction

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size (bp)
				Mg2+	Primer ratio Long:Short	
CTLA-4						
rs3087243	2	Chromosome position: 204447164	AA=G, G>A			
rs3087243g		gCgggCAgggCggCTTCACCACTATTTgggATATAACG	G/A(0.79/0.21)			87(GG)
rs3087243a		gATTACCgTTCACCACTATTTgggATATAACA	GG:GA"AA	2mM	1:3	81,87(AG)
rs3087243crp		CCAATTgATTTATAAAggACTgCTATgTCTgT	0.64:0.29:0.07			81(AA)
CTLA-4						
rs10497873	2	Chromosome position: 204470572	AA=G, G>A			
rs10497873g		gCgggCAgggCggCTTCCCCTCAAAGGTCCAGAG	G/T(0.88/0.12)			81(GG)
rs10497873t		gATTACCgTTCCCCTCAAAGGTCCAGAT	GG:GT:TT	2mM	1:2	75,81(GT)
rs10497873crp		CATGAAAATGATGGCTGTATTCATTGAC	0.78:0.20:0.02			75(TT)

AA=Ancestral Allele

Table 3-11 ***GZEB*** **and** ***STAT4*** **primer sequences, conditions used in Tm shift allele specific PCR reaction**

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size (bp)
				Mg2+	Primer ratio Long:Short	
GZEB						
rs854350	14	Chromosome position: 24317403	AA=G, G>C			
rs854350g		gCgggCAgggCggCCATTgCCCATTggTACCg	G/C(0.66/0.34)			68(GG)
rs854350c		gATTACCgCATTgCCCATTggTACCC	GG:GC:CC	2mM	1:2	62,68(GC)
rs854350crp		TCACTAgTgAAACTAggAACACCAgAgg	0.07:0.56:0.37			62(CC)
STAT4						
rs7574865	2	Chromosome position: 191672878	AA=T, G>T			
rs7574865g		gCgggCAgggCggCgAAAAgTTggTgACCAAAATgTG	G/T(0.67/0.33)			81(GG)
rs7574865t		gATTACCgAAAAgTTggTgACCAAAATgTT	GG:GT:TT	2mM	1:2	75,81(GT)
rs7574865crp		AATCCCCTgAAATTCCACTgAAATAAgAT	0.42:0.49:0.09			75(TT)

AA=Ancestral Allele

Table 3-12 ***PADI-4-92*** **and** ***PADI-4-94*** **primer sequences, conditions used in Tm shift allele specific PCR reaction**

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size (bp)
				Mg2+	Primer ratio Long:Short	
PADI-4-92						
rs874881	1	Chromosome position: 17533086	AA=G, G/C			
rs874881g		gCgggCAgggCggCAAGCTCTACTCTACCTCACCGG	G/C(0.400/0.600)	1.5mM	1:2	71(GG)
rs874881c		gATTACCgAAgCTCTACTCTACCTCACCGC	GG:GC:CC			65,71(GC)
rs874881crp		CggCACTCTAggATCCTggTTG	0.18:0.44:0.38			65(CC)
PADI-4-94						
rs2240340	1	Chromosome position: 17535226	AA=A, G>A			
rs2240340g		gCgggCAgggCggCGGAGATTTCTGAAATCCCATCAG	G/A(0.59/0.41)	2mM	1:3	67(GG)
rs2240340a		gATTACCgGGAGATTTCTGAAATCCCATCAA	GG:GA:AA			61,67(AG)
rs2240340crp		GGACCCTCACCAACCTCTCCT	0.36:0.47:0.18			61(AA)

AA=Ancestral Allele

Table 3-13 ***PRKCQ* and *KAZALD1* primer sequences, conditions used in Tm shift allele specific PCR reaction**

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size (bp)
				Mg2+	Primer ratio Long:Short	
PRKCQ						
rs4750316	10	Chromosome position: 6433266	AA=G, G>C			
rs4750316g		gCgggCAgggCggCCATTTggAAggAATgAgAgAATTAg	G/C(0.87/0.13)			74(GG)
rs4750316c		gATTACCgCATTTggAAggAATgAgAgAATTAc	GG:GC:CC	2mM	1:3	68,74(CG)
rs4750316crp		AAACCAgACCCTCATTgTCACCTAA	0.76:0.22:0.02			68CC)
KAZALD1	10	Chromosome position: 102825244	AA=C, C>T			
rs10786617c		gCgggCAgggCggCAATgCTggTACTgATATCTATAATCTTTCTAC	C/T(0.69/0.31)			74(CC)
rs10786617t		gATTACCgAATgCTggTACTgATATCTATAATCTTTCTAT	CC:CT:TT	2mM	1:2	72,78(CT)
rs10786617crp		CACAgATggCCCTAACCAAggT	0.47:0.44:0.09			72(TT)

AA=Ancestral Allele

Table 3-14 ***CCR6*and *TRA alpha* primer sequences, conditions used in Tm shift allele specific PCR reaction**

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size
				Mg2+	Primer ratio Long:Short	
CCR6						
rs3093024	6	Chromosome position: 167452783	AA=A , A/G			
rs3093024g		gCgggCAgggCggCGAACTGGACGCTCGG	G/A(0.63/0.37)			61(GG)
rs3093024a		gATTACCgGAACTGGACGCTCGCA	GG:GA:AA	1.5mM	1:1	54,61(GA)
rs3093024crp		ACGTGTGACGGATCAGTGAAGAA	0.36:0.52:0.12			54(AA)
TRA alpha						
rs1154155	14	Chromosome position: 22072524	AA=T, T/G			
rs1154155g		gCgggCAgggCggCAAAGGCTATATAAGCATGCTTTCG	G/T(0.54/0.46)			82(GG)
rs1154155t		gATTACCgAAAGGCTATATAAGCATGCTTTCT	GG:GT:TT	2mM	1:3	76,82(TG)
rs1154155crp		GGAAAAAGTCTTAGTGGTTAGCTAAATTCCAC	0.31:0.42:0.27			76(TT)

AA=Ancestral Allele

Table 3-15 SNP rs11162922, and rs1937506 primer sequences, conditions used in Tm shift allele specific PCR reaction

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size
				Mg2+	Primer ratio Long:Short	
rs11162922	1	Chromosome position: 80344646	AA=A , A>G			
rs11162922g		gCgggCAgggCggCAATTCTAGAGTACTTTCAAGATTTAAAAGTCG	A/G(0.78/0.22)			87(GG)
rs11162922a		gATTACCgAATTCTAGAGTACTTTCAAGATTTAAAAGTCA	AA:AG:GG	2mM	1:3	81,87(GA)
rs11162922crp		TCTGCTCAATGGAATGATGTGTATG	0.58:0.40:0.02			81(AA)
rs1937506	Chr13	Chromosome position: 66933372	AA=G, G>A			
rs1937506g		gCgggCAgggCggCCAAATAATgAAgTTATACCTgTggAgAg	G/A (0.86/0.14)			85(GG)
rs1937506a		gATTACCgCAAATAATgAAgTTATACCTgTggAgAa	GG:GA:AA	2mM	1:2	79,85(AG)
rs1937506crp		CCTTTggACATATCCATTACAATgTTCC	0.73:0.24:0.03			79(AA)

AA=Ancestral Allele

Table 3-16 SNP rs1154155 primer sequences, conditions used in Tm shift allele specific PCR reaction

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size
				Mg2+	Primer ratio Long:Short	
rs1154155	14	Chromosome position: 22072524	AA=T, T/G			
rs1154155g		gCgggCAgggCggCAAAGGCTATATAAGCATGCTTTCG	G/T(0.54/0.46)			82(GG)
rs1154155t		gATTACCgAAAGGCTATATAAGCATGCTTTCT	GG:GT:TT	2mM	1:3	76,82(TG)
rs1154155crp		GGAAAAAGTCTTAGTGGTTAGCTAAATTCCAC	0.31:0.42:0.27			76(TT)

AA=Ancestral Allele

Table 3-17 ***PTPN22*** **primer sequences, conditions used in Tm shift allele specific PCR reaction**

Primer	Loci	Sequences (5' – 3')	Alleles and Frequency (CHB HapMap data)	Condition		Product size
				Mg2+	Primer ratio Long:Short	
PTPN22	1	Chromosome position:114179091	AA=G G/A			
rs2476601g		gCgggCAgggCggCACAATAAATGATTCAGGTGTCCG	G/A(0.89/0.11)			75(GG)
rs2476601a		gATTACCgACAATAAATGATTCAGGTGTCCA	GG:GA:AA	1.5mM	1:2	69,75(GA)
rs2476601crp		GATGAAATCCCCCCTCCACTTC	0.96:0.04:0.00			AA(69)
PTPN22	1	Chromosome position:114179091	AA=G C/T			
rs2476601c		gCgggCAgggCggCACAATAAATGATTCAGGTGTCCC	C/T(NA)	2.0mM	1:3	75(CC)
rs2476601t		gATTACCgACAATAAATGATTCAGGTGTCCT	CC:CT:TT			69,75(CT)
rs2476601crp		GATGAAATCCCCCCTCCACTTC	NA			69(TT)

AA= Ancestral Allele

Chapter 4 Results

4.1. Characteristics of the RA patients

Six hundred RA patients are recruited, of whom 120 are men and 480 are women.

For each RA patient, their haplotypes for the 24 SNPs are compared with two normal control subjects. The control cohort is in general 10 years older than the case cohort.

The demographic and clinical data of the study participants are summarized in Table.4.1. The mean age is 58.3 ± 12.2 years in men and 55.2 ± 13.2 years in women. The RA onset age in men is 50.0 ± 13.6 and 45.8 ± 14.6 in women. Duration of the disease in men is 8.6 ± 7.4 years and 9.8 ± 9.0 years in women.

RF, erosions on X ray, rheumatoid nodules and anti-CCP results are presented in Table 4.1 as percentage. The mean CRP level is 20.2 ± 28.2 IU/mL in men and 15.9 ± 22.9 IU/mL in women.

Eighty-two percent of the men are serum anti-CCP positive with mean value of 7157 ± 33934 U/mL and 82% of the women showed serum anti-CCP positive with mean value of 10498 ± 43473 U/mL. eventy-one percent of the men are RF positive with mean RF level of 292 ± 401 mg/L and 63.9 % of the women are RF positive with mean RF 267 ± 343 mg/L. Sixty-two percent men showed both anti-CCP+ and RF+ and 57 % women were both positive for anti-CCP and RF.

ESR in men is 38.0 ± 30.3 mm/hour and 44.5 ± 32.4 mm/hour in women.

Bone erosion on X-ray is 66.3% in men and 65.6 % in women. Rheumatoid nodules among men in are 7.8 % and among women are 5.3%. The number of tender joints in men are 2.1 ± 3.1 and in women are 2.5 ± 3.7. The number of swollen joints

in men are 1.6 ± 2.7 and in women are 1.7 ± 2.6. DAS28 score among men are 3.2 ± 1.2 and among women are 3.4 ± 1.1. Present of a family history for RA among men is 4.2% and among 7.1% women.

Physician global score among men are 4.6 ± 2.3 and 5.0 ± 2.3. Patient global score in men were 4.1 ± 2.5 and 4.5 ± 2.9 in women.

Fifty-eight percent of the women are post-menopausal. Smokers and drinkers are predominately found in men. Sixty-seven percent of the men and 6.4% of the women were ever cigarette smokers. Forty-six percent of the men were ever alcohol drinkers and 8.2 % of the women were ever alcohol drinkers.

Active RA disease status in men is 44.0% in men and 42.2% in women.

Table 4-1 Demographic data of the RA patients

Total (n =600)	Male (n=120)	Female (n=480)	All (n=600)
Age (years)	58.3 ± 12.2	55.2 ± 13.2	55.7 ± 13.0
Onset age (years)	50.0 ± 13.6	45.8 ± 14.6	46.6 ± 14.5
Duration of RA (years)	8.6 ± 7.4	9.8 ± 9.0	9.0 ± 8.5
anti-CCP$^+$ (%)	84.2	82.0	82.2
anti-CCP U/mL	7157 ± 33934	10498 ± 4347	9825 ± 41714
Log anti-CCP	2.93 ± 0.71	2.81 ± 0.81	2.83 ± 0.79
RF$^+$ (%)	71.3	63.9	65.4
RF (mg/L)	292 ± 401	267 ± 343	275.6 ± 358
anti-CCP$^+$ and CRP$^+$ (%)	62.4	57.7	58.6
CRP (IU/mL)	20.2 ± 28.2	15.9±22.9	16.6 ± 24.0
ESR (mm/hour)	38.0 ± 30.3	44.5±32.4	43.2 ± 32.0
Erosion on X-ray (%)	63.2	62.7	62.8
Rheumatoid Nodule + (%)	7.8	5.3	5.8
Physician global score	4.6 ± 2.3	5.0 ± 2.3	4.9 ± 2.3
Patient global score	4.1 ± 2.5	4.5 ± 2.9	4.4 ± 2.8
Tender Joints	2.1 ± 3.1	2.5 ± 3.7	2.4 ± 3.6
Swollen Joints	1.6 ± 2.7	1.7 ± 2.6	1.7 ± 2.6
DAS28	3.2 ± 1.2	3.4 ± 1.1	2.9 ± 1.4
Active RA (%)[1]	44.0	42.2	42.6
Family history for RA (%)	4.2	7.1	6.5
Post-menopausal	-	58.%	-
Ever smoker (%)	67.0	6.4	18.8
Ever alcohol drinker (%)	46.0	8.2	16.0

Data are presented as mean ± SD for continous variables, and frequency (percentage) for categorical variables.

[1] Active RA: RA>2.4

4.2. Genetic risk for rheumatoid arthritis

The –log p-values obtained using the Cochran-Armitage trend test are plotted against chromosomal number for the increased risk of RA for the 24 SNPs (Figure 4-1). Increased risk for RA is associated with polymorphisms in seven SNPs. Three SNPs are associated with increased risk of RA for the whole group of patients, *HLA-DRB1* (p for trend = 1.8×10^{-4}), *HLA-DQA1* (p for trend = 2.0×10^{-16}), and *STAT4* (p for trend = 3.9×10^{-9}). One SNP *PADI-4-92* (p for trend = 0.005) increased the risk of RA in men only, whereas *TNFA1P2* (p for trend = 0.024), *MMEL1-TNFRSF14* (p for trend = 3.3×10^{-4}, and rs11761231 (p for trend = 1.7×10^{-4}) are associated with increased risk for RA in women only.

The –log p-values obtained using odds ratios are presented in the dominant genetic model (Figure 4-2). In the dominant model the polymorphisms in *STAT4* (OR = 0.48, 95 % CI = 0.38 – 0.61, p-value = 9.8×10^{-10}) and *HLA-DQA1* (OR = 0.35, 95 % CI = 0.25 – 0.48, p-value < 1.0×10^{-7}) are associated with increased risk for RA for the whole group of patients. *PADI-4*-94 (OR = 2.16, 95 % CI = 1.28 – 3.66, p-value = 0.002) and *ANAPC4* (OR = 0.59, 95% CI = 0.37 – 0.94, p-value = 0.02) are associated with RA in men only, whereas *TNFA1P2* (OR = 1.36, 95% CI = 1.05 – 1.77, p-value = 0.016), *MMEL1-TNFRSF14* (OR = 0.48, 95% CI = 0.34 – 0.67, p-value = 9.6×10^{-6}), *SLC22A4* (OR = 1.28, 95 % CI = 1.00 – 1.64, p-value = 0.043), and rs11761231 OR = 0.52, 95 % CI = 0.38 – 0.71, p-value = 2.4×10^{-5}) are associated with RA in women only.

In the recessive model (Figure 4-3) the polymorphism in *HLA-DRB1* (OR = 1.8, 95 % CI = 1.34 – 2.44, p-value = 5.7×10^{-5}) and *HLA-DQA1* (OR = 2.17, 95 % CI = 1.71- 2.72, p-value < 1.0×10^{-7}) are associated with risk for RA in both men and women. *STAT4* (OR = 2.38, 95 % CI = 1.28 – 4.41, p-value = 2.8×10^{-3}), *ANAPC4*

(OR = 0.29, 95 % CI = 0.09 – 0.90, p-value = 0.017), and *SLC22A4* (OR = 2.18, 95 % CI = 1.04 – 4.58, p-value = 0.024) are associated with the risk for RA in men only, whereas *CTLA-4* (OR = 2.01, 95 % CI = 1.03 – 3.97, p-value = 0.027), *TNFA1P2* (OR = 0.67, 95 % CI = 0.45 – 1.01, p-value = 0.044), *TNFA1P3* (OR = 0.68, 95 % CI = 0.49 – 0.94, p-value = 0.015) and *IL2RA* (OR = 0.42, 95 % CI = 0.17 – 0.99, p-value = 0.030) are associated with the risk for RA in woman only.

In summary, the increased risk for RA was associated with:

In the whole cohort: *STAT4*, *HLA-DRB1*and *HLA-DQA1*

In men only: *PADI-4*, and *ANAPC4*.

In women only: TNF loci (*MMEL1-TNFRSF14*, *TNFA1P2*, *TNFA1P3), CTLA-4, SLC22A4, IL2RA* and rs111761231.

Figure 4-1 –log p-values were plotted against chromosomal number for the increased risk of RA with the 24 SNPs

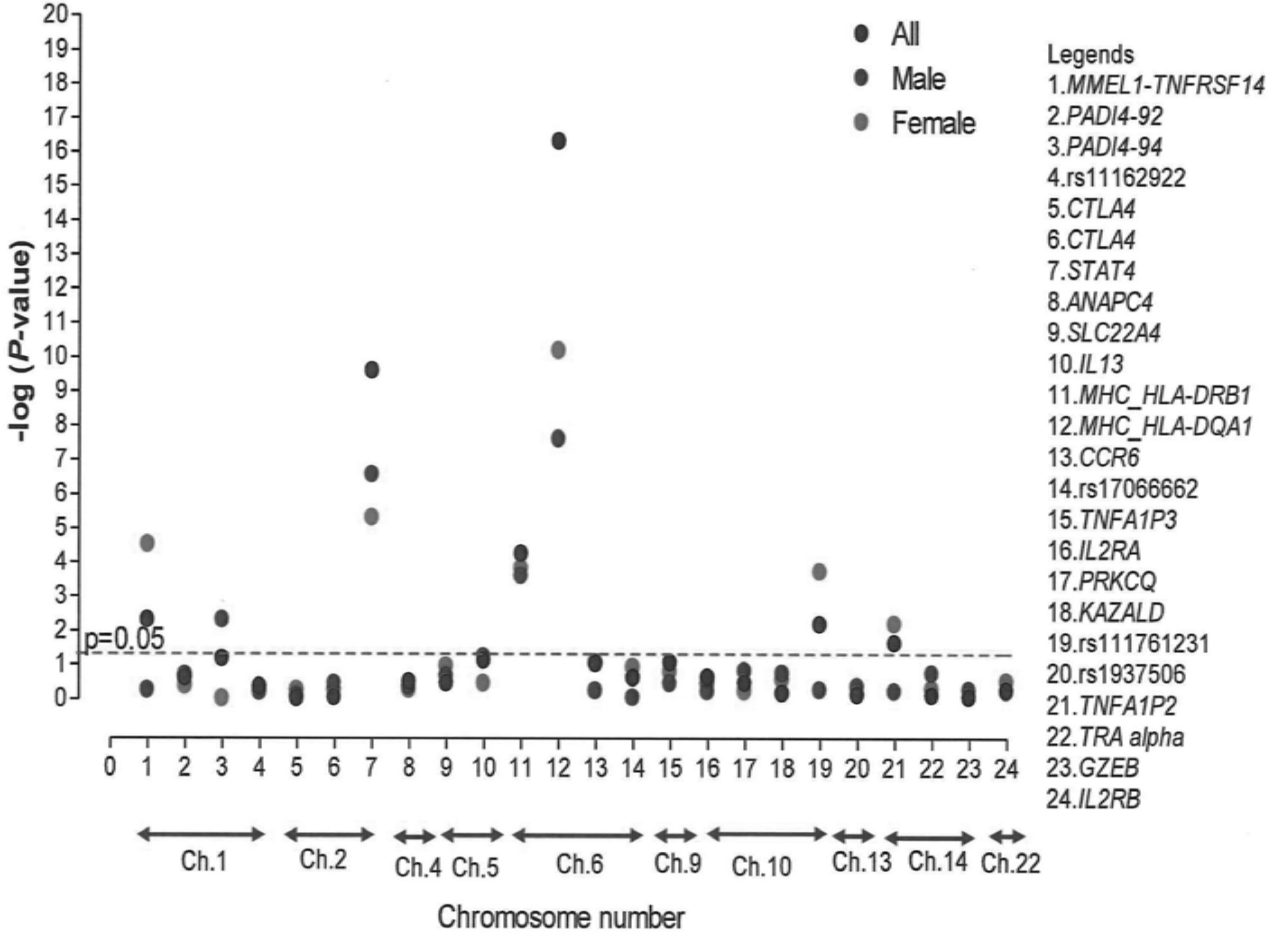


Figure 4-2 Association of RA risk in the dominant genetic model

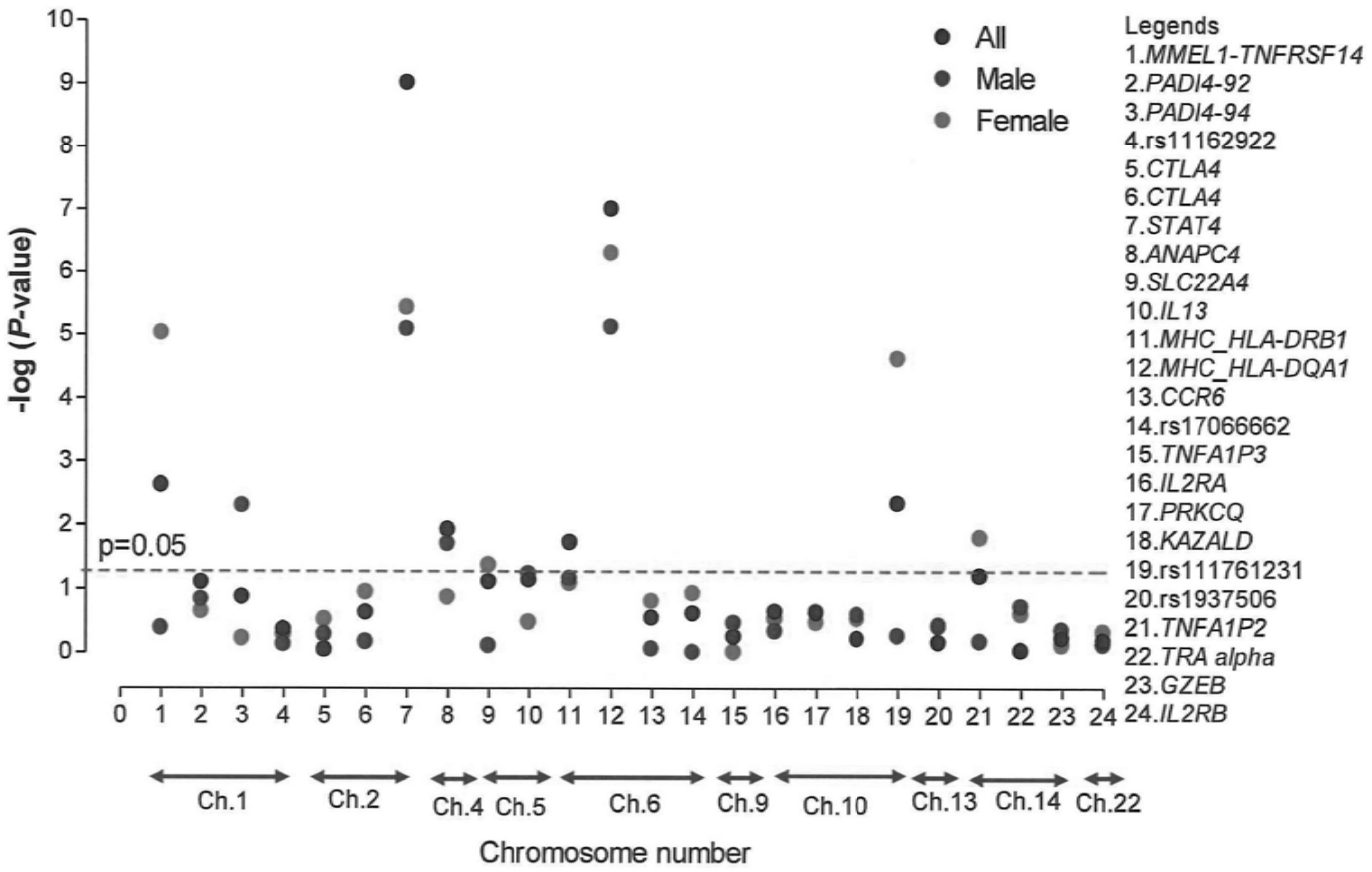
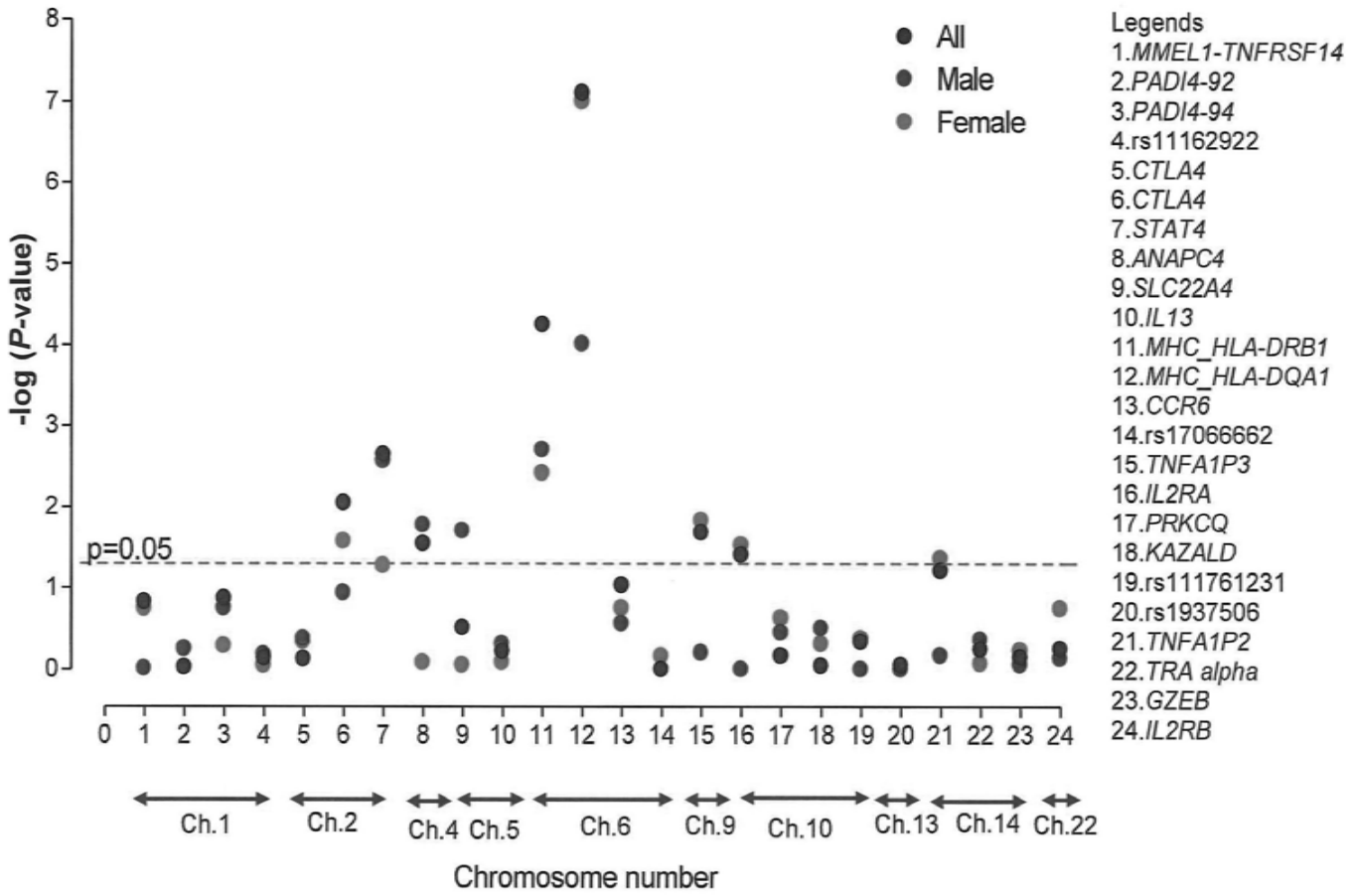


Figure 4-3 Association of RA risk in the recessive genetic model



4.2.1. Association of Major Histocompatibility Complex, Class II, DR beta 1 and Major Histocompatibility Complex, Class II, DQ alpha 1

The genotype frequencies for the two MHC SNPS are shown in Table 4-2. Genotypic frequencies of the two MHC SNPs followed HWE ($p>0.05$). The GG minor allele frequency of rs615672 in *HLA-DRB1* was significantly higher in the case patients then in the normal controls (20% versus 12%; p for trend = 1.8×10^{-4}). Upon subgroup analysis, these associations are retained in both RA men and women, (p for trend = 3.9×10^{-3} and 6.6×10^{-3}, respectively). In the dominant model the G-allele was a risk allele (OR = 1.30, 95 % CI: 1.03-1.63, p-value = 0.021) and when analyzed in a recessive model, RA patients carrying the GG genotype showed an increase in risk for RA (OR = 1.80; 95% CI: 1.34 - 2.44; p-value = 5.7×10^{-5} The increased risk for RA in men was 2.42 (OR = 2.42; 95% CI: 1.31 - 4.47; p–value = 2.0×10^{-3}) in a recessive model, and in women it was 1.64 (OR=1.64, 95% CI: 1.15-2.33; p-value = 3.9×10^{-3}).

The TT minor allele frequency of rs6457617 in *HLA-DQA1* was significantly higher in the case patients then in the normal controls (44% versus 27%; p for trend = 2.0×10^{-16}), and when analyzed in a dominant model, the T-allele in RA patients showed a increase in risk for RA (OR = 2.89, 95% CI: 2.08 – 4.02; p-value < 1.0×10^{-7}). In a recessive model the TT genotype showed an increase in risk for RA (OR = 2.17, 95% CI: 1.71 – 2.72; p-value < 1.0×10^{-7}). Upon subgroup analysis both male and female RA patients showed an increased TT minor allele frequency (p for trend = 4.0×10^{-7}, 1.5×10^{-10}, respectively). In a recessive model, the TT-alleles showed an increased risk for RA in both male and female RA patients (OR = 2.51, 95% CI: 1.53 - 4.12; p-value = 9.7×10^{-3}; OR = 2.05, 95 % CI: 1.58 - 2.69; p-value = 1.0×10^{-7}, respectively). The dominant model showed an increase in risk for RA when

compared to controls (OR = 5.08, 95% CI: 2.25 – 11.89; $p = 7.6 \times 10^{-6}$) in males and (OR = 2.45, 95% CI: 1.69 – 3.55; $p = 5.0 \times 10^{-7}$) in females.

Table 4-2 The allele and genotype frequencies of *HLA-DRB1*and *HLA-DQA1* SNPs in RA patients and controls

SNP	N	Genotype frequencies, n (%)					Dominant Model		Recessive Model	
							OR (95%CI)	P^C	OR (95% CI)	P^C
HLA-DRB1										
rs615672		CC	CG	GG	*Ptrend*[a]	HWE[b]	CG + GG vsCC		GG vs CG + CC	
Case	560	187 (34)	259 (46)	112 (20)	$1.8x10^{-4}$	NS	1.30 (1.03-1.63)	0.021	1.80 (1.34-2.44)	$5.7x10^{-5}$
Control	866	344 (40)	416 (48)	106 (12)						
Male RA	115	34 (30)	54 (47)	27 (23)	$3.9x10^{-3}$	NS	2.57 (1.61-4.14)	$2.7x10^{-5}$	2.42 (1.31-4.47)	$2.0x10^{-3}$
Male control	267	105 (39)	132 (49)	30 (11)						
Female RA	442	153 (35)	204 (46)	85 (19)	$6.6x10^{-3}$	NS	1.25 (0.96-1.63)	$8.2x10^{-2}$	1.64 (1.15-2.33)	$3.9x10^{-3}$
Female Control	595	239 (40)	284 (47)	76 (13)						
HLA-DQAI										
rs6457617		CC	CT	TT			CT + TT vs CC		TT vs CT + CC	
Case	569	57 (10)	262 (46)	250 (44)	$2.0x10^{-16}$	NS	2.89 (2.08-4.02)	$<1.0x10^{-7}$	2.17 (1.71-2.72)	$<1.0x10^{-7}$
Control	813	198 (24)	399 (49)	216 (27)						
Male RA	113	8 (7)	54 (48)	51 (45)	$4.0x10^{-7}$	NS	5.08 (2.25-11.89)	$7.6x10^{-6}$	2.51 (1.53-4.12)	$9.7x10^{-5}$
Male control	251	70 (28)	119 (47)	62 (25)						
Female RA	456	49 (11)	208 (46)	199 (43)	$1.5x10^{-10}$	NS	2.45 (1.69-3.55)	$5.0x10^{-7}$	2.05 (1.58-2.69)	$1.0x10^{-7}$
Female Control	562	128 (23)	280 (50)	154 (27)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test;

MHC-HLA: major histocompatibility complex- human leukocyte antigen

4.2.1.1. The linkage disequilibrium plot of rs615672 and rs6457617

The LD plot of rs615672 and rs6457617 was constructed by Haploview. No association was found between these two SNPs where D' equal to 15 (Figure 4-4).

Figure 4-4 The LD plot between rs615672 and rs6457017

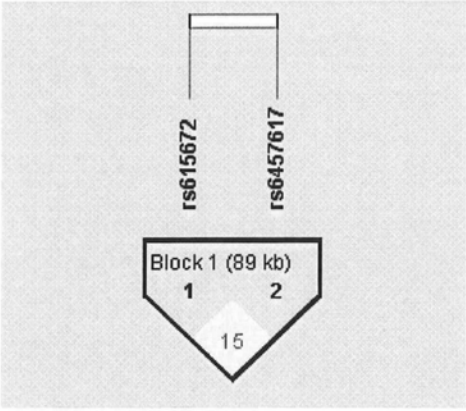


Figure 4-5 The frequencies of the haplotypes of the study samples

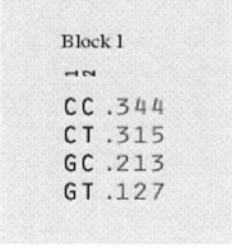


Table 4-3 The Haploview p values of rs615672 and rs6457617

Name	Assoc Allele	Case, Control Ratio Counts	Case, Control Frequencies	Chi square	P value
1 rs615672	G	477:625, 496:1264	0.433, 0.282	68.889	1.04×10^{-16}
2 rs6457617	C	764:376, 825:879	0.670, 0.484	95.865	1.23×10^{-22}

Table 4-4 The p-values for 10,000 permutation tests

#10000 permutations performed.		
Name	Chi Square	Permutation p-value
rs615672	68.889	0.00E+00
rs6457617	95.865	0.00E+00
Block 1: CT	85.009	0.00E+00
Block 1: GC	111.181	0.00E+00
Block 1: GT	2.103	0.3956
Block 1: CC	0.884	0.7187

4.2.2. Association of Solute Carrier Family 22 and Interleukin-13

The genotype frequencies for the two SNPS rs2073838 and rs1800925 are presented in Table 4-5. Genotypic frequencies of the two SNPs followed HWE ($p > 0.05$). There are no statistical significance for the genotype AA-allele in RA patients in rs2073838, but there was an increase in percentage of AA minor allele genotype in male RA patients (14% versus 7%, p for trend = NS), and in a recessive model the minor AA allele showed an increased risk of two-fold in male RA patients (OR = 2.18, 95 % CI: 1.10 - 4.35; p-value = 0.024). However, in the dominant model a small decrease in risk is observed with the AA alleles in female RA patients (OR = 0.78, 95% CI: 0.61 - 1.00; p-value = 0.043).

No genetic associations are found for rs1800925 in RA patients when compared to controls. The minor TT allele frequency of rs1800925 was similar in the case patients and in the normal controls (Table 4-5).

Table 4-5 The allele and genotype frequencies of *SLC22A4-SLC22A5-IL-13* region in RA patients and controls

SNP	n	Genotype (%)					Dominant Model		Recessive Model	
							OR (95 % CI)	P^b	OR (95 % CI)	P^b
SLC22A4										
rs2073838		GG	GA	AA	*Ptrend*[a]	HWE[b]	GA + AA vs GG		AA vs GA + GG	
Case	561	293 (52)	207 (37)	61 (11)	NS	NS	0.83 (0.67-1.03)	NS	1.20 (0.83-1.72)	NS
Control	930	442 (48)	402 (43)	86 (9)						
Male RA	113	52 (46)	45 (40)	16 (14)	NS	NS	1.06 (0.67-1.68)	NS	2.18 (1.04-4.58)	0.024
Male control	299	142 (47)	136 (46)	21 (7)						
Female RA	448	241 (54)	162 (36)	45 (10)	NS	NS	0.78 (0.61-1.00)	0.043	0.97 (0.65-1.48)	NS
Female Control	631	300 (48)	266 (42)	65 (10)						
IL-13										
rs1800925		CC	CT	TT			CT + TT vs CC		TT vs. CT + CC	
Case	595	438 (74)	144 (24)	13 (2)	NS	NS	0.82 (0.65-1.03)	NS	0.84 (0.42-1.67)	NS
Control	1235	859 (70)	344 (28)	32 (2)						
Male RA	120	91 (76)	27 (23)	2 (2)	NS	NS	0.63 (0.38-1.04)	NS	0.58 (0.09-2.95)	NS
Male control	319	212 (66)	98 (31)	9 (3)						
Female RA	475	347 (73)	117 (25)	11 (2)	NS	NS	0.89 (0.69-1.15)	NS	0.92 (0.42-2.00)	NS
Female Control	916	647 (70)	246 (27)	23 (3)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

4.2.3. Association of Tumour necrosis factor pathway

Genotypic frequencies of the four SNPs *TNFA1P2*, *TNFA1P3*, *MMELI-TNFRSF14* and rs17066662 from the TNF pathway are presented in Table 4-6. Genotypic frequencies of the four SNPs followed HWE ($p > 0.05$). The AA minor allele frequency of rs2771369 in *TNFA1P2* was significantly lower in the case patients then in the normal controls (10% versus 14%; p for trend = 0.024). Upon subgroup analysis, this association is only seen in female RA patients (p for trend = 0.0064). In a dominant genetic model there was a small decrease for the risk of RA in the AA genotype (OR = 0.74, 95 % CI: 0.57-0.95; p = 0.016) and when analyzed in a recessive model, RA patients show that the minor AA genotype was a protective allele (OR = 0.67, 95 % CI: 0.45-1.01; p–value = 0.044).

No association is found in rs3761847 in *TNFA1P3* for RA patients. However, upon subgroup analysis, the minor GG genotype in the recessive model reduces the risk for RA in female patients (OR=0.68, 95% CI: 0.49-0.94; p-value=0.015).

In the *MMELI-TNFRSF14* loci, the AA minor allele frequency of rs6684865 is significantly higher in the case patients then in the normal controls (23% versus 20%; p for trend = 0.005). Upon subgroup analysis, this association is only seen in female RA patients (p for trend = 3.3×10^{-4}). In a dominant genetic model, the A allele shows a two fold increase for the risk of RA (OR = 2.11, 95 % CI: 1.49-2.98; $p = 9.6 \times 10^{-6}$. The frequency of the GG genotype in the male control (25%) is lower than female controls (35%). However, no statistical difference in the genotype frequency between the male and female controls is observed (p=NS).

There are no genetic associations in rs17066662. The percentage of the rare minor TT genotype was similar in both RA patients and normal controls. Upon subgroup analysis, no associations are observed between RA patients and normal

controls in both dominant and recessive genetic models.

These results showed there was an association between RA in female patients in the TNF pathway.

Table 4-6 The allele and genotype frequencies of *TNFA1P2*, and *TNFA1P3* of the *TNF* pathway in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^C	OR (95% CI)	P^C
TNFA1P2 rs2771369		GG	GA	AA	*P trend*[a]	HWE[b]	GA+AA vs GG		AA vs. GA+GG	
Case	577	276 (48)	242 (42)	59 (10)	0.024	NS	0.82 (0.65-1.02)	NS	0.73 (0.51-1.03)	NS
Control	804	344 (43)	351 (44)	109 (14)						
Male RA	114	59 (52)	43 (38)	12 (11)	NS	NS	0.91 (0.58-1.45)	NS	0.86 (0.40-1.82)	NS
Male control	275	136 (49	106 (39)	33 (12)						
Female RA	463	217 (47)	199 (43)	47(10)	0.0064	NS	0.74 (0.57-0.95)	0.016	0.67 (0.45-1.01)	0.044
Female Control	529	208 (39)	245 (46)	76(14)						
TNFA1P3 rs3761847		AA	AG	GG			AG+GG vs AA		GG vs. AG + AA	
Case	559	170 (30)	299 (53)	90 (16)	NS	NS	0.94 (0.74-1.18)	NS	0.73 (0.55-0.96)	0.021
Control	1019	296 (29)	510 (50)	213 (21)						
Male RA	112	34 (30)	59 (53)	19 (17)	NS	NS	0.80 (0.48-1.32)	NS	0.87 (0.47-1.59)	NS
Male control	299	77 (26)	165 (55)	57 (19)						
Female RA	447	136 (31)	240 (54)	71 (16)	NS	NS	1.00 (0.77-1.30)	NS	0.68 (0.49-0.94)	0.015
Female Control	720	219 (30)	345 (48)	156 (22)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

Table 4-6(*continued*) The allele and genotype frequencies of rs17066662,and *MMEL1-TNFRSF14*of the *TNF* pathway in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^c	OR (95% CI)	P^c
TNF pathway rs17066662		CC	CT	TT	*P trend*[a]	HWE[b]	CT+TT vs CC		TT vs. CT + CC	
Case	560	482(86)	75(13)	3(1)	NS	NS	1.20 (0.87-1.66)	NS	1.23 (0.22-6.51)	NS
Control	919	810(88)	105(11)	4(1)						
Male RA	112	96(86)	16(14)	0(0)	NS	NS	1.00 (0.51-1.93)	NS	0.00 (0.00-45.61)	NS
Male control	293	251(85)	41(14)	1(1)						
Female RA	448	386(86)	59(13)	3(1)	NS	NS	1.34 (0.91-1.97)	NS	1.40 (0.23-8.69)	NS
Female Control	626	559(88)	64(10)	3(1)						
MMEL1-TNFRSF 14 rs6684865		GG	GA	AA	*P trend*[a]	HWE[b]	GA + AA vs GG		AA vs. GA + GG	
Case	553	121 (22)	302 (55)	130 (23)	0.005	NS	1.51 (1.15-1.99)	$2.4x10^{-3}$	1.23 (0.92-1.64)	NS
Control	595	177 (30)	299 (50)	119 (20)						
Male RA	108	31 (29)	56 (52)	21 (19)	NS	NS	0.82 (0.49-1.38)	NS	0.98 (0.54-1.77)	NS
Male control[d]	299	74 (25)	166 (55)	59 (20)						
Female RA	445	90 (20)	246 (55)	109 (25)	$3.3x10^{-4}$	NS	2.11 (1.49-2.98)	$9.6x10^{-6}$	1.28 (0.88-1.85)	NS
Female Control	296	103 (35)	133 (45)	60 (20)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test; d male control vs. female , p trend=NS

rs17066662 is a tagging SNP for rs10499194 at the *TNFIP3* region

4.2.4. Association of Interleukin-2 receptor beta and interleukin-2 receptor gamma

The genotype frequencies in rs2104286 in *IL-2RA* and rs743777 in *IL-2RB* are presented in Table 4-7. Genotypic frequencies of the two SNPs followed HWE ($p>0.05$). The GG minor allele frequency of rs2104286 in *IL-2RA* is similar in both RA patients and normal controls (p for trend > 0.05). Upon subgroup analysis, a small statistical association is found in a recessive genetic model. There is a small decrease in risk of RA for GG minor genotype in woman RA patients (OR = 0.42, 95 % CI: 0.17 - 0.99; p-value = 0.030).

No genetic association is found among RA patients and controls in rs743777 in *IL-2RB*. RA patients are sub-grouped according to sex and both dominant and recessive genetic models yielded negative results.

Table 4-7 The allele and genotype frequencies of the *IL-2RA*,and *IL-2RB* in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^C	OR (95% CI)	P^C
IL-2RA										
rs2104286		AA	GA	GG	*P trend*[a]	HWE[b]	GA + GG vs AA		GG vs. AG + AA	
Case	589	398 (68)	180 (30)	11 (2)	NS	NS	1.18 (0.94-1.49)	NS	0.49 (0.23-1.02)	0.039
Control	885	642 (71)	228 (26)	33 (4)						
Male RA	118	78 (66)	37 (31)	3 (3)	NS	NS	1.18 (0.73-1.92)	NS	0.78 (0.16-3.21)	NS
Male control	278	194 (70)	75 (27)	9 (3)						
Female RA	471	320 (68)	143 (30)	8 (2)	NS	NS	1.15 (0.88-1.50)	NS	0.42 (0.17-0.99)	0.030
Female Control	607	430 (71)	153 (25)	24 (4)						
IL-2RB										
rs743777		AA	GA	GG	*P trend*[a]	HWE[b]	AA vs. AG+GG		GG vs. AG+AA	
Case	592	440 (74)	138 (23)	14 (3)	NS	NS	1.05 (0.83-1.35)	NS	1.23(0.57-2.62)	NS
Control	932	702 (75)	212 (23)	18 (2)						
Male RA	119	90 (75)	27 (23)	2 (2)	NS	NS	0.93 (0.55-1.56)	NS	0.55 (0.08-2.78)	NS
Male control	299	222 (74)	68 (23)	9 (3)						
Female RA	473	350 (74)	111 (24)	12 (2)	NS	NS	1.10 (0.83-1.46)	NS	1.80 (0.70-4.69)	NS
Female Control	633	480 (76)	144 (23)	9 (1)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

4.2.5. Association of Cytotoxic T lymphatic Antigen 4

The genotype frequencies of rs3087243 and rs10497873 in *CTLA-4* are presented in Table 4-8. Genotypic frequencies of the two SNPs followed HWE ($p > 0.05$). The frequencies of the minor AA allele in rs307243 are similar to that of the controls (p for trend > 0.05). No statistical associations are found in both dominant and recessive genetic models.

The allele frequencies in rs10497873 are similar in both RA cases and normal controls (GG (67 %), GT (27 %), TT (6 %) in RA patients versus GG (64 %), GT (33%), TT (3 %) in controls, p for trend >0.05). Genotypic frequencies of the SNPs followed HWE ($p > 0.05$). A recessive genetic model shows the genotype TT increased the risk of RA (OR = 1.98, 95% CI: 1.14 - 3.44: p value = 0.009). Upon subgroup analysis by sex, it revealed that such association is only found in female RA patients (OR = 2.01, 95% CI: 1.03 - 3.97: p-value = 0.027).

Table 4-8 The allele and genotype frequencies of 2 SNPs in the *CTLA-4* gene of RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^C	OR (95% CI)	P^C
CTLA-4 rs3087243		GG	GA	AA	*P trend*[a]	HWE[b]	GG vs. GA + AA		AA vs. GA + GG	
Case	546	327 (60)	191 (35)	28 (5)	NS	NS	1.01 (0.79-1.30)	NS	1.10(0.61-1.96)	NS
Control	553	333 (60)	194 (35)	26 (5)						
Male RA	108	66 (61)	39 (36)	3 (3)	NS	NS	1.16 (0.71-1.88)	NS	0.55 (0.12-2.09)	NS
Male control	282	182 (65)	86 (30)	14 (5)						
Female RA	438	261 (60)	152 (35)	25 (6)	NS	NS	0.85 (0.62-1.17)	NS	1.31 (0.62-2.81)	NS
Female Control	271	151 (56)	108 (40)	12 (4)						
CTLA-4 rs10497873		GG	GT	TT			GT +TT vs GG		TT vs. GT + GG	
Case	574	386 (67)	156 (27)	32 (6)	NS	NS	0.88 (0.70-1.10)	NS	1.98 (1.14-3.44)	0.009
Control	932	599 (64)	306 (33)	27 (3)						
Male RA	116	74 (64)	34 (30)	8 (7)	NS	NS	1.10 (0.68-1.76)	NS	2.12 (0.74-6.00)	NS
Male control	296	195 (66)	91 (31)	10 (3)						
Female RA	458	312 (68)	122 (27)	24 (5)	NS	NS	0.81 (0.63-1.06)	NS	2.01 (1.03-3.97)	0.027
Female Control	636	404 (63)	215 (34)	17 (3)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

4.2.6. Association of Peptidylarginine Deiminase 4

The genotype frequencies for rs874881 in *PADI-4-92* and rs2240340 in *PADI-4-94* are presented in Table 4-9. Genotypic frequencies of the two SNPs followed HWE ($p > 0.05$). Risk for RA is not associated with polymorphisms in *PADI-4* (p for trend > 0.05). Upon subgroup analysis, the major GG genotype was significantly higher in the male RA cases then normal controls in rs2240340 (p for trend = 0.005). The frequency of the GG genotype in the male controls (19%) is lower than in female controls (35%) (p for trend = 3.4×10^{-5}).

In a dominant model, the A allele genotype of the rs2240340 revealed there is a 2-fold increase in the risk for RA in male subjects (OR = 2.16, 95 % CI: 1.28 – 3.66: p-value = 0.002).

Table 4-9 The allele and genotype frequencies of *PADI-4-92 and PADI-4-94* in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^C	OR (95% CI)	P^C
PADI-4-92 rs874881		CC	CG	GG	*P trend*[a]	HWE[b]	CG + GG vs CC		GG vs. CG + CC	
Case	531	161 (31)	273 (51)	97 (18)	NS	NS	1.23 (0.97-1.56)	NS	0.99 (0.75-1.31)	NS
Control	907	316 (35)	424 (47)	167 (18)						
Male RA	108	30 (28)	52 (48)	26 (24)	NS	NS	1.43 (0.86-2.38)	NS	1.16 (0.67-2.02)	NS
Male control	299	106 (35)	129 (43)	64 (21)						
Female RA	423	131 (31)	221 (52)	71 (17)	NS	NS	1.18 (0.89-1.55)	NS	0.99 (0.70-1.40)	NS
Female Control	608	210 (35)	295 (49)	103 (17)						
PADI-4-94 rs2240340		GG	GA	AA	*P trend*[a]	HWE[b]	GA + AA vs GG		AA vs. GA + GG	
Case	541	181 (34)	273 (50)	87 (16)	NS	NS	0.84 (0.66-1.06)	NS	0.81 (0.61-1.09)	NS
Control	896	266 (30)	459 (51)	171 (19)						
Male RA	107	36 (34)	54 (50)	17 (16)	0.005	NS	0.46 (0.27-0.78)	0.002	0.67 (0.36-1.24)	NS
Male control[d]	295	56 (19)	174 (59)	65 (22)						
Female RA	434	145 (33)	219 (51)	70 (16)	NS	NS	1.07 (0.82-1.40)	NS	0.90 (0.64-1.27)	NS
Female Control	601	210 (35)	285 (47)	106 (18)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test; d male control vs. female control P trend = 3.4×10^{-5}

4.2.7. Association of Granzyme

The genotype frequencies of rs854350 in GZEB are presented in Table 4-10. The allele frequencies in rs854350 in *GZEB* are similar in both RA cases and normal controls {GG (33%), GC (48%), CC (19%) in RA patients versus GG (34%), GC (46%), CC (20%) in controls, p for trend >0.05}. Both dominant and recessive genetic models yielded negative associations.

4.2.8. Association of Signal Transducer and Activator of Transcription 4

The genotype frequencies of rs7574865 in *STAT4* are presented in Table 4-10. Genotypic frequencies of the SNP followed HWE ($p > 0.05$). The TT minor allele frequency is significantly higher in the case patients then in the normal controls (19% versus 13%; p for trend = 3.9×10^{-9}). Upon subgroup analysis, these associations are retained in both RA men and women, (p for trend = 3.6×10^{-6} and 2.0×10^{-5}, respectively). In a dominant genetic model, the major T allele shows a two fold increase for RA risk (OR = 2.08, 95 % CI: 1.63 – 2.66; p-value < 1.0×10^{-7}). Subgroup analysis reveals that this effect is retained in both male and female RA patients (OR = 3.51, 95 % CI: 1.92 – 6.49; p = 8.2×10^{-6}, and OR = 1.87, 95% CI: 1.42 – 2.47; p = 3.7×10^{-6}, respectively). When analyzed in a recessive genetic model, the minor TT genotype increases the risk of RA (OR = 2.38, 95 % CI: 1.28 - 4.41; p = 2.8×10^{-3} in men and OR = 1.41, 95 % CI: 0.98 - 2.03; $p > 0.05$ in women).

Table 4-10 The allele and genotype frequencies of the *GZMB* and *STAT4* in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^C	OR (95% CI)	P^C
GZEB rs854350		CC	GC	GG	*P trend*[a]	HWE[b]	GC + GG vs CC		GG vs. GC + CC	
Case	558	185(33)	266(48)	107(19)	NS	NS	1.06 (0.84-1.34)	NS	0.95 (0.72-1.25)	NS
Control	876	302(34)	399(46)	175(20)						
Male RA	111	33(30)	56(50)	22(20)	NS	NS	1.20 (0.73-1.99)	NS	1.04 (0.58-1.88)	NS
Male control	282	95(34)	133(47)	54(19)						
Female RA	447	152(34)	210(47)	85(19)	NS	NS	1.04 (0.80-1.36)	NS	0.92 (0.67-1.27)	NS
Female Control	594	207(35)	266(45)	121(20)						
STAT4 r7574865		GG	GT	TT			GT +TT vs GG		TT vs. GT + GG	
Case	509	135 (27)	279 (55)	95 (19)	3.9×10^{-9}	NS	2.08 (1.63-2.661)	$<1.0\times10^{-7}$	1.58 (1.16-2.15)	2.3×10^{-3}
Control	900	386 (43)	400 (44)	114 (13)						
Male RA	96	17 (18)	55 (57)	24 (25)	3.6×10^{-6}	NS	3.51 (1.92-6.49)	8.2×10^{-6}	2.38 (1.28-4.41)	2.8×10^{-3}
Male control	293	126 (43)	131 (45)	36 (12)						
Female RA	413	118 (29)	224 (54)	71 (17)	2.0×10^{-5}	NS	1.87 (1.42-2.47)	3.7×10^{-6}	1.41 (0.98-2.03)	NS
Female Control	607	260 (40)	269 (41)	78 (19)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

4.2.9. Association of Protein kinase C theta

The genotype frequencies in rs4750316 are presented in Table 4-11. Genotypic frequencies of the SNP followed HWE ($p > 0.05$). The CC minor allele frequency is similar in both case patients and in the normal controls (p for trend > 0.05). No associations are found in both dominant and recessive genetic models. Upon subgroup analysis, an increase risk for RA in men is noted, but it is statistically insignificant (OR = 2.68, 95 % CI: 0.43 - 16.92; p-value = NS).

4.2.10. Association of Kazal type serine protease inhibitor domain

The genotype frequencies for SNP rs10786617 are presented in Table 4-11. Genotypic frequencies of the SNP follow HWE ($p > 0.05$). The TT minor allele frequency is similar in both case patients and in the normal controls (p for trend $>$ 0.05). No associations are found in both dominant and recessive genetic models.

Table 4-11 The allele and genotype frequencies of *PRKCQ* and *KAZALD1* genes in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^c	OR (95% CI)	P^c
PRKCQ										
rs4750316		GG	GC	CC	*P trend*[a]	HWE[b]	GC +CC vs GG		CC vs. GC + GG	
Case	565	453 (80)	105 (19)	7 (1)	NS	NS	1.18 (0.89-1.56)	NS	0.83 (0.30-2.23)	NS
Control	869	718 (83)	138 (16)	13 (1)						
Male RA	114	86 (75)	25 (22)	3 (3)	NS	NS	1.36 (0.79-2.35)	NS	2.68 (0.43-16.92)	NS
Male control	301	243 (81)	55 (18)	3 (1)						
Female RA	451	367 (81)	80 (18)	4 (1)	NS	NS	1.17 (0.83-1.64)	NS	0.50 (0.13-1.74)	NS
Female Control	568	475 (84)	83 (15)	10 (2)						
KAZALD1										
rs10786617		CC	CT	TT			CT + TT vs CC		TT vs. CT + CC	
Case	567	226 (40)	255 (45)	86 (15)	NS	NS	0.95 (0.76-1.18)	NS	0.98 (0.73-1.33)	NS
Control	903	348 (39)	416 (46)	139 (15)						
Male RA	115	37 (38)	57 (48)	21 (14)	NS	NS	1.30 (0.80-2.11)	NS	1.35 (0.73-2.48)	NS
Male control	288	110 (32)	137 (50)	41 (18)						
Female RA	452	189 (42)	198 (44)	65 (14)	NS	NS	0.88 (0.68-1.13)	NS	0.89 (0.62-1.26)	NS
Female Control	615	238 (39)	279 (45)	98 (16)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

4.2.11. Association of Chemokine (C-C motif) Receptor 6

The genotype frequencies for SNP rs3093024 are shown in Table 4-12. Genotypic frequencies of the SNPs followed HWE ($p > 0.05$). There are no statistically significant difference between RA patients and normal controls. No associations are found in the dominant and recessive genetic models.

4.2.12. Association of T-Cell Receptor alpha

The genotype frequencies for SNP rs1154155 are shown in Table 4-12. Genotypic frequencies of the SNPs followed HWE ($p > 0.05$). The minor TT genotype frequency is similar in both RA patients and normal (p for trend > 0.05). There are no statistically significant differences between RA patients and normal controls. No associations are found in the dominant and recessive genetic models.

Table 4-12 The allele and genotype frequencies of *CCR6* and *TRA alpha* in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^C	OR (95% CI)	P^C
CCR6										
rs3093024		GG	GA	AA	*P trend*[a]	HWE[b]	GA+AA vs GG		AA vs. GA+GG	
Case	571	156 (27)	285 (50)	130 (23)	NS	NS	1.13 (0.89-1.44)	NS	1.24 (0.96-1.62)	NS
Control	924	276 (30)	471 (51)	177 (19)						
Male RA	116	31 (28)	53 (48)	27 (24)	NS	NS	0.96 (0.58-1.62)	NS	1.34 (0.77-2.32)	NS
Male control	294	80 (27)	157 (53)	57 (20)						
Female RA	460	125 (27)	232 (50)	103 (23)	NS	NS	1.21 (0.92-1.59)	NS	1.23 (0.90-1.67)	NS
Female Control	630	196 (31)	314 (50)	120 (19)						
TRA alpha										
rs1154155		GG	GT	TT	*P trend*[a]	HWE[b]	GT +TT vs GG		TT vs. GT + GG	
Case	577	177 (31)	297 (51)	103(18)	NS	NS	1.01 (0.80-1.27)	NS	0.92 (0.71-1.22)	NS
Control	904	279 (31)	453 (50)	172 (19)						
Male RA	116	36 (31)	61 (53)	19 (16)	NS	NS	0.73 (0.44-1,21)	NS	0.80 (0.43-1.45)	NS
Male control	299	74 (25)	166 (55)	59 (20)						
Female RA	461	141 (31)	236 (51)	84 (18)	NS	NS	1.16 (0.89-1.52)	NS	0.97 (0.70-1.34)	NS
Female Control	605	205 (34)	287 (47)	113 (19)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

4.2.13. Association of Anaphase-promoting complex 4

The genotype frequencies for SNP rs3816587 in *ANAPC4* are presented in Table 4-13. Genotypic frequencies of the SNPs followed HWE ($p > 0.05$). The minor TT genotype frequency is similar in both RA patients and in normal controls (p for trend > 0.05). No associations are found even after subgroup analysis by sex. In a dominant genetic model it shows that the T allele increase the risk for RA (OR = 1.32, 95 % CI: 1.06 - 1.65; $p = 0.012$), but in the recessive genetic model, the TT genotype also reduced the risk for RA (OR = 0.65, 95% CI: 0.43 - 0.96; $p = 0.029$).

Table 4-13 The allele and genotype frequencies of *ANAPC4* in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^C	OR (95% CI)	P^C
ANAPC4 rs3816587		CC	CT	TT	*P trend*[a]	HWE[b]	CT+TT vs CC		TT vs. CT + CC	
Case	589	251 (43)	297 (50)	41 (7)	NS	NS	1.32 (1.06-1.65)	0.012	0.65 (0.43-0.97)	0.028
Control	741	367 (50)	297 (40)	77 (10)						
Male RA	117	45 (39)	68 (58)	4 (3)	NS	NS	1.68 (1.06-2.68)	0.020	0.29 (0.09-0.90)	0.017
Male control	279	143 (51)	106 (38)	30 (11)						
Female RA	472	206 (44)	229 (48)	37 (8)	NS	NS	1.22 (0.93-1.59)	NS	0.75 (0.48-1.21)	NS
Female Control	462	224 (49)	191 (41)	47 (10)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

4.2.14. Association of SNPs rs11162922 and rs1937506

The genotype frequencies for SNPs rs11162922 and rs1937506 are shown in Table 4-14. Genotypic frequencies of the SNPs followed HWE ($p > 0.05$). There are no statistically significant differences between RA patients and normal controls. No associations are found in the dominant and recessive genetic models.

Table 4-14 The allele and genotype frequencies of rs11162922 and rs1937506 in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^C	OR (95% CI)	P^C
rs11162922		AA	GA	GG	*P trend*[a]	HWE[b]	GA + GG vs AA		GG vs. GA + AA	
Case	568	295 (52)	231 (41)	42 (7)	NS	NS	0.92 (0.74-1.14)	NS	0.93 (0.61-1.41)	NS
Control	900	449 (50)	380 (42)	71 (8)						
Male RA	114	58 (51)	48 (42)	8 (7)	NS	NS	0.93 (0.59-1.47)	NS	0.83 (0.33-2.04)	NS
Male control	277	136 (49)	118 (43)	23 (8)						
Female RA	454	237 (52)	183 (40)	34 (8)	NS	NS	0.92 (0.72-1.19)	NS	0.97 (0.61-1.57)	NS
Female Control	623	313 (50)	262 (42)	48 (8)						
rs1937506		GG	GA	AA			AG+AA vs GG		AA vs. AG + GG	
Case	588	436 (74)	143 (24)	9 (2)	NS	NS	1.05 (0.82-1.34)	NS	0.94 (0.38-2.30)	NS
Control	925	694 (75)	216 (23)	15 (2)						
Male RA	118	92 (78)	24 (20)	2 (2)	NS	NS	0.80 (0.47-1.37)	NS	0.99 (0.13-.5.84)	NS
Male control	292	216 (74)	71 (24)	5 (2)						
Female RA	470	344 (73)	119 (25)	7 (1)	NS	NS	1.13 (0.85-1.50)	NS	0.94 (0.32-2.70)	NS
Female Control	633	478 (76)	145 (23)	10 (2)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

4.2.15. Association of SNP rs11761231

The genotype frequencies of rs11761231 are shown in Table 4-15. Genotypic frequencies of the SNP followed HWE ($p > 0.05$). There is a small decrease in the major TT genotype among RA patients from 78 % in controls to 72 % in RA subjects (p for trend = 8.5×10^{-3}). Subgroup analysis showed this association was only found among female RA patients (83 % versus 73 %; control vs. RA, respectively; p for trend = 1.7×10^{-4}). In a dominant genetic model that the C allele increases risk for RA among female patients (OR = 1.93, 95 % CI: 1.40 – 2.66; $p = 2.4 \times 10^{-5}$).

Table 4-15 The allele and genotype frequencies of rs11761231 in RA patients and controls

SNP	n	Genotype frequency, n (%)					Dominant Model		Recessive model	
							OR (95% CI)	P^C	OR (95% CI)	P^C
rs11761231		TT	TC	CC			TC + CC vs TT		CC vs. CT + TT	
Case	567	407 (72)	140 (25)	20 (3)	$8.5x10^{-3}$	NS	1.43 (1.11-1.86)	0.004	1.26 (0.67-2.43)	NS
Control	781	613 (78)	146 (19)	22 (3)						
Male RA	112	83 (74)	26 (23)	3 (3)	NS	NS	0.86 (0.51-1.45)	NS	0.95 (0.20-4.02)	NS
Male control	284	202 (71)	74 (26)	8 (3)						
Female RA	455	324 (73)	114 (23)	17 (4)	$1.7x10^{-4}$	NS	1.93 (1.40-2.66)	$2.4x10^{-5}$	1.34 (0.62-2.91)	NS
Female Control	497	411 (83)	72 (14)	14 (3)						

SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test; b, Hardy-Weinberg equilibrium; c, uncorrected chi-square test

4.2.16. Association of Protein Tyrosine Phosphatase, non-receptor type 22

Two sets of primers are used to determine the role of rs2476601 (*PTPN22* gene) in Chinese RA patients, the forward primer set with A or G at the 3 prime ends produced no PCR products. The primers set with C or T at their 3 prime position showed a single melting peak at the lower temperature which was assigned as T. These results suggested that the SNP rs2476601 was not polymorphic in Chinese and the allele for this SNP in Chinese was a "T" allele.

4.2.17 Logistic regression analysis

Logistic regression was performed to study the relative importance of of the following factors in determining RA: age, sex and the 24 SNPs.

The results of multiple regression, regression coefficients, SE, significant and OR are shown in tables 4-16.

RA disease was used as the independent variable, with age, sex, and the 24 SNPs as the dependent variables.

Age, sex, T allele in *STAT4* and C allele *ANAP4* were significantly negatively associated with RA, and CC genotype in *HLA-DQA1* positively associated with RA.

Table 4-16 Variables predicting RA in logistic regression analysis

Variables	beta	SE	Sig	OR
Age	-0.216	0.049	0.000	0.805
Sex (male/ female)	-34.381	2506.129	0.989	0.000
rs6457617 (*HLA-DQA1*) (CC vs CG + GG	1.841	0.771	0.017	6.301
rs7574865 (*STAT4*) (GG vs GT + TT))	-2.029	0.864	0.019	0.131
rs3816587 (*ANAP4*) (TT vs CT + CC)	-2.620	0.792	0.001	0.073
Constant	50.113	2506.132	0.984	0.000

4.3. Gene to gene interaction

We found that *DRB1* allele interacted with *DQA1* ($p = 0.0075$), *PADI-4-92* ($p=0.0072$), *PADI-4-94* ($p = 0.0094$) and *IL2RB* ($p = 4.22 \times 10^{-5}$) in predisposition to RA Table 4-17.

In addition to the 3 interactions by *HLA-DRB* loci, three other interactions are also established. They are *SLC22A4* and *ANAPC4* ($p = 0.0039$), *IL13* and *CTLA-4* ($p = 0.0036$), *and MMEL1-TNFRSF14* and *TRA-alpha* ($p = 7.16 \times 10^{-39}$). The 3 x 3 gene-gene interaction data are presented in Figures 4-6 to 4-13.

Table 4-17 The gene to gene interaction in Chinese with rheumatoid arthritis

SNP1	ID	SNP2	ID	OR_INT	STAT	p-value
rs615672	*MHC-HLA DRB*	rs6457617	*MHC-HLA-DQA*	0.711	7.1	0.0075
rs615672	*MHC-HLA DRB*	rs743777	*IL2RB*	0.469	16.8	4.22×10^{-5}
rs615672	*MHC-HLA DRB*	rs874881	*PADI-4-92*	0.725	7.2	0.0072
rs615672	*MHC-HLA DRB*	rs2240340	*PADI-4-94*	0.724	6.8	0.0094
rs615672	*MHC-HLA DRB*	rs11162922	Gene desert	0.700	7.2	0.0072
rs2073838	*SLC22A4*	rs3816587	*ANAPC4*	1.53	8.3	0.0039
rs1800925	*IL13*	rs3087243	*CTLA-4*	1.833	8.5	0.0036
rs6684865	*MMEL1_TNFRSF14*	rs1154155	*TRA-alpha*	0.1038	170.1	7.16×10^{-39}

p-value = logistic regression model

Figure 4-6 The gene-gene interaction between *HLA-DRB1*and *HLA-DQA*

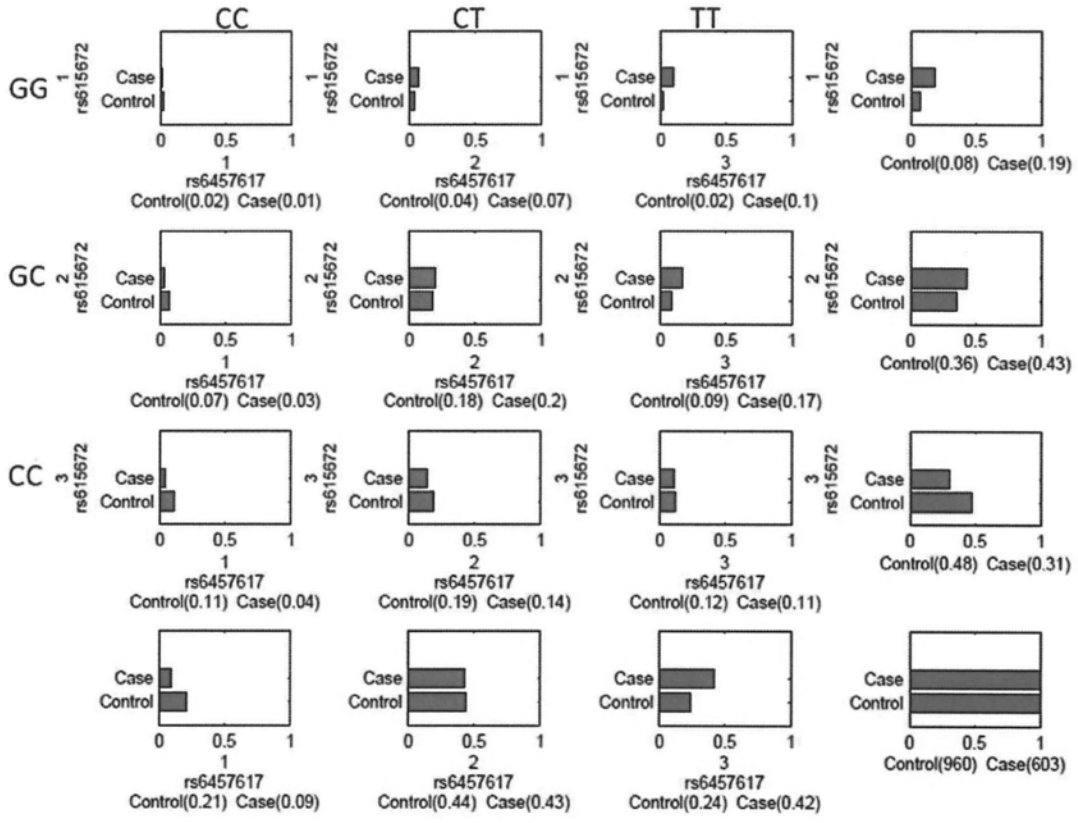


Figure 4-7 The gene-gene interaction between *HLA-DRB1*and *IL-2RB*

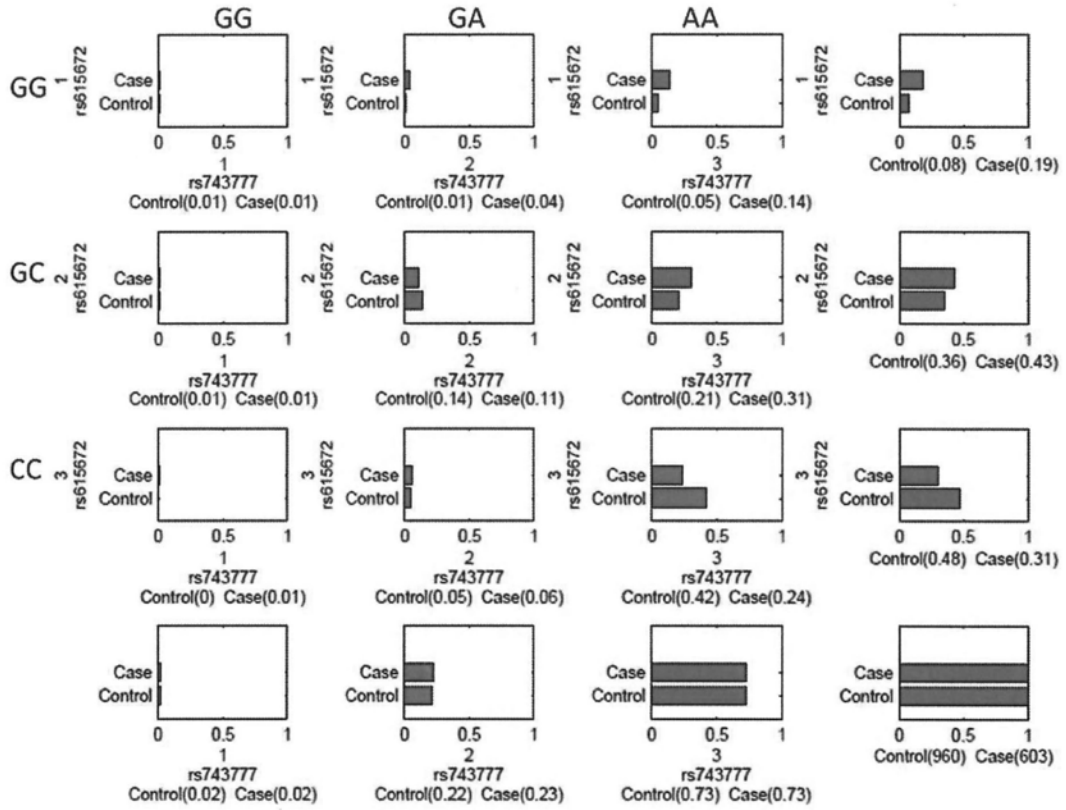


Figure 4-8 The gene-gene interaction between *HLA-DRB1*and *PADI-4-92*

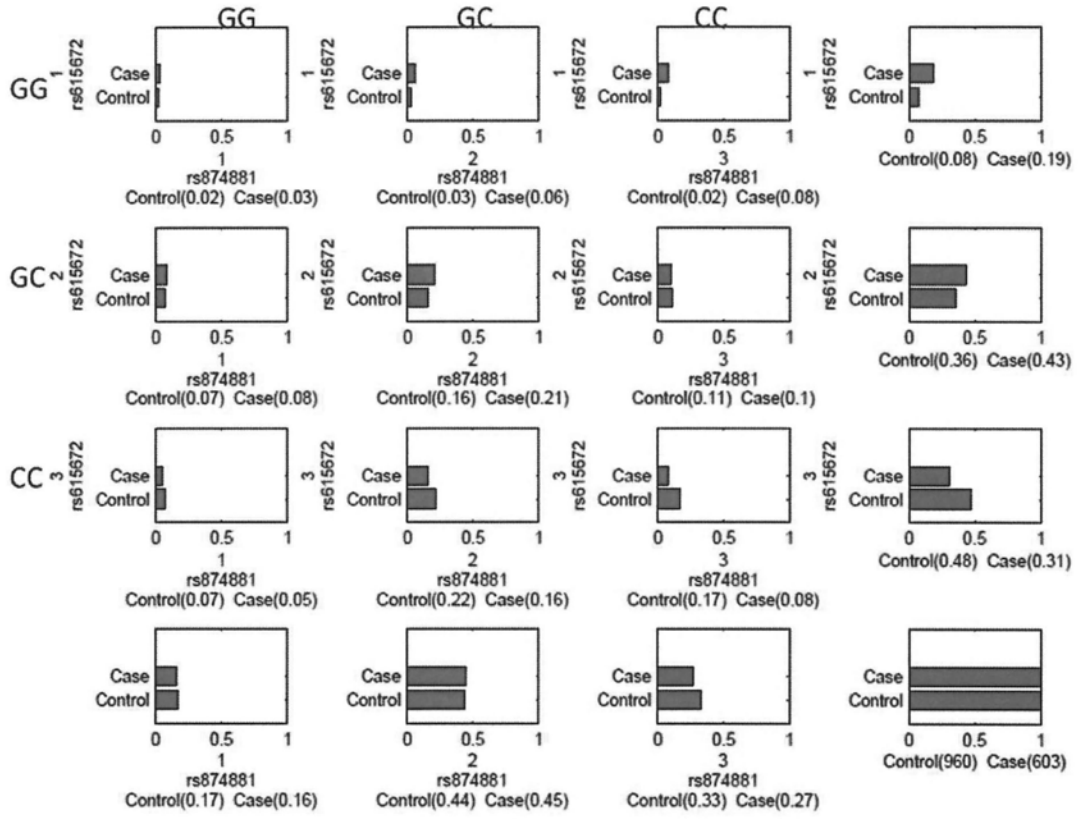


Figure 4-9 The gene-gene interaction between *HLA-DRB1*and *PAD-I-4-94*

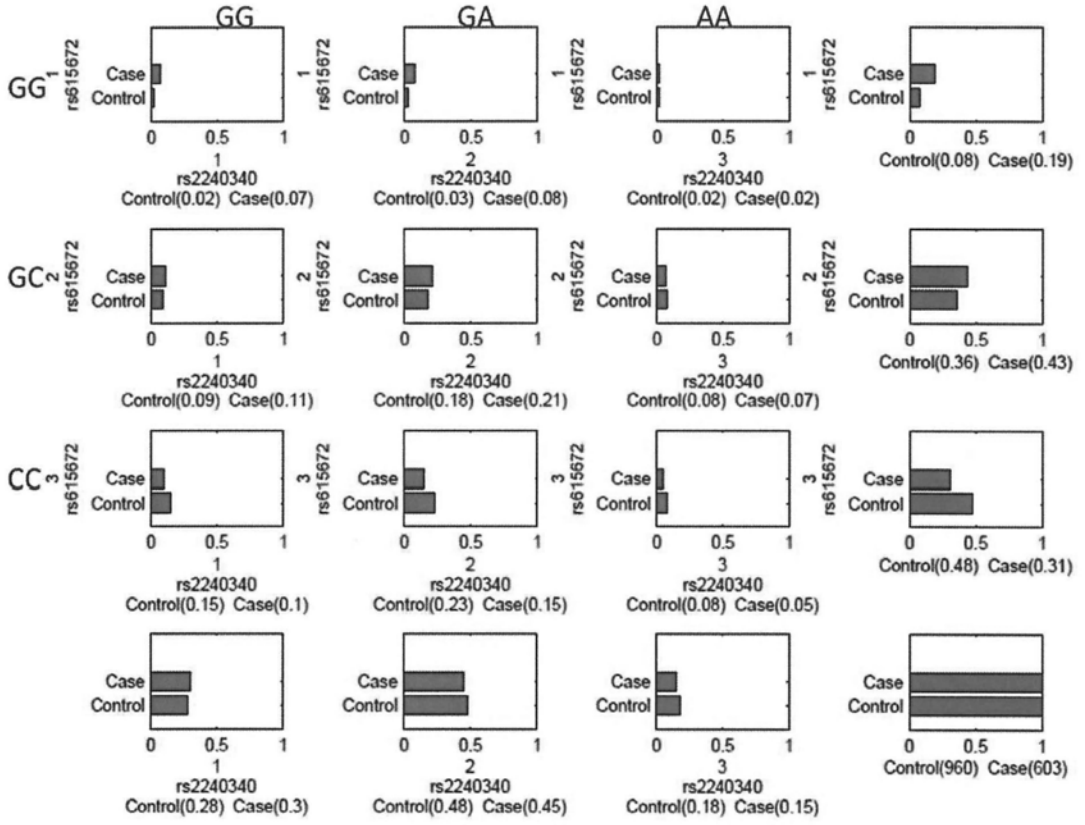


Figure 4-10 The gene-gene interaction between *HLA-DRB1*and SNP rs1162922

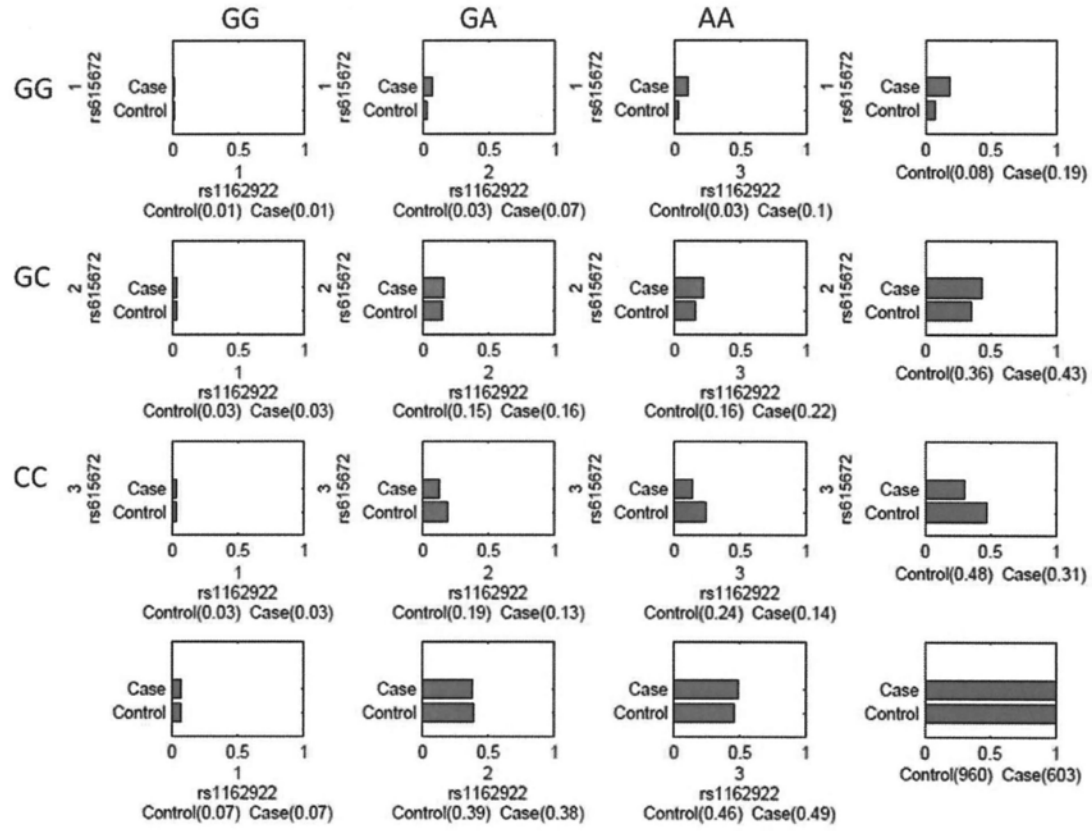


Figure 4-11 The gene-gene interaction between *SLC22A4* and *ANAPC4*

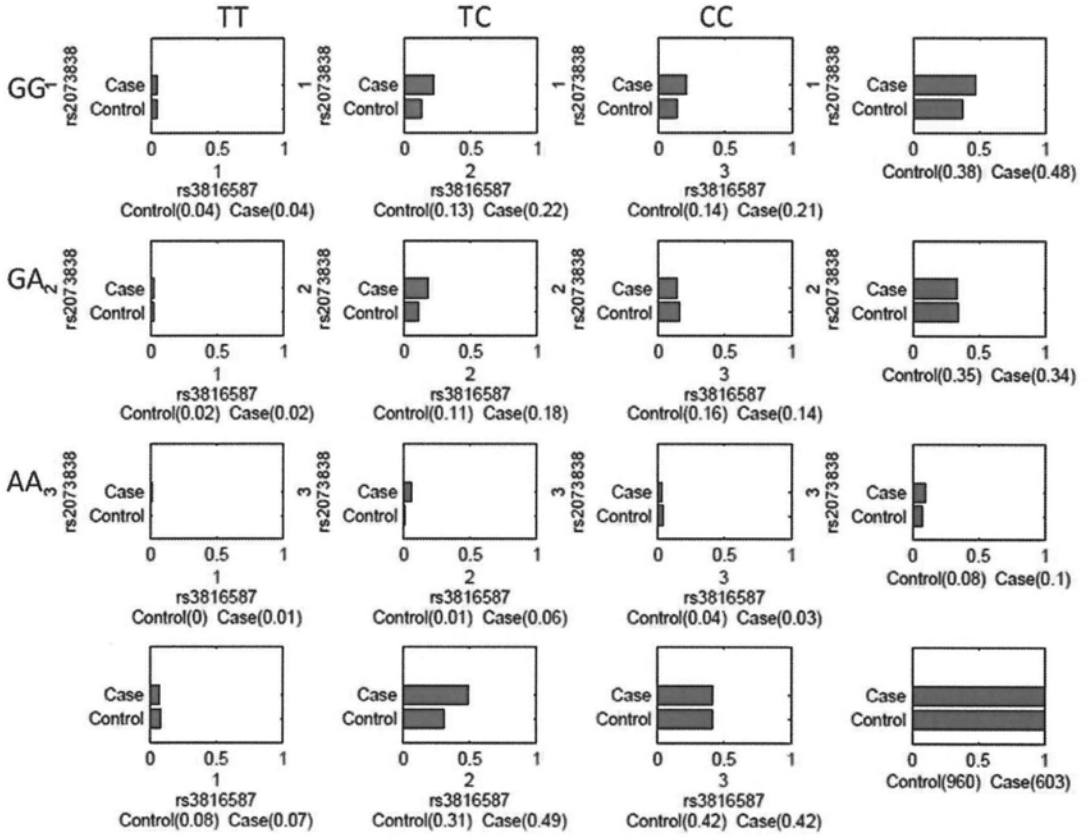


Figure 4-12 The gene-gene interaction between *IL-13* and *CTLA-4*

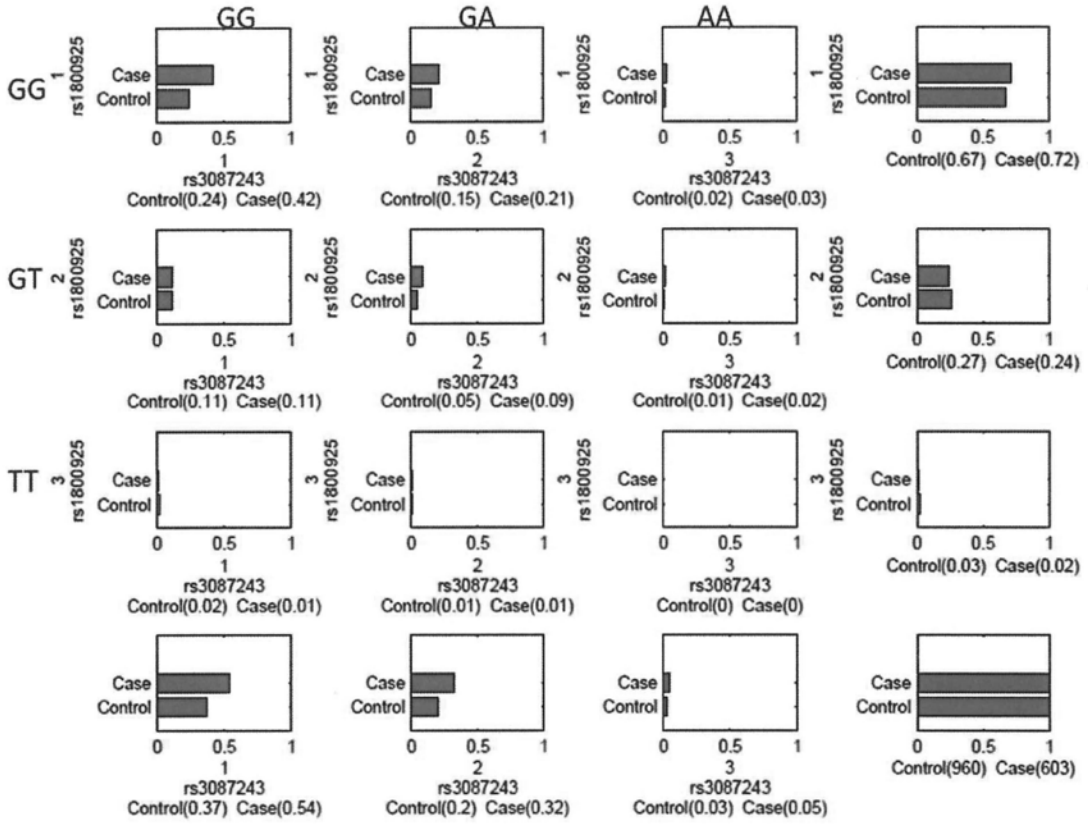
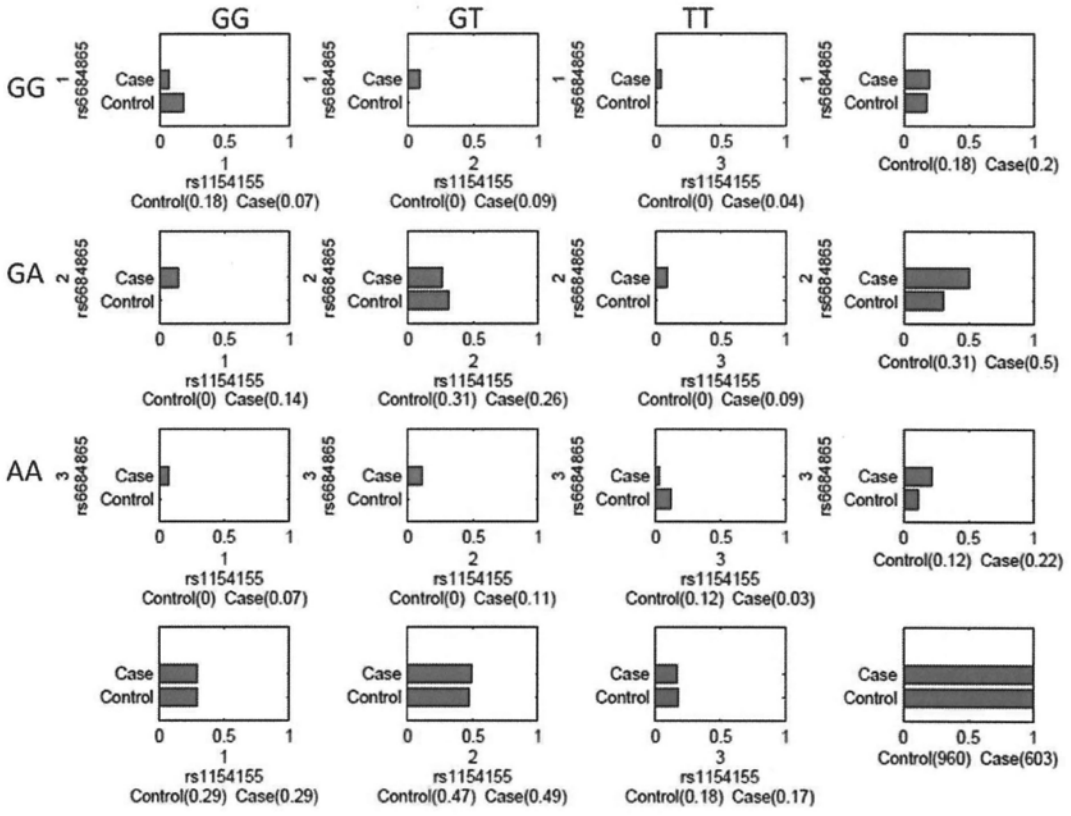


Figure 4-13 The gene-gene interaction of *MMEL1-TNFRSF14* and *TRA-alpha*



4.4. Association of serum biomarkers and rheumatoid arthritis

4.4.1. Bone erosion in RA patients with and without anti-citrullinated protein antibodies positive antibody

There was no association between bone erosion and anti-CCP status in RA patients (Table 4-18, p=NS). Upon sub-group analysis, a statistical difference was found among in female bone eosion + and female bone erosion – group (p=0.0027). There were no statistical differences in male bone erosion + and female bone erosion + groups (p=ns), or male bone erosion – and female boner erosion – groups (p=NS).

Table 4-18 Demographic data on bone erosion in rheumatoid arthritis patients with and without anti-citrullinated protein antibodies

	anti-CCP +	anti-CCP-	p-value
N, number, (%)	366 (82)	80 (18)	
Age, mean, SD (years)	55.8 ± 12.48	54.7 ± 15.25	NS
Sex, women (%)	79.9	82.5	NS
Age of onset (years)	46.4 ± 15.9	45.6 ± 17.8	NS
Disease duration (years)	9.4 ± 8.7	9.1 ± 8.4	NS
Bone erosion on X-ray			
All bone erosion (+)	250 (68)	42 (53)	NS
All bone erosion (-)	116 (32)	38 (47)	
Male bone erosion (+)[a]	51 (65)	9 (69)	NS
Male bone erosion (-)[b]	27 (35)	4 (31)	
Female bone erosion (+)	199 (69)	33 (49)	0.0027
Female bone erosion (-)	89 (31)	34 (51)	

Uncorrected chi-square test, a: male bone erosion + vs. female bone erosion p = NS, b: male bone erosion – vs. female bone erosion – p = NS

4.4.2.1. Genetic association in RA patients with and without bone erosion

The -log p-values for the genetic association in the 24 studied SNPs in the two groups of RA patients with and without bone erosion is presented in Figure 4-14.

Bone erosion was associated with the GG-allele in *HLA-DRB1* loci in female RA patients ($p = 0.037$).

4.4.2.2. Genetic association in RA patients with bone erosion and normal controls

The -log p-values for the genetic association in RA patients with bone erosion and normal controls are presented in Figure 4-15. Risk of RA is associated with polymorphisms in *STAT4, HLA-DRB1, HLA-DQA1and GZEB* loci ($p > 0.05$). Upon subgroup analysis risk of bone erosion associated with polymorphism in *MMEL1-TNFRSF14*, *TNFA1P3,* rs111761231 and *TNFA1P2* are associated in female RA patients and *PADI-4-94* is associated in male RA patients.

4.4.2.3. Genetic association in RA patients without bone erosion and normal controls

The -log p-value for the genetic association of RA patients without bone erosion and normal controls is presented in Figure 4-16. Risk of bone erosion is associated with polymorphism in *STAT4, and HLA-DQA1* loci ($p > 0.05$). Upon subgroup analysis *MMEL1-TNFRSF14*, rs17066662, rs111761231 and *TNFA1P2* are associated with risk of bone erosion in female RA patients and *HLA-DRB1*and *TNFA1P3* are associated with risk of bone erosion in male RA patients.

Figure 4-14 Genetic association in two RA patients groups with and without bone erosion

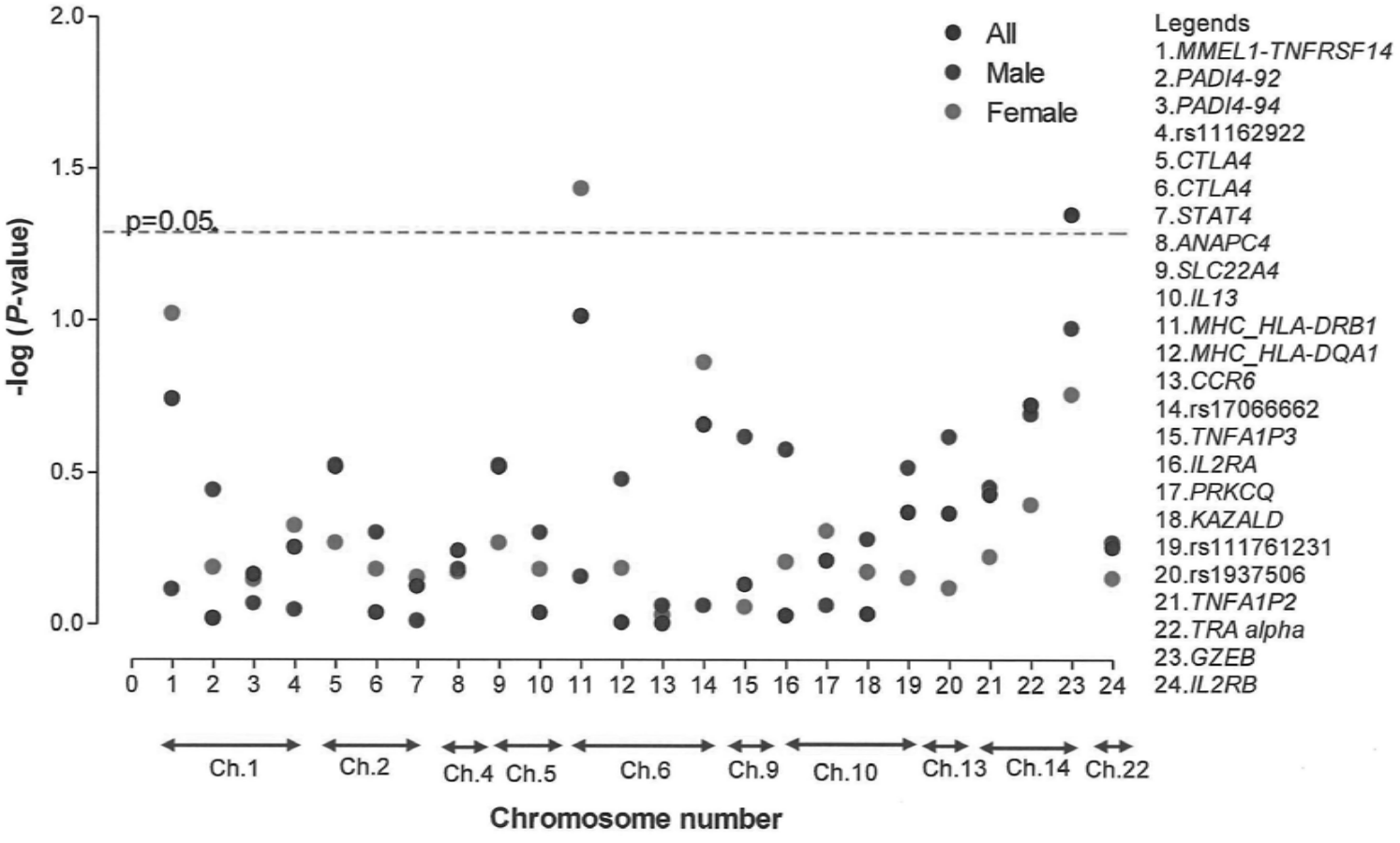


Figure 4-15 Genetic association in RA patients with bone erosion and normal controls

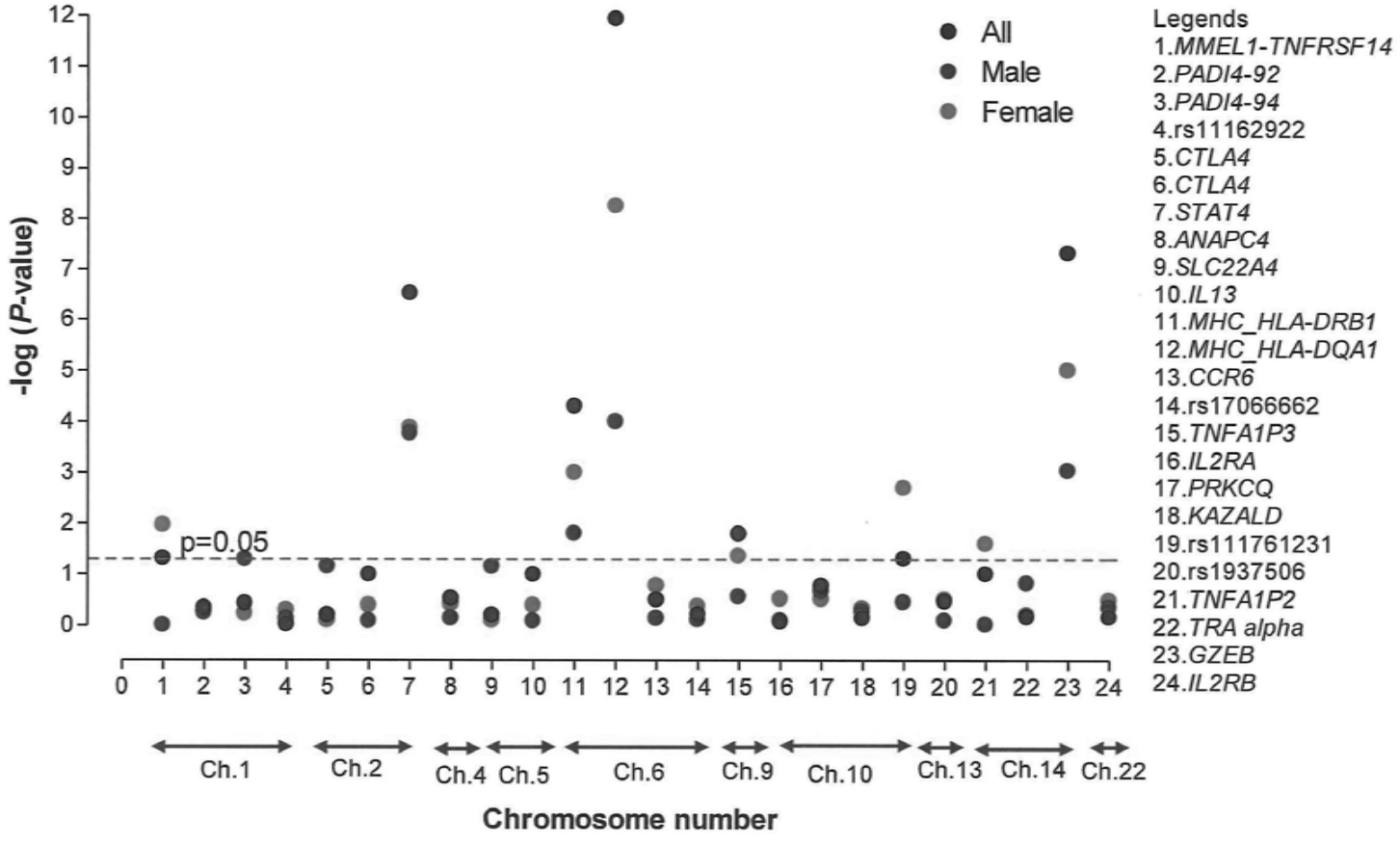
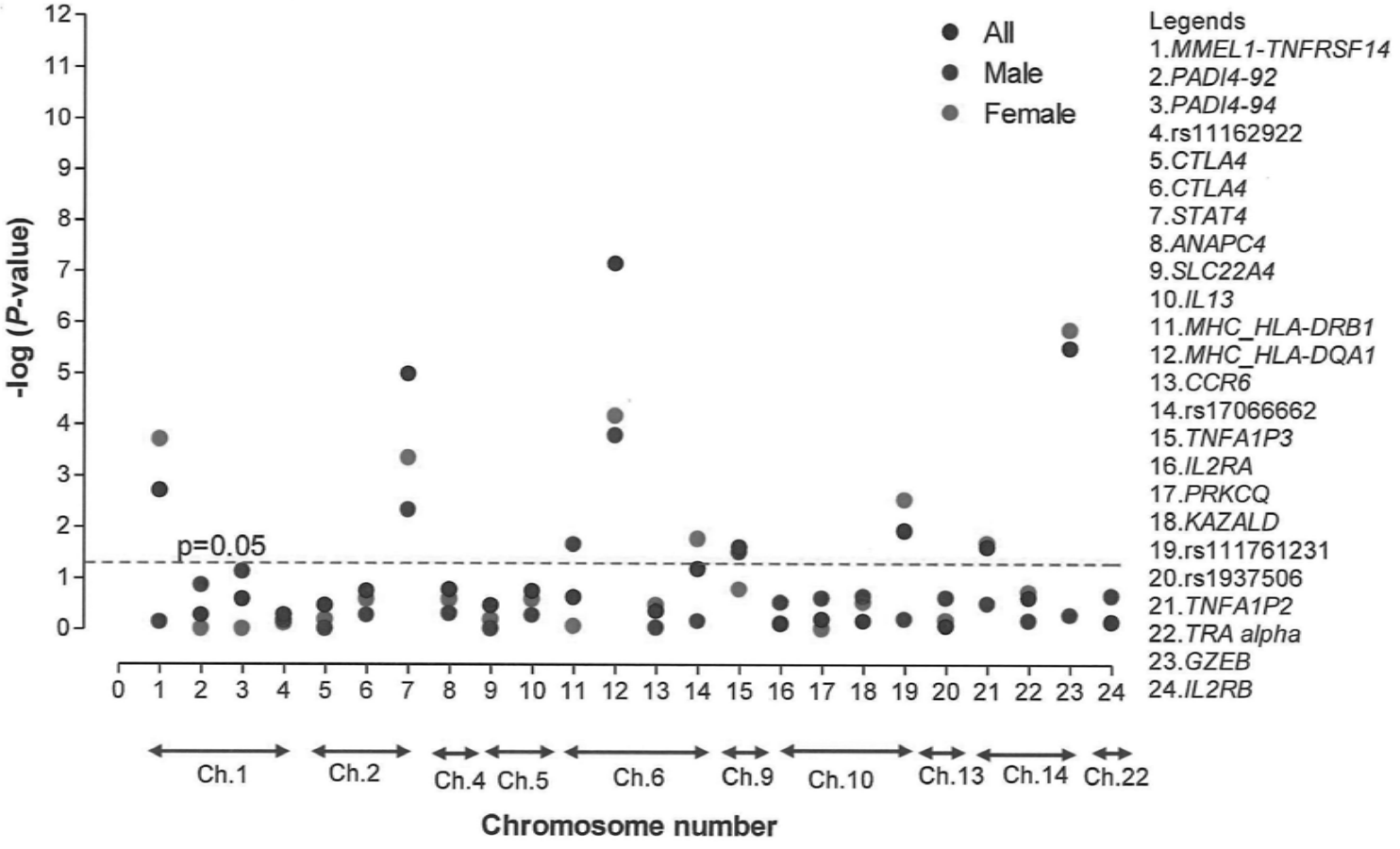


Figure 4-16 Genetic association in RA patients without bone erosion and normal controls



4.4.3. Genetic association of bone erosion and SNP10786617 Kazal type serine protease inhibitor domain

The interactions of bone erosion and SNP rs10786617, *KAZALDI* are studied. The results are presented in Table 4-19. *KAZALDI* (rs10786617) is not associated with bone erosions, There are no statistical differences in genotype frequencies between bone erosion and non-bone erosion subgroups (P for trend > 0.05) among male or female RA patients. There are no associations between RA patients with bone erosion to control subjects or RA patients with non bone erosion to control subjects (p for trend > 0.05).

Table 4-19 Bone erosion and *KAZALDI* genotype in Chinese RA men and women and controls

		KAZALDI, genotype, n (%)				
	n	CC	CT	TT	*P* trend[a]	*P* trend[b]
Bone erosion (+)	306	119 (38.9)	144 (47.1)	43 (14.0)	NS	NS
Bone erosion (-)	161	69 (42.9)	64 (39.8)	28 (17.3)		NS
Controls	903	348 (38.5)	416 (46.1)	139 (15.4)		
Male bone erosion (+)	66	22 (33.3)	34 (51.5)	10 (15.2)	NS	NS
Male bone erosion (-)	34	11 (32.4)	15 (44.1)	8 (23.5)		NS
Male controls	288	110 (38.2)	137 (47.6)	41 (14.2)		
Female bone erosion(+)	240	97 (40.4)	110 (45.8)	33 (13.8)	NS	NS
Female bone erosion(-)	127	58 (45.7)	49 (38.6)	20 (15.7)		NS
Female controls	497	411 (82.7)	72 (14.5)	14 (2.8)		

a: Cochran-Armitage trend test between the two patient group with bone erosion and without bone erosion

b: Cochran-Armitage trend test comparing to controls

4.4.4. Genetic association of bone erosion and *HLA-DRB1*

Table 4-20 shows the genotype frequencies of *HLA-DRB1*. Genotypic frequencies of the SNP followed HWE ($p > 0.05$). The minor GG genotype frequency is similar in both groups of RA patients with and without bone erosion (p for trend > 0.05). However, upon subgroup analysis, the minor GG genotype frequency is significant higher in female RA patients with bone erosion then those without bone erosion (p for trend = 0.037). The minor GG genotype is significantly higher in the RA patients with bone erosion then normal controls (p for trend = 4.98×10^{-5}). Upon subgroup analysis, there are statistical differences between RA patients with bone erosion and normal controls in men (p for trend = 0.016) and women (p for trend = 0.001). No statistical association is found between the non-bone erosion female RA patients and normal controls (p for trend > 0.05), and a small statistical association is noted in RA men with bone erosion and normal controls (p for trend = 0.022).

In the dominant genetic model, the G allele is a risk allele for bone erosion compared to normal controls (OR = 1.48, 95 % CI: 1.11 – 1.98; p-value = 0.005). Subgroup analysis shows this effect only exists in female RA patients with bone erosion when compared to normal controls (OR = 1.45, 95% CI: 1.04 – 2.06; p-value = 0.021).

In the recessive genetic model, the GG genotype shows a two-fold increase in risk for bone erosion in RA patients and normal controls (OR = 1.99, 95% CI: 1.40-2.82; $p = 5.78 \times 10^{-5}$). Upon subgroup analysis, this risk is retained in both male RA bone erosion patients (OR = 2.28, 95 % CI: 1.08 - 4.77; p-value = 0.0168) and female RA bone erosion patients (OR = 1.88, 95 % CI: 1.25 - 2.83; p-value = 0.001).

No association is found among non bone erosion RA patients to controls.

Table 4-20 Bone erosion and *HLA DRB1* genotype in Chinese RA men and women and controls

		Genotype, n (%)					Dominant model			Recessive model		
							OR (95%CI)			OR (95% CI)		
HLA-DRB1	n	CC	CG	GG	*P*trend[a]	*P*trend[b]	CG+GG vs CC	*P*[c]	*P*[c]	GG vs. CG+CC	*P*[c]	*P*[c]
Bone erosion (+)	309	95 (23)	147 (47)	67 (22)	NS		1.35 (0.89-2.06)	NS		1.36 (0.81-2.31)	NS	
Bone erosion (-)	160	60 (37)	73 (46)	27 (17)		4.98 x 10^{-5}	1.48 (1.11-1.98)		0.005	1.99 (1.40-2.82)		5.78 x 10^{-5}
Controls	866	344 (40)	416 (48)	106 (12)		NS	1.10 (0.76-1.58)		NS	1.46 (0.89-2.36)		NS
Male bone erosion (+)	67	19 (28)	33 (49)	15 (23)	NS		1.05 (0.35-2.86)	NS		0.69 (0.25-1.95)	NS	
Male bone erosion (-)	34	10 (29)	14 (41)	10 (30)		0.016	1.64 (0.88-3.07)		NS	2.28 (1.08-4.77)		0.0168
Male controls	267	105 (39)	132 (50)	30 (11		0.022	1.56 (0.68-3.65)		NS	3.29 (1.32-8.09)		0.003
Female bone erosion(+)	242	76 (31)	114 (47)	52 (22)	0.037		1.44 (0.90-2.31)	NS		1.75 (0.93-3.33)	NS	
Female bone erosion (-)	126	50 (40)	59 (47)	17 (13)		0.001	1.45 (1.04-2.00)		0.021	1.88 (1.25-2.83)		0.0012
Female controls	599	239 (40)	284 (47)	76 (13)		NS	1.01 (0.67-1.52)		NS	1.07 (0.59-1.95)		NS

OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test between the two patient group with bone erosion and without bone erosion

b: Cochran-Armitage trend test comparing to controls

c: Uncorrected chi-square test

4.4.5. Genetic association of bone erosion and *HLA-DQA1*

Table 4-21 shows the genotype frequencies of *HLA-DQA1*. Genotypic frequencies of the SNP follow HWE ($p > 0.05$). The minor TT genotype frequency in *HLA-DQA1* loci is similar in both RA patient groups that with and without bone erosion (p for trend > 0.05). The minor TT genotype frequency was higher in RA patients with or without bone erosion then normal controls (p for trend = 1.15×10^{-12}, and 7.35×10^{-8}, respectively).

No statistical associations are found in either dominant or recessive genetic models between bone erosion RA and normal controls. *HLA-DQA1* did not contribute to bone erosion in RA patients.

Table 4-21 Bone erosion and *MHC-HLA DQA1* genotype in Chinese RA men and women and controls

		Genotype, n (%)					Dominant model OR (95%CI)			Recessive model OR (95% CI)		
HLA-DQA1	n	CC	TC	TT	*P*trend[a]	*P*trend[b]	CT+TT vs CC	P^c	P^c	TT vs. CT+CC	P^C	P^C
Bone erosion (+)	315	26 (8)	151 (48)	138 (43)	NS		1.41 (0.71-2.77)	NS		0.88(0.59-1.32)	NS	
Bone erosion (-)	160	18 (11)	67 (41)	75 (50)		$1.15x10^{-12}$	3.58 (1.48-5.65)		$<1.0x10^{-7}$	2.15(1.63-2.85)		$<1.0 \times 10^{-7}$
Controls	813	198 (24)	399 (49)	216 (27)		$7.35x10^{-08}$	2.54(1.48-4.41)		$2.6x10^{-4}$	2.44(1.70-3.50)		3.0×10^{-7}
Male bone erosion (+)	66	4 (6)	35 (53)	27 (41)	NS		0.97(0.12-6.66)	NS		0.62(0.25-1.53)	NS	
Male bone erosion (-)	34	2 (6)	14 (41)	18 (53)		$1.00x10^{-4}$	5.99(2.00-20.16)		$1.9x10^{-4}$	2.11(1.15-3.87)		0.0091
Male controls	251	70 (28)	119 (47)	62 (25)		$1.65x10^{-4}$	6.19(1.39-38.39)		$5.5x10^{-3}$	3.43(1.56-7.58)		0.0006
Female bone erosion(+)	249	22 (9)	116 (47)	111 (44)	NS		1.50(0.72-3.12)	NS		0.97(0.62-1.53)	NS	
Female bone erosion (-)	126	16 (13)	53 (42)	57 (45)		$5.55x10^{-9}$	3.04(1.84-5.07)		$2.4x10^{-6}$	2.13(1.54-2.94)		1.5×10^{-6}
Female controls	562	128 (20)	280 (43)	154 (37)		$6.92x10^{-5}$	2.03(1.13-3.70)		0.012	2.19(1.44-3.32)		8.7×10^{-5}

OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test with bone erosion and without bone erosion

b: Cochran-Armitage trend test comparing to controls

c: uncorrected chi-square test

4.4.6. Association of serum anti-CCP status and SNP genotypes

In order to deduce the genetic association of serum anti-CCP in RA patients within the 24 studied loci, the RA patients are sub-grouped into those with serum anti-CCP antibody (anti-CCP+) and those without serum anti-CCP antibody (anti-CCP-). Three statistical tests are included: anti-CCP+ versus anti-CCP- RA patients; anti-CCP+ RA patients versus normal controls; and anti-CCP- RA patients versus normal controls. The -log p-values for these genetic associations are presented in Figures 4-17 to 4-19.

Figure 4-17 Genetic association of serum anti-CCP antibody in anti-CCP(+) and anti-CCP(-) RA patients

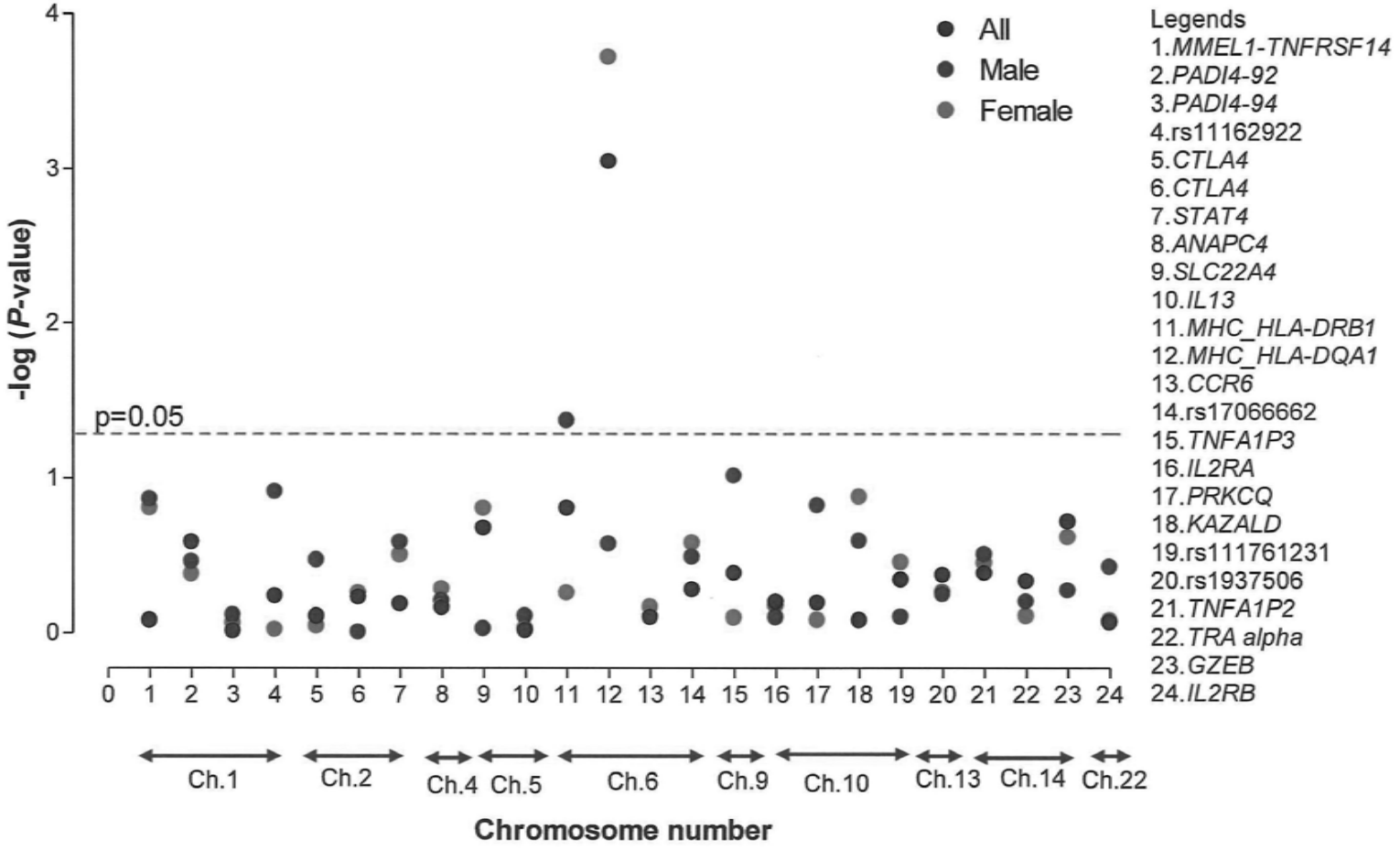


Figure 4-18 Genetic association of serum anti-CCP antibody in anti-CCP(+) RA and normal controls

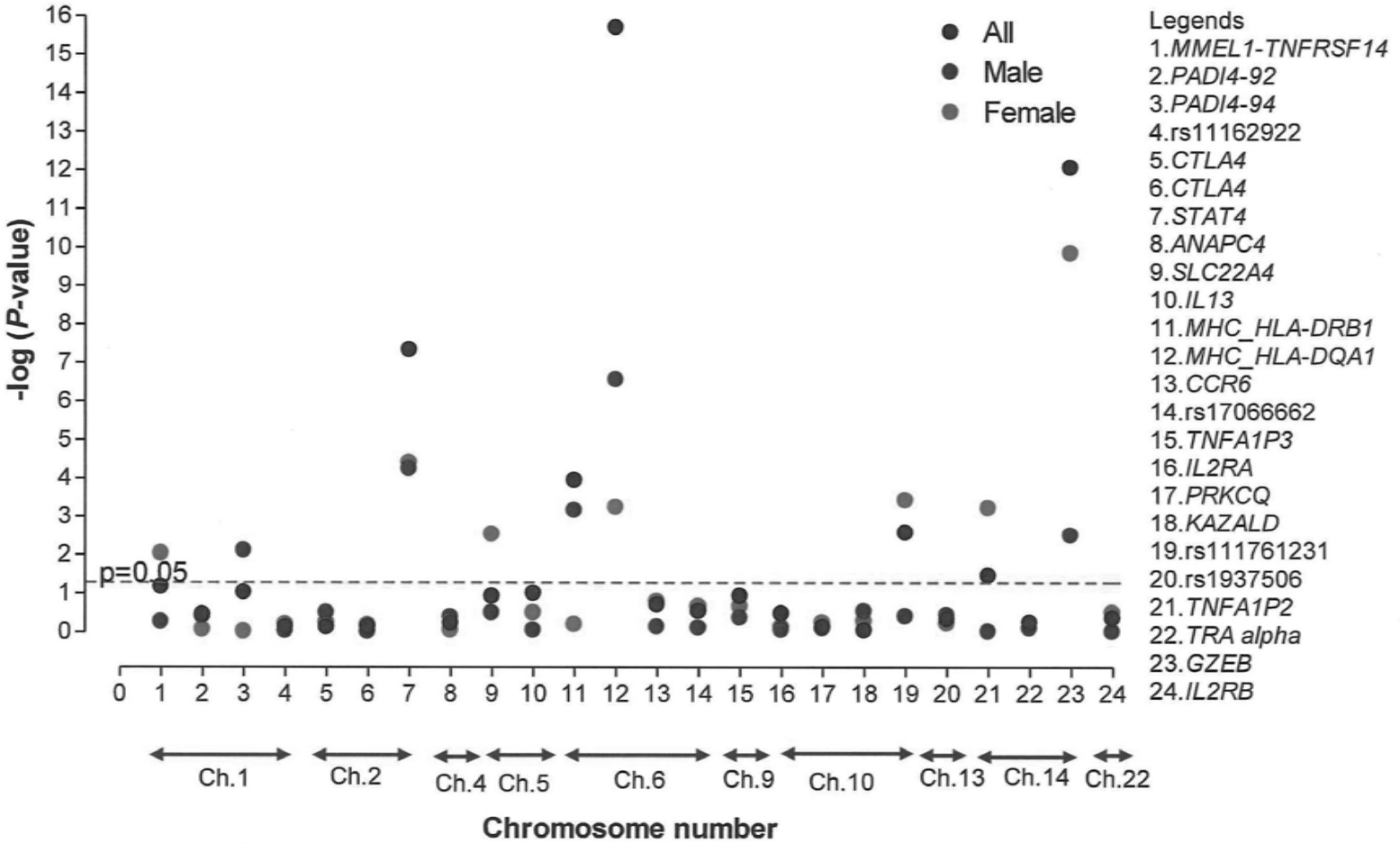
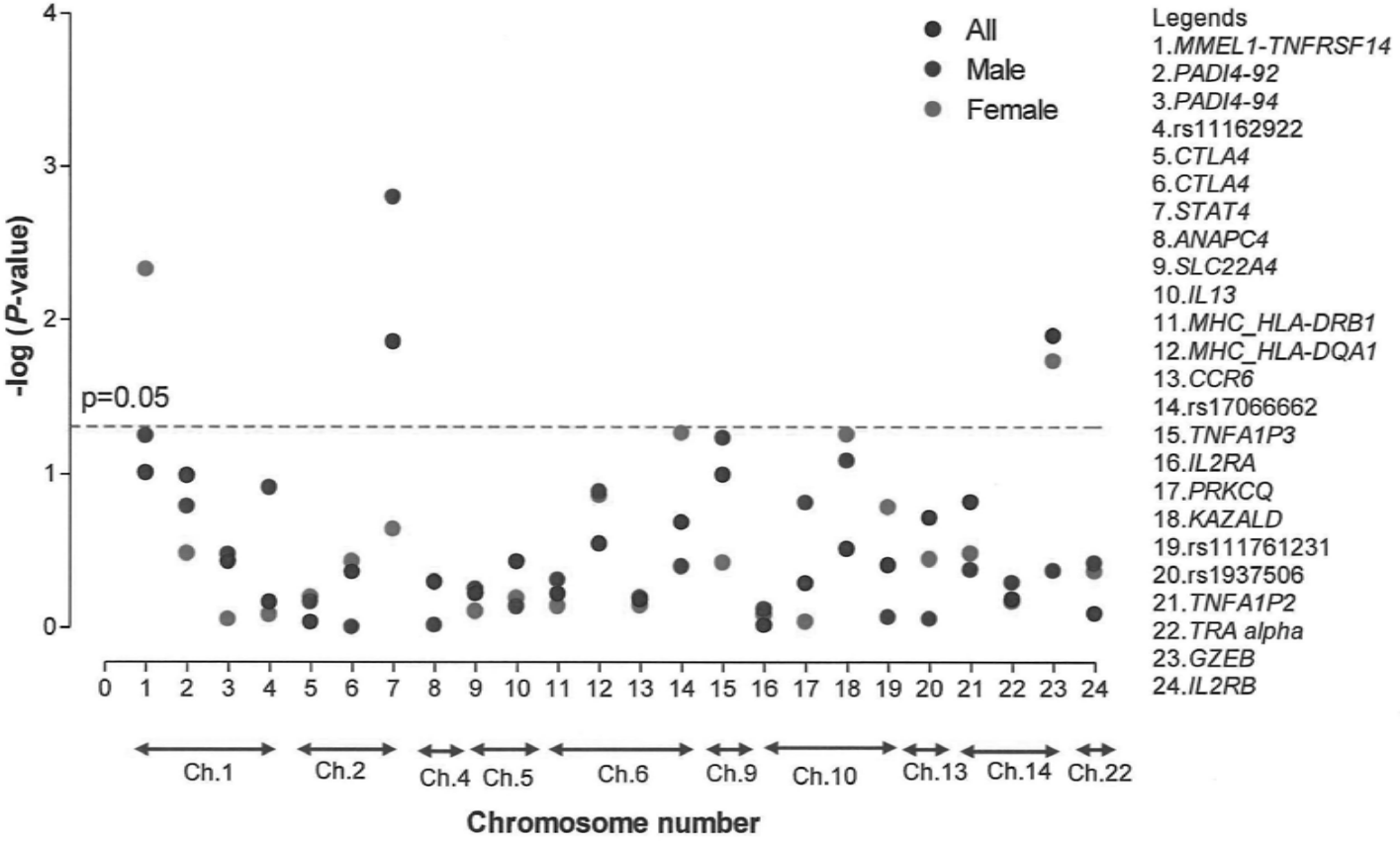


Figure 4-19 Genetic association of serum anti-CCP antibody in anti-CCP(-) RA and normal controls



4.4.6.1. Association of serum anti-CCP status and HLA loci

The anti-CCP antibody in RA patients was associated with the genetic polymorphism frequencies with rs615672 in *HLA-DRB1* (p for trend test < 0.050), in Table 4-22 and with rs6457617 in *HLA-DQA1* (p for trend test < 0.050) in Table 4-23.

Table 4-22 showed that the genetic frequency of the minor GG allele with rs615672 in *HLA-DRB1* was higher in anti-CCP positive RA men then in normal controls (p for trend = 0.043). The genetic frequencies of the minor GG genotype in anti-CCP positive RA patients are higher than the normal controls (20% versus 12 %, p for trend = 1.15×10^{-4}). Upon subgroup analysis, both male and female RA patients with anti-CCP showed higher frequencies of GG genotype in RA case than normal controls (p for trend = 0.0007 and 0.0089, respectively).

In a dominant model, the genetic frequency of the major G allele in *HLA-DRB1* reduced the risk for anti-CCP in anti-CCP positive RA patients compared to anti-CCP negative RA patients (OR = 1.68, 95% CI: 1.03 – 2.76; p- value = 0.027). Upon subgroup analysis, this association was not found in either men or women anti-CCP positive and normal controls (p for trend > 0.05).

Table 4-22 Association of anti-CCP status and *HLA-DRB1*genotype in RA patients and normal controls

		Genotype, n (%)					Dominant model			Recessive model		
							OR (95%CI)			OR (95% CI)		
HLA-DRB1	n	CC	CG	GG	*P*trend[a]	*P*trend[b]	CG+GG vs CC	P^c	P^c	GG vs. CG+CC	P^c	P^c
anti-CCP (+)	417	132 (32)	202(48)	83 (20)	NS	$1.15x10^{-4}$	1.68 (1.03-2.76)	0.027		0.98 (0.54-1.81)	NS	
anti-CCP (-)	89	39 (44)	32 (36)	18 (20)		NS	1.42 (1.10-1.84)		0.005	1.78 (1.28-2.47)		2.86×10^{-4}
Normal controls	866	344 (40)	416 (48)	106 (12)			0.84 (0.53-1.34)		NS	1.82 (1.00-3.27)		0.0328
Male anti-CCP (+)	88	22 (25)	44 (50.0)	22 (25)	0.043	0.0007	2.63 (0.75-0.22)	NS		4.67 (0.58-100.4)	NS	
Male anti-CCP (-)	15	7 (47)	7 (47)	1 (6)		NS	1.94 (1.10-3.47)		0.015	2.63 (1.36-5.08)		0.0015
Male controls	267	105 (39)	132 (50)	30 (11)			0.74 (0.24-2.35)		NS	0.56 (0.03-4.36)		NS
Female anti-CCP(+)	329	110 (33)	158 (48)	61 (19)	NS	0.0089	1.52 (0.88-2.61)	NS		0.76 (0.40-1.47)	NS	
Female anti-CCP(-)	126	32 (43)	25 (34)	17 (23)		NS	0.32 (0.99-1.77)		NS	1.57 (1.07-2.30)		0.016
Female controls	599	239 (40)	284 (47)	76 (13)			0.87 (0.52-1.46)		NS	2.05 (1.09-3.85)		0.016

OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test between anti-CCP positive and anti-CCP negative groups

b: Cochran-Armitage trend test between anti-CCP positive and anti-CCP negative groups to controls

c: Uncorrected chi-square test

Table 4-23 shows the genetic frequency of the minor TT genotype with rs6457617 in *HLA-DQA1* was higher in anti-CCP positive RA women then in normal controls (p for trend = 9.0×10^{-4}). Upon subgroup analysis, this effect is retained only in the anti-CCP positive RA women when compared to anti-CCP negative RA women (p for trend = 1.9×10^{-3}). Genotypic frequencies of the SNP followed HWE ($p > 0.05$). In the dominant genetic model, the T allele increases risk for anti-CCP positive status (OR = 2.98, 95% CI: 1.47-6.02; p value = 4.15×10^{-4}). Upon subgroup analysis, this effect was retained only in the anti-CCP positive RA women when compared to normal controls (OR = 0.34, 95 % CI: 0.17 - 0.68; p value = 7.4×10^{-4}). In the recessive model, the minor TT genotype increased the risk for anti-CCP positive status in women RA patient (OR = 1.71, 95 % CI: 0.98 – 3.01; p = 4.5×10^{-2}).

Compared to the normal controls, the frequencies of the minor TT genotype in *HLA-DQA1* was higher in RA patients with anti-CCP (p for trend = 2.0×10^{-16}). Upon subgroup analysis this association was retained in both male and female anti-CCP positive RA subgroups when compared with normal controls (p for trend = 2.8×10^{-7} and 1.5×10^{-10} respectively). The genetic frequencies for the minor TT genotypes with *HLA-DQA1* in anti-CCP negative RA are similar to the normal controls (p for trend > 0.05). In the dominant genetic model, the frequency of the T allele in *HLA-DQA1* increases the risk of anti-CCP status in RA patients when compared with control (OR = 3.59, 95% CI: 2.41 - 5.35; p-value < 1.0×10^{-7}). Upon subgroup analysis, this association was retained in both male and female anti-CCP antibody positive RA patients when compared to normal controls (OR = 6.19, 95% CI: 2.29 – 18.12; p-value = 2.5×10^{-5}, OR = 3.05, 95 % CI: 1.96-4.76; p-value = 1.0×10^{-7} respectively). In the recessive model, the TT genotype increased the risk for

anti-CCP positive status to over two-fold when compared to normal controls (OR = 2.28, 95% CI: 1.77 - 2.94; $p < 1.0 \times 10^{-7}$). Upon subgroup analysis, this risk was retained in both male and female subgroups (OR = 2.98, 95 % CI: 1.73 - 5.14; $p = 2.1 \times 10^{-5}$ and OR = 2.09, 95 % CI: 1.56 - 2.80; $p = 3.0 \times 10^{-7}$ respectively).

Table 4-23 Association of anti-CCP status and *HLA DQA1* genotype in RA patients and controls

		Genotype, n (%)					Dominant model			Recessive model		
							OR (95%CI)			OR (95% CI)		
HLA-DQA1	n	CC	TC	TT	*P*trend[a]	*P*trend[b]	CT+TT vs CC	P^c	P^c	TT vs. CT+CC	P^C	P^C
anti-CCP (+)	425	35 (8)	198 (47)	192 (45)	$9.0x10^{-4}$		2.90 (1.50-5.57)	$4.1x10^{-4}$		1.70 (1.03-2.82)	0.027	
anti-CCP (-)	92	19 (21)	43 (47)	30 (32)		$2.0x10^{-16}$	3.59 (2.41-5.35)		$<1.0x10^{-7}$	2.28 (1.77-2.94)		$<1.0x10^{-7}$
Controls	813	198 (24)	399 (49)	216 (27)		NS	1.24 (0.71-2.18)		NS	1.34 (0.82-2.17)		$<1.0x10^{-7}$
Male anti-CCP (+)	85	5 (6)	38 (45)	42 (49)	NS		2.29 (0.27-15.56)	NS		1.63 (0.49-5.60)	NS	
Male anti-CCP (-)	16	2 (12)	8(50)	6 (38)		$2.8x10^{-7}$	6.19 (2.29-18.12)		$2.5x10^{-5}$	2.98 (1.73-5.14)		$2.1x10^{-5}$
Male controls	251	70 (28)	119 (47)	62 (25)		NS	2.71 (0.57-17.72)		NS	1.83 (0.56-5.76)		NS
Female anti-CCP(+)	340	30 (9)	160 (47)	150 (44)	$1.9x10^{-3}$		2.98 (1.47-6.02)	$7.4x10^{-4}$		1.71 (0.98-3.01)	$4.5x10^{-2}$	
Female anti-CCP(-)	76	17 (22)	35 (46)	24 (32)		$1.5x10^{-10}$	3.05 (1.96-4.76)		$1.0x10^{-7}$	2.09 (1.56-2.80)		$3.0x10^{-7}$
Female controls	562	128 (20)	280 (43)	154 (37)		NS	1.02 (0.56-1.90)		NS	1.22 (0.70-2.11)		NS

OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test between anti-CCP positive RA and anti-CCP negative RA patients

b: Cochran-Armitage trend test between anti-CCP positive RA and anti-CCP negative RA to normal controls

c: Uncorrected chi-square test

4.4.6.2. Association of serum anti-CCP status and TNF pathway

The genotype frequencies of *TNFA1P2*, *TNFA1P3*, *MMEL1-TNFRSF14* and rs17066662 are similar in anti-CCP positive or anti-CCP negative status RA patients (p for trend > 0.05). Upon subgroup analysis, the genotype frequencies of the risk GG genotype in rs2771369 (*TNFA1P2*) (Table 4-24) and the risk AA genotype in rs6684865 (*MMEL1-TNFRSF14*) (Table 4-25) are higher among anti-CCP positive RA patients when compared to normal controls (p for trend = 0.0006 and 0.008, respectively). These associations are only found in female RA patients. In the dominant model the A allele in rs2771369 reduce risk for RA (OR=0.73, 95% CI: 0.55 - 0.97; p = 0.025). In a recessive model, the AA genotype is a protective genotype (OR = 0.40, 95 % CI: 0.23 - 0.67; p = 0.0002). In the dominant model, the A allele among anti-CCP positive RA patients in rs6686865 increased risk for RA when it is compared to normal controls (OR = 1.90, 95% CI: 1.32-2.76; p = 0.0003) and the anti-CCP negative subgroup reduced risk for RA (OR = 2.29, 95% CI: 1.17 - 4.51; p = 0.0086).

The genotypic frequencies of polymorphism in *TNFA1P3* (rs3761847) and *rs17066662* did not show any association with anti-CCP status (data not shown). No associations are found in the dominant or the recessive genetic models.

Table 4-24 Association of anti-CCP status and *TNFA1P2* (rs2771369) genotype in RA patients and controls

		Genotype, n (%)					Dominant model			Recessive model	
							OR (95%CI)			OR (95% CI)	
TNFA1P2	n	GG	GA	AA	*P*trend[a]	*P*trend[b]		*P*[c]	*P*[c]	AA vs. GA+GG	*P*[c]	*P*[c]
							GA+AA vs GG					
anti-CCP (+)	422	203(48)	189(45)	30(7)	NS							
anti-CCP (-)	92	44(48)	33(36))	15(16)		0.035	0.81 (0.63-1.03)		NS	0.49 (0.31-0.76)		$7.2x10^{-4}$
Normal controls	804	344(43)	351(44)	109(14)		NS						
Male anti-CCP (+)	86	45(52)	32(37)	9(10)	NS							
Male anti-CCP (-)	17	7(41)	7(41)	3(18)		NS						
Male controls	275	136(49)	106(39)	33(12)		NS						
Female anti-CCP(+)	336	158(47)	157(47)	21(6)	NS							
Female anti-CCP(-)	75	37(49)	26(35)	12(16)		0.0006	0.73 (0.55-0.97)		0.025	0.40 (0.23-0.67)		0.0002
Female controls	529	208(39)	245(46)	76(14)		NS						

OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test between anti-CCP positive and anti-CCP negative RA patients

b: Cochran-Armitage trend test between anti-CCP positive or anti-CCP negative and normal controls

c: Uncorrected chi-square test

Table 4-25 Association of anti-CCP status and *MMEL1-TNFRSF14* (rs6684865) in RA patients and controls

		Genotype, n (%)					Dominant model			Recessive model		
							OR (95%CI)			OR (95% CI)		
MMEL1-TNF SF14	n	GG	GA	AA	*P*trend[a]	*P*trend[b]	GA+AA vs GG	*P*[c]	*P*[c]	AA vs. GA+GG	*P*[c]	*P*[c]
anti-CCP (+)	409	95(23)	226(55)	88(23)	NS							
anti-CCP (-)	90	21(23)	45(50)	24(27)		NS						
Normal controls	595	177(30)	299(50)	119(20)		NS						
Male anti-CCP (+)	80	23(29)	42(53)	15(19)	NS							
Male anti-CCP (-)	16	7(44)	8(50)	1(6)		NS						
Male controls	299	74(25)	166(55)	59(20)		NS						
Female anti-CCP(+)	329	72(22)	184(56)	73(22)	NS							
Female anti-CCP(-)	74	14(19)	37(50)	23(31)		0.008	1.90 (1.32-2.76)		0.0003	1.12 (0.75-1.68)-		NS
Female controls	296	103(35)	133(45)	60(20)		0.004	2.29 (1.17-4.51)		0.0086	1.77 (0.97-3.25)		0.046

OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test between anti-CCP positive and anti-CCP negative RA patients

b: Cochran-Armitage trend test between anti-CCP positive or anti-CCP negative and normal controls

c: Uncorrected chi-square test

4.4.7. Genetic association of rheumatoid factor status and genotypes

The RA patients are sub-grouped into those with RF and those without RF. The genotype of the 24 SNPs in the two groups of RA patients are compared to each other (Figure 4-20) and to normal controls (Figures 4-21 to 4-22).

Significant associations are found in four SNPs. They are *MMEL1_TNFSF14* in females only, *ANAPC4* and rs1937506 in males only, *SLC22A4* in all patients . However, upon subgroup analysis by sex, the association of serum RF status in *SLC22A4* polymorphism was removed.

There are nine SNPs associated with serum RF positive status when compared with normal controls ($p > 0.05$: Figure 4-21) and nine SNPs associated with negative RF status when compared with normal controls ($p > 0.05$: Figure 4-22).

Figure 4-20 The –log p-values of the 24 SNPs in rheumatoid arthritis patients with and without rheumatoid factor

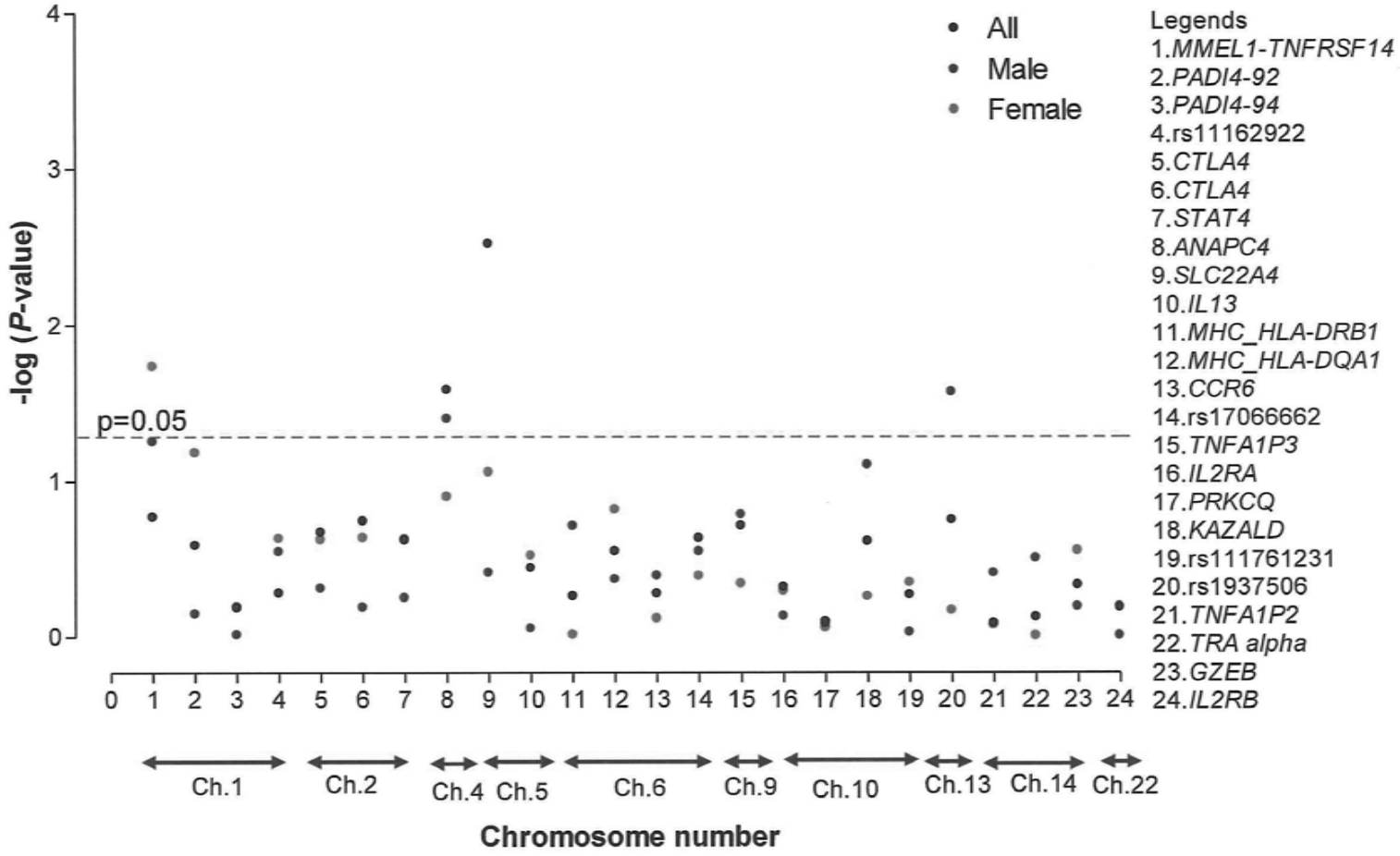


Figure 4-21 Genetic association of RA with RF and normal controls

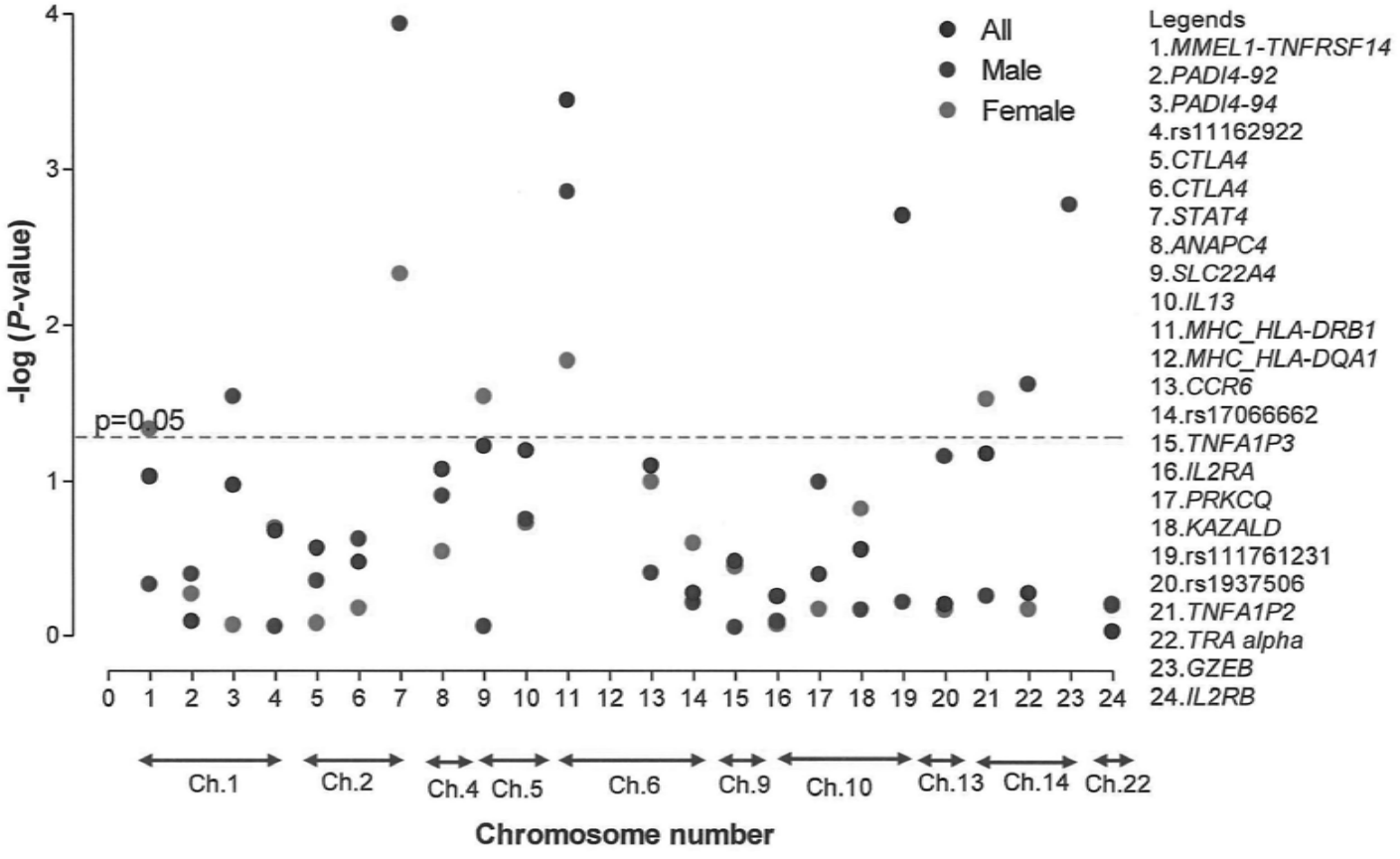
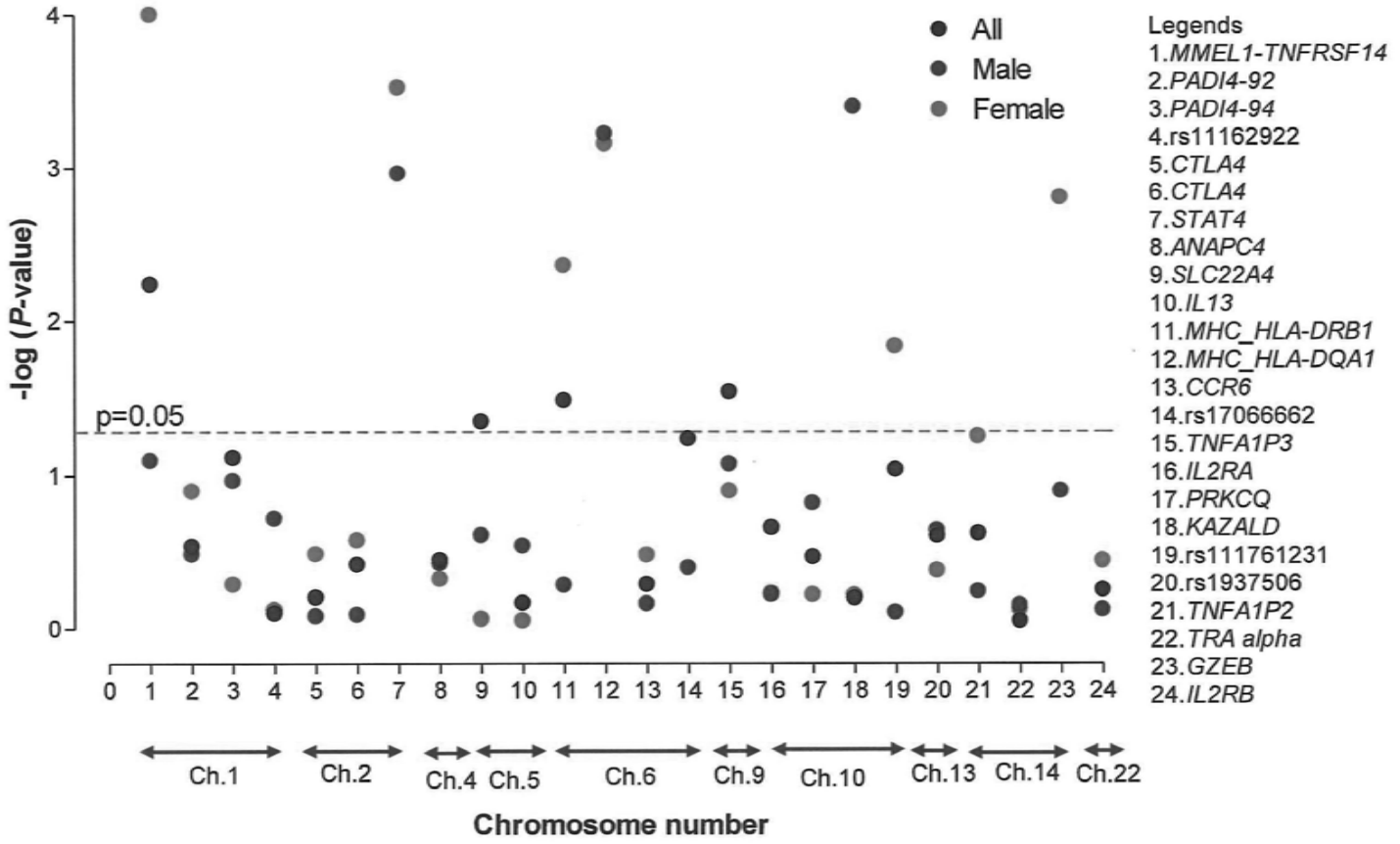


Figure 4-22 Genetic association of negative RF status in RA and normal controls



There are no significant associations between genetic polymorphisms and RF status in *TNFA1P2* (Table 4-26), *TNFA1P* (Table 4-27) and *MMELI-TNFRSF14* (Table 4-28). Upon subgroup analysis, a modest association (*p* for trend = 0.001) was found between RF positive RA female patients and the polymorphism in *MMELI-TNFRSF14* loci. Both RF+ RA and RF- RA patients are found to be significantly different from the normal controls (p for trend = 0.046, and 3.01×10^{-6}, respectively). In a dominant genetic model, the A allele increased risk for RF (OR = 1.83, 95% CI: 1.23-2.72, p = 0.0015 and OR=2.50, 95% CI: 1.50-4.20, p=0.00016, respectively).

Table 4-26 Association of RF status and rs2771369 in *TNFA1P2*

		Genotype, n (%)					Dominant model			Recessive model		
							OR (95%CI)			OR (95% CI)		
TNFA1P2	n	GG	GA	AA	*P*trend[a]	*P*trend[b]	GA+AA vs GG	*P*[c]	*P*[c]	AA vs. GA+GG	*P*[c]	*P*[c]
RF (+)	346	162(47)	151(44)	33(9)	NS							
RF (-)	181	88(49)	70(39))	23(12)		NS						
Normal controls	804	344(43)	351(44)	109(14)		NS						
Male RF (+)	75	39(52)	29(39	7(9)	NS							
Male RF(-)	30	14(47)	11(37)	5(16)		NS						
Male controls	275	136(49)	106(39)	33(12)		NS						
Female RF(+)	271	123(45)	122(45)	26(10)	NS							
Female RF(-)	151	74(49)	59(39)	18(12)		NS						
Female controls	529	208(39)	245(46)	76(14)		0.03	0.67 (0.46-0.99)		0.033	0.81 (45-1.44		NS

OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test between RF positive and RF negative RA patients

b: Cochran-Armitage trend test between RF positive or RF negative and normal controls

c: Uncorrected chi-square test

Table 4-27 Association of RF status and rs3761847 in *TNFA1P3*

		Genotype, n (%)					Dominant model			Recessive model		
							OR (95%CI)			OR (95% CI)		
TNFA1P3	n	GG	GA	AA	*P*trend[a]	*P*trend[b]	GA+AA vs GG	*P*[c]	*P*[c]	AA vs. GA+GG	*P*[c]	*P*[c]
RF (+)	336	100(30)	176(52)	60(18)	NS							
RF (-)	175	57(33)	96(55)	22(13)		NS						
Normal controls	1019	296(29)	510(50)	213(21)		0.029	0.85 (0.59-1.21)		NS	0.54 (0.33-0.89)		0.01
Male RF (+)	75	21(28)	38(51)	16(21)	NS							
Male RF(-)	29	11(38)	15(52)	3(10)		NS						
Male controls	299	77(26)	165(55)	57(19)		NS						
Female RF(+)	261	79(30)	138(53)	44(17)	NS							
Female RF(-)	146	46(31)	81(56)	19(13)		NS						
Female controls	720	219(30)	345(48)	156(22)		NS						

OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test between RF positive and RF negative RA patients

b: Cochran-Armitage trend test between RF positive or RF negative and normal controls

c: Uncorrected chi-square test

Table 4-28 Association of RF status and rs6684865 in *MMEL1-TNFRSF14*

		Genotype, n (%)					Dominant model			Recessive model		
							OR (95%CI)			OR (95% CI)		
MMEL1-TNF SF14	n	GG	GA	AA	*P*trend[a]	*P*trend[b]	GA+AA vs GG	P^c	P^c	AA vs. GA+GG	P^c	P^c
RF (+)	326	74(23)	184(56)	68(21)	NS							
RF (-)	177	37(21)	91(51)	49(28)		NS						
Normal controls	595	177(30)	299(50)	119(20)		0.005	1.60 (1.05-2.45)		0.021	1.53 (1.02-2.29)		0.0296
Male RF (+)	69	16(23)	36(52)	17(25)	NS							
Male RF (-)	29	11(38)	15(52)	3(10)		NS						
Male controls	299	74(25)	166(55)	59(20)		NS						
Female RF(+)	257	58(23)	148(58)	51(20)	0.001		0.73 (0.42-1.26)	NS		0.55(0.34-0.90)	0.011	
Female RF(-)	148	26(18)	76(51)	46(31)		0.046	1.83 (1.23-2.72)		$1.15x10^{-3}$	0.97(0.63-1.51)		NS
Female controls	296	103(35)	133(45)	60(20)		$3.01x10^{-6}$	2.50 (1.50-4.20)		$1.60x10^{-4}$	1.77(1.10-2.85)		0.011

OR, odds ratio; CI, confidence interval; NS, not statistically significant

a: Cochran-Armitage trend test between RF positive and RF negative RA patients

b: Cochran-Armitage trend test between RF positive or RF negative and normal controls

c: Uncorrected chi-square test

4.4.8. Cigarette smoking

There are no statistically significant differences in genotype frequencies between smoker RA patients and non-smoker RA patents (Figure 4-23) in the 24 loci except in rs17066662, which is a TNF tagged SNP.

Figure 4-24 shows the genotype frequencies in the 24 studied loci between smoker RA patients to normal controls. There are statistically significant differences in genotype frequencies between smoker RA patients and normal controls in the following loci in *HLA-DRB1* ($p = 0.017$), *HLA-DQA1* ($p = 5.39 \times 10^{-7}$), *TNFA1P3* ($p = 0.032$), *GREB* ($p = 0.0038$), and *STAT4* ($p = 0.00023$).

Figure 4-25 shows the genotype frequencies in the 24 studied loci between non-smoker RA patients to normal controls. There are statistically significant differences in genotype frequencies between non-smoker RA patients and normal controls in *HLA-DRB1* ($p = 0.004$), *HLA-DQA1* ($p = 5.22 \times 10^{-13}$), *TNFA1P3* ($p = 0.008$), *GREB* ($p = 7.7 \times 10^{-10}$), *STAT4* ($p = 6.21 \times 10^{-8}$) and rs11761231 ($p = 0.006$).

Figure 4-23 Genetic association of cigarette smoking status in RA

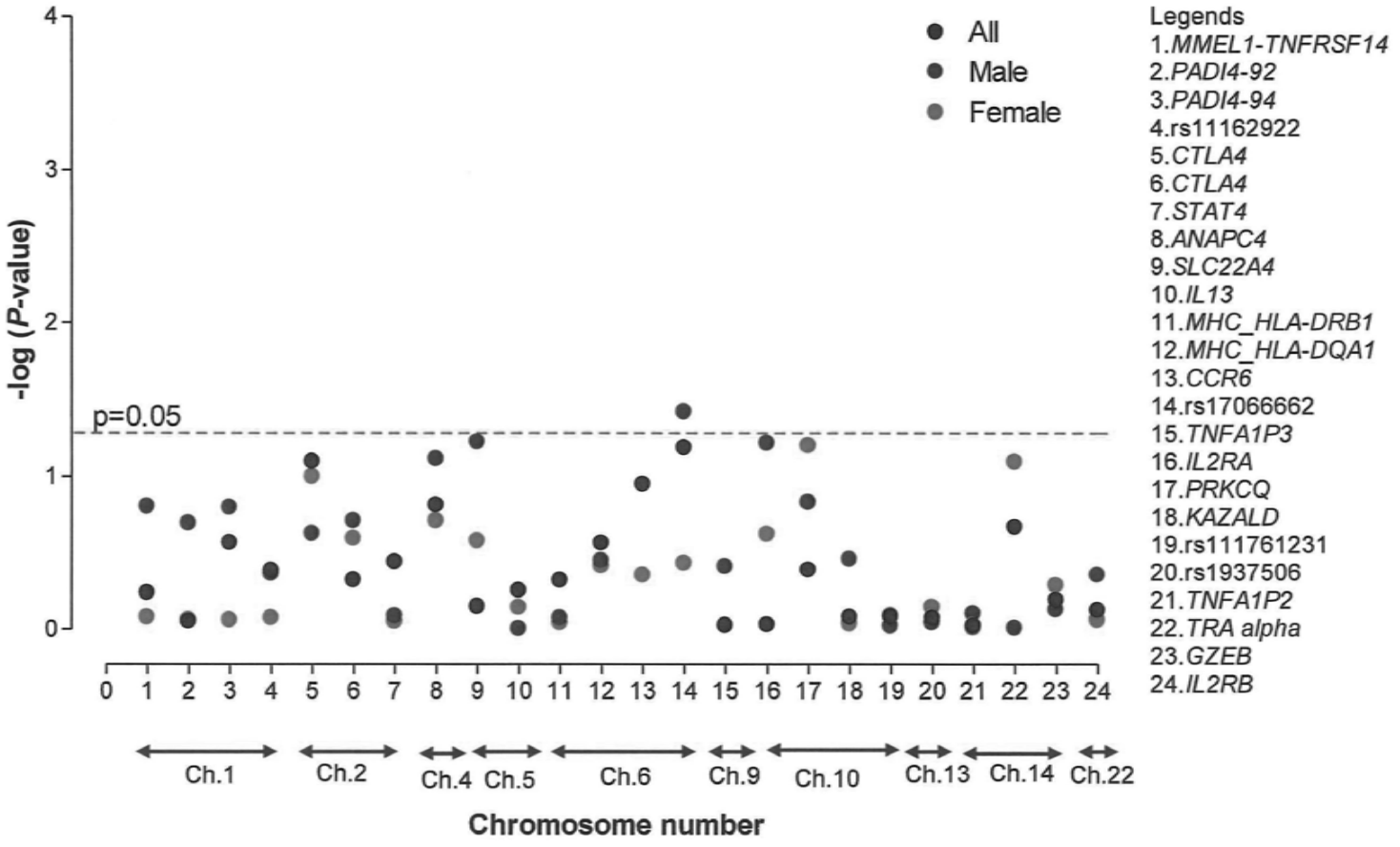


Figure 4-24 Genetic association of ever cigarette smoker in RA and controls

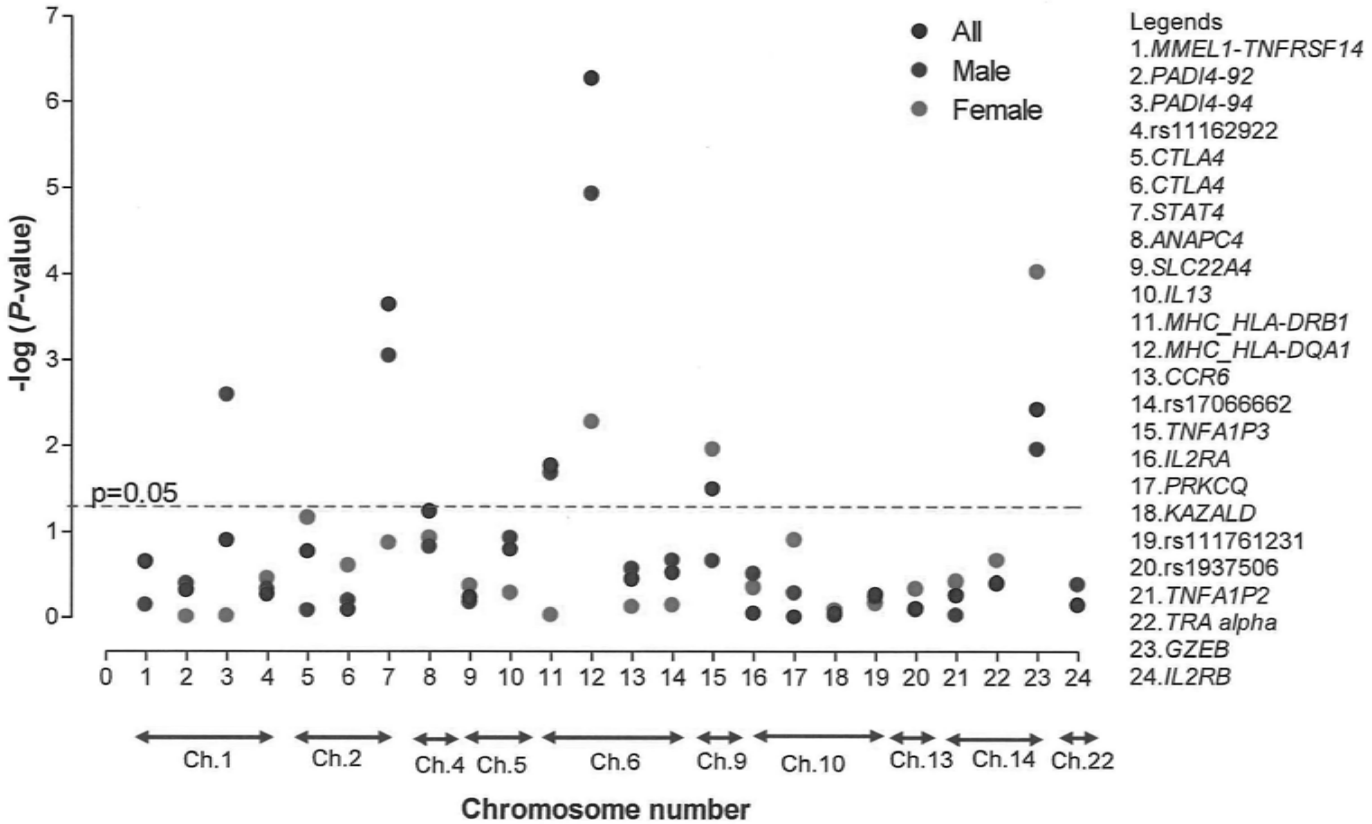
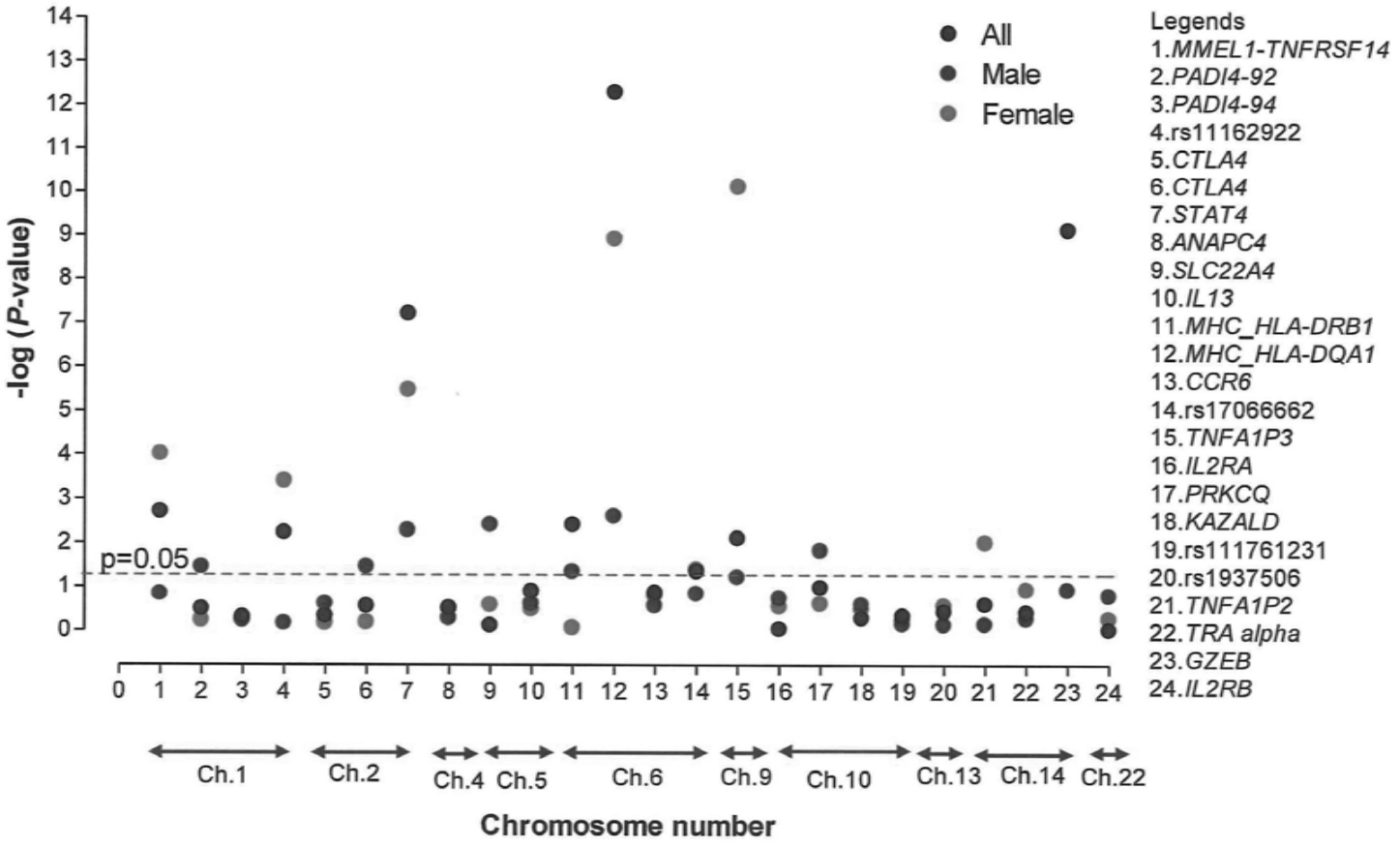


Figure 4-25 Genetic association of never cigarette smoker in RA and controls



Chapter 5 Discussion

This thesis is divided into three parts.

1. We examined the genetic predisposition of 25 SNPs in Chinese with rheumatoid arthritis compared to normal controls.
2. We examined the gene-gene interactions between these SNPs.
3. We also examined the associations in these 25 SNPs with serum biomarkers, such as anti-CCP antibodies, RF, bone erosions and smoking status.

5.1.1. Characteristics of the rheumatoid arthritis patients

With reference to the results of the demographic data presented in Table 4-1, women are more susceptible to RA. The prevalence of RA in men and women was 1:4. This ratio was similar to that of previously reported studies found elsewhere (Carmona et al.2002).

Compared to men, women suffered from RA at an earlier age. In this study the onset age for RA in men are 4.2 older than in women (50.0 years compared to 45.8 years old), and we find more SNPs in women associated with RA.

5.1.2. Biomarkers

The percentage of RA patients with detected anti-CCP level was 85%. This was similar to previously reported studies in Chinese (Liao et al.2011; Zhao, J. et al.2010), Japanese (Tamai et al.2010), and in Caucasians (Inanc et al.2007; Potter et al.2009). However, it was higher than previous reported studies, 42% (Bongi et al.2004) and 55% (Forslind et al.2004). The differences were due to small sample sizes (n=54, 379, respectively) and shorter disease duration in both studies (less than 6 months).

Furthermore, the anti-CCP levels were measured with the first generation anti-CCP test kits (Bongi et al.2004; Forslind et al.2004), whereas anti-CCP-2 kits were used in other studies (Liao et al.2011; Potter et al.2009; Tamai et al.2010; Zhao, J. et al.2010).

RF was present in 65% of the RA patients and this figure was similar to previously reported studies (De Vries et al.2005; Inanc et al.2007; Quinn et al.2006; Silveira et al.2007; Steuer et al.2008).

The percentage of RA patients with anti-CCP and RF are 55%, which was similar to previously reported studies (Lee, DM and Schur2003; Sauerland et al.2005; Steuer et al.2008). The sensitivity of the diagnosis of RA with anti-CCP is 70-78%, with RF is 59-79% and with both markers is over 91% (Lee, DM and Schur2003).

5.1.3. Bone erosion on X-ray

In this study, 35 established RA cases have no bone erosion on X-rays, therefore, the percentage of bone erosion on X-rays is under-estimated in this study.

There are 66% of the RA patients who showed bone erosion on X-ray, which is different from the bone erosion data of the previous studies of 49% (Bongi et al.2004) and 55% (Syversen et al.2010). The differences are probably due to the small sample sizes in those two studies, 35 and 253 respectively, and a more likely reason is the mean years of RA duration are less than 1 to 2.3 years, whereas the RA disease duration mean years in this study was over 9 years.

5.1.4. Other demographic data

In this study, the number of tender and swollen joints, DAS-28 and rheumatoid nodules are similar in both men and women. However, compared to established data,

our data on tender joints, swollen joints and DAS-28 are significantly lower than that in Japanese (Kawashiri et al.2010), Turks (Samanci et al.2005) and Caucasians (Liu et al.2008; Miriovsky et al.2010; Potter et al.2009). The differences might be partly explained by the small sample sizes in some of these studies where n values range from 45 to 89 patients. In contrast, when compared with other studies of similar sample size, our data on DAS-28 was lower than that of Potter el al (Potter et al.2009) but similar to that of Miriovsky et al. (Miriovsky et al.2010). This suggested that the number of tender and swollen joints and DAS28 are not good parameters for the diagnosis of RA since they are based on physician opinion. As for the DAS28 score, there are different methods used in the literature and variations between test centers.

5.1.5. Family history

There are more women who had a family history for developing RA than men. However, the actual numbers are small (7.1%). Established familial studies based on the Norfolk Arthritis Register suggested the increased risk was only two-fold (Jones et al.1996). On the other hand, Criswell et al. (Criswell et al.2005) showed that 89% of the affected familial individual are females. Such observations may support the fact that females are more likely to develop RA than males. Criswell et al. also pointed out that this female tendency may be due to the fact that females are more knowledgeable about their family history and are likely to participate in any etiological studies (Criswell et al.2005).

5.1.6. Cigarette smoking and alcohol drinking.

The number of smokers and drinkers are mainly found in men. There are fewer smokers and drinkers in females. In this study, smokers and drinkers were defined as

ever users according to the current and past use/intake of cigarettes or alcohol. No attempts were made to quantify these in terms of years of duration or usage. A more detailed questionnaire could have been used as cigarette smoking may be a confounding factor for RA. Therefore, cigarette pack years, total number of cigarette smoked may be employed. Alcohol intake is a confounding factor, the number of grams of alcohol intake per week and duration in years are also important.

5.2. Genetic associations

Genome wide association case and control studies are clearly effective in improving our understanding of the genetic basis of rheumatoid arthritis. They are considered as an established tool for high-throughput study and the established list of SNPs would provide a base for the study of RA in other ethnic and minority groups.

Established results from the WTCCC (WTCCC2007) and NARAC (Remmers et al.2007) studies established that the SNPs associated with RA are within GWAS p-values of 10×10^{-6} to 10×10^{-8}. However, the WTCCC study found nine new SNPs and some of their GWAS p-values are less than 10×10^{-6}. This suggested that it is possible to identify rare SNPs with low allele frequencies. In this case-control study we replicated 25 of these established SNPs and found that 24 of them are polymorphic in the Chinese population but one SNP is not polymorphic in the Chinese population.

Of the 24 polymorphic SNPs, seven of them showed p for trend < 0.05. When the data are analyzed by the dominant model two more SNPs are identified to be associated with RA. Likewise three SNPs are identified in the recessive models. These suggested that it is important to study SNPs with significant GWAS p values as well as insignificant p values. These results will be discussed in the follow pages.

Logistic regression for RA disease showed age, sex, T allele in *STAT4* and C allele *ANAP4* were significantly negatively associated with RA, and CC genotype in *HLA-DQA1* was positively associated with RA.

In the next part of the thesis, the results of MHC and non-MHC SNPs will be discussed.

5.2.1. Association of Major Histocompatibility Complex, Class II, DR beta 1

loci and DQ alpha 1

This study in Chinese RA patients confirmed that rs6457617 in *HLA-DQA1* and rs615672 in *HLA-DRB1* are significantly associated with RA. Both genes are located on chromosome 6p21. The *DRB1* locus has long been shown to be associated with RA (Stastny, Peter1978). Due to its size and complexity, it remains as one of the unknown and yet best studied region for RA. The first full sequence was done in 1999 and it showed that the MHC region consisted of 224 loci, with 128 of these loci expressed, and 40% of the expressed proteins have immunological properties. *HLA-DRB1* is a confirmed susceptibility locus in other autoimmune diseases such as T1D, SLE, AS, PsA etc (Fernando et al.2008). The structural functions related to antigen processing and presentation to HLA class I (HLA: A, B, C) and Class II (HLA: DP, DQ, DR) genes are emerging. To date, more than 100 diseases have been mapped to HLA genotypes. Some disorders involved single HLA genes such as AS (HLA-B27), and others required multiple HLA genes such as in celiac disease and RA. The MHC region shows extensive allelic variation and linkage disequilibrium and a small number of genes.

In this study we established not only that *HLA-DRB1* is involved in RA but also the *HLA-DQA*. They are both located on chromosome six and only 89kb apart. A linkage disequilibrium plot was carried out and the results suggested that they are not in linkage disequilibrium but they are independent from one another.

5.2.2. Association of non Major histocompatibility Complex loci

In Caucasians, the WTCCC GWAS study showed that *PTPN22* is the most important non-MHC SNP associated with RA and another study showed that a common mutation rs2476601 in *PTPN22* was related to four autoimmune diseases, RA, SLE, T1D and Hashimoto thyroiditis (Criswell et al.2005). However, our

present study suggested that it is non-polymorphic in the Chinese population. In the early HapMap data it was suggested to be polymorphic in the Chinese but it was subsequently found to be non-polymorphic. (Rodriguez-Rodriguez et al.2011). For other *PTPN22* SNPs including rs33996649, no Chinese Hapmap data is published, the pooled six countries reported T allele frequency is 0.025 in Caucasians (Rodriguez-Rodriguez et al.2011). rs66799677 is a tagged SNP to rs2476601, it is non polymorphic in Chinese. However, for a recently reported SNP rs2488257, the reported G allele rate ranged 0.3 to 0.9 and this is a new candidate SNP for the Asian population (Thompson et al.2010). Another SNP rs1893217 in *PTPN22* gene with minor allele frequency ranging from 4 to 12% is another option (Zhernakova et al.2011).

The present study focused on the previously reported regions of *SLC22A4-SLC22A5-IL13*, TNF pathway, *IL2RA* and *IL2RB*, *CTLA-4, PADI-4, GZEB, STAT4, PRKCQ, KAZALD, CCR6, TRA alpha, ANAPC4*, and three other SNPs of undetermined functions rs11162922, rs1937506 and rs11761231.

5.2.2.1. Association of Solute carrier family 22 and Interleukin-13

Established studies showed that the polymorphisms in the *SLC22A4-SLC22A5-IL13* loci are associated with autoimmune disease (Barton, A. et al.2005; Martinez et al.2006; Yamada et al.2004; Yamazaki et al.2004). In 2008, Li et al. established that the polymorphisms in rs11568506 in *SLC22A4* and rs1800925 in *IL-13* are associated with psoriasis and Crohn disease. Furthermore, the rs2073838 in *SLC22A4* has been shown to be associated with SLE and RA (Yamada et al.2004). However, rs11568506 is monomorphic in Chinese. The *IL-13* region was located downstream of *SLC22A4-SLC22A5*. Therefore in the present study we studied rs2073838 (*SLC22A4)* and rs1800925 (*IL-13*) together. Their p for trend values are >

0.05. Both SNPs did not show any significance in the association of RA. No association was found in rs1800925 in *IL-13*. However, in the recessive model, the risk AA genotype of the rs2073838 suggested a 2-fold increase in the risk of RA in male patients only, but this association could not be replicated in the female subgroup. Our data are different from that of Caucasian based studies (Barton, A. et al.2005; Newman et al.2005) whereas in those studies the frequency of the risk AA genotype was 0.3% compared to 11% for the present study. Other Japanese studies also found no genetic association with RA (Kawasaki et al.2004). The connections of *SLC22A4* and RA are further investigated by gene-gene interactions with *Runx1* loci (Maeda et al.2007; Martinez et al.2006; Yamada et al.2004). Yet again, inconclusive results are drawn and the role of either loci with RA remained unknown. Therefore, the genetic association for the risk of RA with polymorphism in *SLC22A4* was probably mild.

5.2.2.2. Association of Tumour Necrosis Factor pathway

Four SNPs in the TNF pathways are replicated in this study. They are rs2771369 *TNFA1P2* in chromosome 14, rs3761847 *TNFA1P3* in chromosome 9, rs6684865 *MMELI-TNFRSF14* in chromosome 1 and rs17066662 in chromosome 6, whereas rs17066662 is a tagged SNP of rs10499194 (*TNFAIP3-OLIG3*) at 6q23 (Van Der Helm-Van Mil et al.2010). The rs10499194 loci had been shown to have a strong genetic association with RA in Caucasians (Plenge et al.2007), Patsopoulos et al. established that the genetic association of rs10499194 with RA was small (p-value 5.6×10^{-3}), and in 2010, Shimane et al. established that it was polymorphic in the Japanese population with a minor allele frequency in the range of 5-7%. The odds ratio was 1.3 and the p-value was small 8.4×10^{-4} (Shimane et al.2010). The TT minor allele frequency of its tagged SNP rs17066662 was 1% in our study, and there

was no association with the risk for RA with rs17066662. The differences in these observations may reflect the variations within the study cohorts and the linkage of the SNP to unknown loci. The p for trend tests for both *TNFA1P2* and *MMELI-TNFRSF14* are significant ($p < 0.05$). Upon subgroup analysis, this association was only found in females. Furthermore, in the subgroup analysis a small association was observed in *TNFA1P3*. It was interesting to know that the three SNPs within the TNF pathway showed significant association with increased risk for RA in female but not in the male RA patients. This gender difference in the genetic association of TNF loci was observed in only one Mexican study on rs1800625 (*TNF alpha- 380*) (Jimeez-Morales et al.2009). On the other hand due to the small sample size of the male RA patients (n = 120), the power of estimation was small. Previously established studies on RA patients treated with monoclonal antibodies to anti-TNF alpha showed reduction in DAS28 scores (Cui et al.2010; Liu et al.2008). This may suggest the TNF pathway is equally important in both males and females. To date, there was no genetic study that demonstrated a direct genetic association of TNF loci with RA, but all established studies are associated with small molecules within the TNF pathway. In this study we showed that three SNPs of the TNF pathway are genetically associated with RA in female patients.

5.2.2.3. Association of Interleukin-2 Receptor Alpha and Interleukin-2 Receptor Beta

rs2104286 in *IL2-RA* and rs743777 in *IL2-RB* are established by the WTCCC study to be associated with RA. They had GWA p-values in the range of 10×10^{-6} to 10×10^{-8} (Kurreeman et al.2009; WTCCC2007). In a recessive model of rs2104286 in *IL-2RA*, a small odds ratio was observed in the female subgroup (OR = 0.42, p =

0.03). No genetic association was found in Chinese RA patients with rs743777 in the *IL-2RB* loci. The differences may be explained by the different genotype frequency found between the two studies, but the frequencies are similar to that of published data. Likewise, we also showed that there was a sex differentiation in the risk of RA in the *IL-2RA* loci in Chinese patients with RA.

5.2.2.4. Association of Cytotoxic T Lymphatic Antigen 4

Our study showed that both SNPs rs3087243 and rs10497873 in *CTLA-4* are not associated with RA in Chinese. Our findings are similar to those established by Kim et al., where they reported that two intronic SNPs rs231777 and rs231779 in the *CTLA-4* loci are not associated with RA in Koreans (Kim, YO et al.2010). Moreover, Daha et al. pooled six Caucasian and two Asian studies published between 2004 and 2008 to study the association of rs3087243 and RA (Daha et al.2009). Half of the pooled studies did not show any genetic association between rs3087243 and RA. However, in a recessive model from our study, the minor TT alleles of rs10497873 increased the risk of RA by two-fold. The discrepancy between these studies may suggest that rs3087243 was not the functional marker for RA. The previous findings between this SNP and RA may have resulted from another proximal marker, which is in linkage disequilibrium with rs3087243.

5.2.2.5. Association of Peptidylarginine Deiminase 4

The enzyme PADI-4 was involved in the conversion of arginine to citrulin in the process of citrulation, which induced the auto-antibodies to CCP (anti-CCP). It has been suggested that PADI-4 might play a role in RA since 85% of the RA patients are anti-CCP positive. It was also suggested that cigarette smoking induced

the formation of anti-CCP auto-antibodies and induced RA onset. In the present study, we observed no genetic associations in RA patients with smoking or anti-CCP producers. Hence, there was no association between smoking, anti-CCP status and *PADI-4* genotype.

In the present study, a weak association was found between the polymorphism in *PADI-4-94* (rs2240340) in male RA patients (p for trend = 0.005), the odds ratio for the GG genotype was over two for the risk of RA. A recent Japanese study with 2000 subjects suggested that rs2240340 is a risk allele among RA patients but on the other hand in one Chinese study with 600 RA cases that studied 5 SNPs within the *PADI-4* loci it failed to confirm any associations (Chen, R et al.2010). Those established results suggested that the polymorphism of *PADI-4* might be weakly associated with RA. However, in our gene-gene interaction study, we established that there was a genetic association between the *PADI-4* loci and *HLA DRB1*.

In summary, there are ethnic differences between the genetic associations of *PADI-4* in RA. Based on established reports, this association was found only in Asian populations (Fan et al.2008; Suzuki et al.2003; Takata et al.2008) and it was extremely rare in Caucasian populations (Barton, A. et al.2004; Burr et al.2010; Mori, M et al.2005; Plenge et al.2005).

5.2.2.6. Association of Signal Transducer and Activator of Transcription 4

The TT genotype in rs7574865 was demonstrated as a common risk allele in RA, SLE, systemic sclerosis and other autoimmune diseases. In the present study, the TT risk allele increased the RA risk by a factor of 2.4 and the p for trend was 3.9×10^{-9}. The odds ratio in this study was higher than other established studies which ranged from OR = 1.15 to 1.3 (Lee, HS et al.2007; Remmers et al.2007). This difference

cannot be explained by smaller sample size or the use of normal controls alone, but it is more like to be an ethnicity issue, the Hapmap data on the minor risk T allele is higher in Chinese (35%) compared to 20% in Caucasian and 15% in Africans. Moreover, STAT4 plays a central role in the activation of T cells into Th1 and Th17 helper T cells (figure 2.2), whereas the activated Th1 and Th17 T cells in turn regulate plasma B-cells and inflammatory cytokines production. The Th1 and Th2 cells are regulated by macrophages. The exact mechanism of how *STAT4* SNPs are involved in RA is unknown, since the four *STAT4* SNPs are intronic. Hence, the role of *STAT4* SNPs might be in gene regulation or gene splicing.

5.2.2.7. Association of Granzyme, Protein Kinase C, theta, and Kazal Type Serine Protease Inhibitor Domain

In the present study, there was no association between rs4750316 and with RA, which was first established by the WTCCC study (WTCCC2007). The minor C allele frequency in both Caucasians and Asians is similar (13.3% in Chinese, 17.8 in Japanese and 18.3% in Caucasians) (Consortium2005; Consortium2007) and moderate p for trend value was obtained in the WTCCC study (p-trend = 5.55×10^{-5}).

5.2.2.8. Association of Chemokine (C-C motif) Receptor 6

In the present study, we failed to replicate the association of *CCR6* polymorphism and the risk of RA, which was first reported by Kochi et al. (Kochi et al.2010) and Stahl et al by meta analysis (Stahl et al.2010). The reasons for being unable to replicate the significance in rs3093024 might be due to the fact that both Kochi et al and Stahl et al involved very large sample sizes (cases = 7,069 and 5,539; controls = 20,729 and 20169, respectively) compared to our 600 cases and 900

controls. Moreover, the reported increased risk of RA from rs3093024 was associated with serum IL-17 levels (Kochi et al.2010). Hence, *IL-17* may be involved in the underlying mechanism in the development of RA. In order to explain this observation, serum IL-17 level may be measured in our study and the association of *IL-17* locus and RA may also be examined.

5.2.2.9. Association of T-cell Receptor Alpha

No association was found between *TRA alpha* polymorphism and the risk of RA. In the present study, there was no significant association between RA risk and rs1154155. The association of *TRA alpha* was first reported in an autoimmune disease narcolepsy (Hallmayer et al.2009). Since T cell receptor alpha played a central role in the interaction of both T and B cells, it therefore seemed to be a good candidate SNP to study. However, our observation yielded negative results for its association with RA. The difference may be explained by the different frequency of the risk minor C allele in Caucasian (14%), Asian (47%) and African (8%) populations (Consortium2005; Consortium2007; Hallmayer et al.2009)

5.2.2.10. Association of SNPs rs11162922, rs1937506 and rs11761231.

These three SNPs are reported by the WTCCC study. In the present study, no association was found between the polymorphism frequencies and the risk for RA in the former two SNPs, both in male and female patients with RA. Upon subgroup analysis and in both recessive and dominant genetic models no significant risk was found. However with rs11761231, the risk for RA disease was only found in female RA patients (p for trend = 8.5×10^{-3}). The TT genotype appears to be a protective

genotype among female patients. This finding was similar to that of the WTCCC report.

5.3. Gene–gene interaction

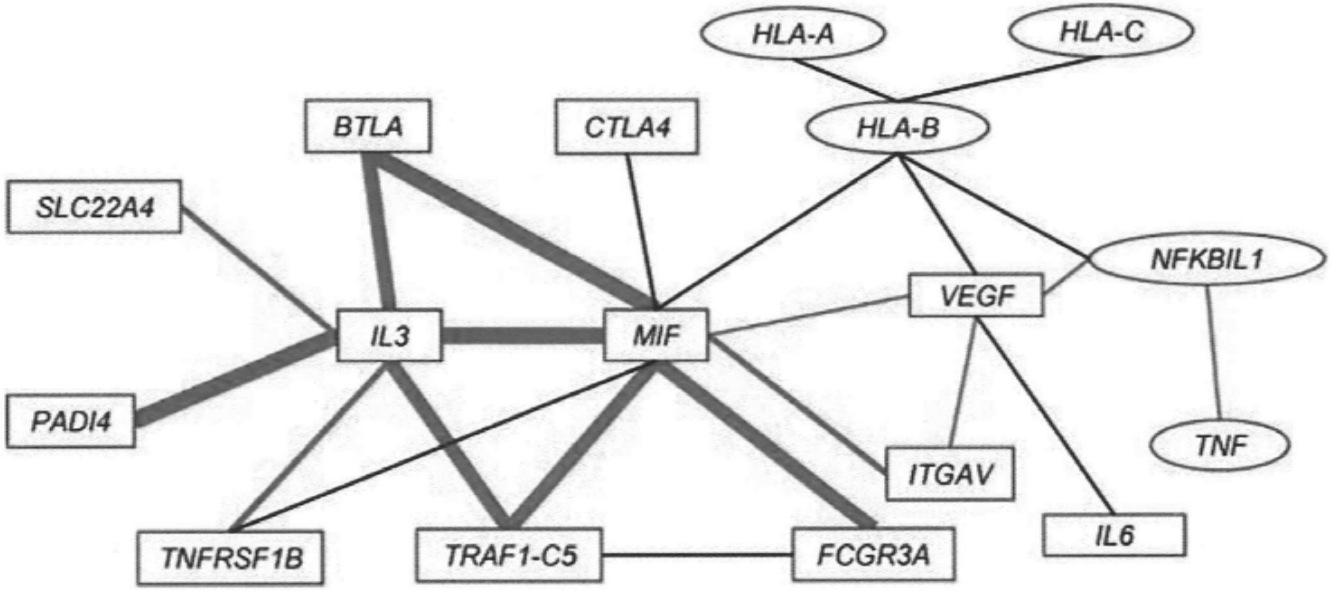
To date based on single SNP and haplotype analysis, about 60 SNPs are associated with RA genetics. HLA regions showed the most importance of all regions. Non-HLA SNPs are *PTPN22*, *STAT4*, *PADI-4*, *SLC22A4*, *TNF*, *CTLA-4*, *IL2RA*, *IL2RB* and other minor loci. However, these findings did not reflect the possible interactions between multiple SNPs. From 2010, several established studies have focused on the epistasis of loci for complex diseases (Feng and Zhu2010; Jung et al.2009; Kallberg et al.2007; Zhang, M et al.2009). In this study we carried out gene-gene interaction analysis of RA based on two HLA and 22 non-HLA SNPs that are derived mainly from GWAS studies.

In this study as expected, the HLA regions accounted for most influence on RA phenotype. We found four interaction signals with HLA regions: between *HLA-DRB1* and *HLA-DQA1, HLA-DRB1* and *PADI-4, HLA-DRB1* and *IL2RA,* and between *HLA-DRB1* and rs1162922. The signal within the HLA region was of two independent loci. This finding suggested there was an interaction between HLA-DRB1 and HLA-DQA in MHC Class II antigens in RA genetics. This finding was similar to that of Liu et al. (Liu et al.2011). We observed that there are interactions between HLA Class II region with *PADI-4*, *IL2RA* and rs1162922 loci of undetermined function.

Outside the HLA region, there are three pairs of genetic interactions, *SLC22A4* and *ANAPC4*, *IL13*and *CTLA-4*, and between *MMELI-TNFRSF14* and *TRA-alpha.* However, the sample sizes in some of the subgroups are small, so that more data will

be needed to confirm the influence of these interactions. It was unexpected to see that no gene-gene interactions are found with *STAT4, TNFA1P2*, and *TNFA1P3*. These single SNPs are well-validated and significant in our SNP analysis. Similar observations are noted by Liu et al. (Liu et al.2011) but Manning et al. showed one interaction between *TRAF1* and *TAP2* (Manning et al.2009), and based on the NARAC data set, 14 significant interactions are identified to include *PADI-4*, *TNFSF1B*, *SLC22A4*, and *TNF* (Huang, CH et al.2009; Qiao et al.2009), where *SLC22A4* was also an important interaction in our study.

Figure 5-1 Gene-gene interactions in rheumatoid arthritis identified by Huang el al. (Huang, CH et al.2009)



5.4. Genetic association of serum biomarkers and RA

In the present study, several serum biomarkers and indices are studied. They are bone erosion, anti-CCP status and level, RF status and level, and cigarette smoking status.

5.4.1. Genetic association of bone erosion in rheumatoid arthritis

The SNP rs10786617 in *KAZALD1* was previously suggested to be associated with bone erosion in RA patients. With data from chapter 4.4.3, there are no genetic associations between *KAZAlDI* genotype and bone erosion. There is no evidence to suggest that bone erosion (+) group and bone erosion (-) group are genetically different. There was also no evidence to suggest that the normal controls are different from the bone erosion (+) or the bone erosion (-) groups. The Kazal type serine protease inhibitor domain did not appear to play any role in the bone erosion RA patients according to these data.

5.4.1.1. Major histocompatibility Complex, *HLA-DRB1* association with bone erosion

The RA patients are classified according to bone erosion and non-bone erosion groups. The genotypes in each group are compared between themselves and to normal controls. Bone erosion is associated with *HLA-DRB1*gene in female RA patients only (p for trend =0.037; Table 4-19), but this association could not be replicated in male patients with RA. There was a significant difference between the bone erosion positive group and normal controls (p for trend = 0.001). Both of the recessive and dominant genetic models showed that the CC genotype was a protective allele whereas the GG genotype was a risk allele for bone erosion. Figure

4-14 showed the 24 polymorphic SNPs studied in this study with their –log p for trend values and their genetic association. No association was found between bone erosion and *HLA-DQA1* loci Table 4-20.

Figure 4-15, presented the genetic association of bone erosion in RA patients with bone erosion to normal controls. Figure 4-16, presented the genetic association of non-bone erosion RA patients to controls. It demonstrated that after subgroup analysis by bone erosion, it did not remove the association of RA risk in the following SNPs *MMEL1-TNF, PADI-4, STAT4, MHC-HLA-DRB1, HLA-DQA1, TNFA1P3, rs111761231 TNFA1P2, and GREB*. This suggested that even after stratification the genetic association with RA in those loci remained unchanged.

5.4.2.1. Association of anti-CCP status and SNP genotypes

Anti-CCP level had been suggested to be associated with *HLA-DRB1*shared epitopes. In the present study, the anti-CCP status was studied in the genotypes of the 24 SNPs. The results are presented in Figure 4-17. Presence of anti-CCP was associated with *MHC-HLA* genes, namely *MHC-HLA DRB1* and *HLA-DQA1*. No genetic associations are found in any other non-MHC loci. The G-allele in the *DRB1* loci was a risk factor for anti-CCP positive RA (Table 4-21) and the TT-allele in the *DQA1* loci was a risk factor for anti-CCP positive RA (Table 4-22). The association of *DQA1* with anti-CCP level was a new finding and it had not been reported elsewhere. The associations of anti-CCP in RA patients are sex-specific, in which the HLA Class II beta chain (*HLA-DRB1*) affected the male patients, whereas the HLA Class II alpha chain (*HLA-DQA*) affected the females.

The anti-CCP had no association with *PADI-4* genotype even through it was the enzyme responsible for the production of anti-CCP.

Figure 4-18, presented the genetic associations of anti-CCP positive RA patients compared to normal controls in the 24 studied SNPs. Figure 4-19, presented the genetic associations of anti-CCP negative RA patients compared to controls in the 24 studied SNPs. The genetic associations for the risk of RA in the following loci TNF (*MMEL1-TNFRSF14, TNFA1P2* and *TNFA1P3*), the HLA loci (*HLA-DRB1,* and *HLA-DQA1), PADI-4, STAT4, rs11176123, and GREB* are retained even after stratification according to anti-CCP status (Figures 4-18 to 4-19).

5.4.2.2. Association of anti-CCP and Tumour Necrosis Factor loci

There are strong associations of AA-allele in *TNFA1P2* loci (Table 4-23) between anti-CCP positive RA females and normal controls and the AA-allele in *MMELI_TNFSF14* loci (Table 4-24) between anti-CCP positive RA females and normal controls. These results reflected that loci within the TNF pathway are candidate loci for the risk of RA.

5.4.3.1. Genetic association of RA patients with or without RF

In RA patients, three SNPs (female only: *MMEL1_TNFSF14;* male only: *ANAPC4* and rs1937506) (Figure 4-20) are associated with the presence of serum RF. Our results presented in Figure 4-20 showed that the genetic associations of RF in RA patients and loci polymorphisms are sex-specific.

5.4.3.2. Genetic association of RF and Tumour Necrosis Factor loci

No association was observed between patients with serum RF in *TNFA1P2* or *TNFA1P3* loci. Among RA female patients, a modest genetic association was observed between RF positive versus RF negative groups compared to controls in *MMELI-TNFRSF*14 loci. The number of subjects at each subgroup was small and

there was no RF status in the control group available for analysis. Therefore, further work will be needed to elucidate the association of RF and TNF loci.

5.4.4. Association of cigarette smoking status in rheumatoid arthritis

It has been suggested that cigarette smoking might play a role in RA disease activity. Smoking RA patients in general are worse in term of disease activity. It has also been suggested that cigarette smoking might induce the production of anti-CCP autoantibodies. In the present study, the number of ever smokers was 67.0 % (n = 80) in males and 6.4% (n = 31) in females (Table 4.1). The genotypes of the SNPs are studied. These results are presented in Figure 4.-23. There was only one TNF tagged SNP rs17066662 with a p - value less than 0.05 in the male subgroup. Due to the small sample size in this subgroup, it may be concluded that there was no association between smoking status and genotype in RA patients. The worse disease activity in RA patients among smokers was not associated with *PADI-4* or anti-CCP status but perhaps by other confounding factors it was associated with smoking.

In Figures 4-24 to 4-25, the cigarette smokers in RA patients compared to controls and the non-cigarette smokers in RA patients are compared to controls. These data demonstrated that after subgroup analysis by smoking status, the association of *MMEL1-TNFSF14, PADI-4, STAT4, MHC-HLA-DRB1, HLA-DQA1, TNFA1P3, rs111761231 TNFA1P2*, and *GREB* in the risk of RA still remained. This suggested those associations are strong.

Chapter 6 Conclusion

RA is a chronic disease affecting 0.35% of the local population (Lau et al.1993), the current Hong Kong population stands at 7,067,800 (Hksar2011), hence, 2,473,730 patients may be affected by RA. This figure is modified with increased influx of people coming from neighbouring cities. A new survey may be needed in order to establish the current RA prevalence, since the last one was done in 1993 (Lau et al.1993) and the current treatments for RA are moving towards the costly biological agents.

This study revealed the prevalence of RA in women is four times higher than in men. Women developed RA disease a few years earlier than men. The number of SNPs attribute to RA is higher in women. This study affirms both *HLA-DRBA1* and *HLA-DQA1* loci are associated with increased risk of RA. Polymorphism in *STAT4* loci is the most important non MHC loci that increased risk of RA. The universal increased risk from these three SNPs is found in Chinese as well as that in Caucasians.

In men, RA is associated with polymorphisms in *PADI4*, *ANAPC4*,and *SLC22A4*, whereas in women it is associated with polymorphisms in SNPs within the TNF pathway, (*MMEL1-TNFRSF14*, *TNFA1P2*, *TNFA1P3), CTLA-4, SLC22A4, IL2RA* and rs1117612311. There is no evidence to suggest GZEB, PRKCQ, KALZALD, CCR6 TRA alpha, rs11162922, and rs1937506 increase the risk for RA.

Age, sex, *HLA-DQA1*, *STAT4* and *ANAPC4* are predictors for RA by logistical multiple regression. Ageing, being a female subject and C allele polymorphism in *ANAPC4* increased the risk for RA, wheres CC genotype in *HLA-DQA1* reduced the risk of RA in Chinese.

RA specific biomarkers were found in 82.2% of all patients. Serum anti-CCP is associated with *HLA-DRB1* and *HLA-DQA1* loci within the MHC region. There is no evidence to suggest anti-CCP is associated with PADI4 loci of cigarette smoking. RF was found in 65.4% of all patients. It is asscoiated with polymorphisms in *MMEL1_TNFSF14* loci in RA women only; *ANAPC4* and rs1937506 loci in RA men only,and *SLC22A4* in all RA patients.

Bone erosion is found in 62.8% of all RA cases, However, this may be an underestimated value since bone erosion data were not available in 5.8% (n=35) of the RA patients. Bone erosion is not associated with polymorphism in *KAZALD1* loci. The GG genotype in *HLA-DRB1* loci is associated with increased risk of bone erosion.

6.2. Future work and prospects

The genetic replication studies on GWAS candidate genes and risk of RA provided new insights to the pathogenesis of RA in Chinese. Genetic polymorphisms of the MHC-HLA alleles are thought to play an important role in RA pathogenesis, anti-CCP status, and the risk of bone erosion, especially with the *HLA DRBA1* loci. Results with the seldom mentioned *HLA-DQA1* gene indicated that it is a new region of interest in term of the pathogenesis of the disease.

The universal increased risk of *STAT4* in many autoimmune diseases suggested the complexity of the pathogenesis of RA is in common with other autoimmune diseases. Other candidate genes, such as these within the TNF pathway, are located in different chromosomes but all demonstrated that only the female patients are affected, whereas other SNPs only induce the risk of RA in males. This study also clearly showed that the risk of RA is partly contributed to by sex, and ageing. However, the underlying mechanisms are still unknown. Therefore, functional studies on these candidate genes involved in the TNF pathway and the *MHC-HLA* interactions are indispensable for elucidating the risk of RA.

Lately, 5 new SNPs are shown to be associated with RA in the association of study rs23622 in C5orf40, rs10892279 in DDX6 gene, rs864537 in CD247, rs2298428 in URE2LS gene and rs1893217 in *PTPN22* gene. The role of C5orf40 may be associated with bone erosion and the latter rs1893217 in *PTPN22* will provide a lot of interesting debate as it is the only SNP in *PTPN22* in Chinese with minor allele frequency of greater than 5%. Moreover, in our study 42.6% of RA patients were with active disease, it would be interesting to determine if there was any association between active disease status and loci relating to the cytokines, including, IL-2, TNF, STAT4, etc.

Additional data analysis are planned, such as the use of medications over a time period and the association study on the response to medication treatment or intervention versus genotype.

Limitations of this study includes the lack of anti-CCP, RF, CRP, and DAS-28 scores in the controls, that the results were not adjusted for confounders such as cigarette smoking and the small sample size in the RA men sub-group.

References:

Adams BS, Cha HC, Cleary J, et al. (2003). "DEK binding to class II MHC Y-box sequences is gene- and allele-specific." Arthritis Res Ther **5**(4): R226-33.

Adebajo AO (1990). "Rheumatology in the Third World." Ann Rheum Dis **49**(10): 813-6.

Ahn K, Haynes C, Kim W, Fleur RS, Gordon D and Finch SJ (2007). "The effects of SNP genotyping errors on the power of the Cochran-Armitage linear trend test for case/control association studies." Ann Hum Genet **71**(Pt 2): 249-61.

Alwan WH, Dieppe PA, Elson CJ and Bradfield JW (1988). "Bone resorbing activity in synovial fluids in destructive osteoarthritis and rheumatoid arthritis." Ann Rheum Dis **47**(3): 198-205.

Anagnostopoulos I, Zinzaras E, Alexiou I, et al. (2010). "The prevalence of rheumatic diseases in central Greece: a population survey." BMC Musculoskelet Disord **11**: 98.

Armitage P (1955). "Tests for Linear Trends in Proportions and Frequencies." Biometrics **11**: 375-86.

Arnett FC, Edworthy SM, Bloch DA, et al. (1988). "The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis." Arthritis Rheum **31**(3): 315-24.

Arora T, Liu B, He H, et al. (2003). "PIASx is a transcriptional co-repressor of signal transducer and activator of transcription 4." The Journal of biological chemistry **278**(24): 21327-30.

Ates A, Kinikli G, Turgay M, Akay G and Tokgoz G (2007). "Effects of rheumatoid factor isotypes on disease activity and severity in patients with rheumatoid arthritis: a comparative study." Clin Rheumatol **26**(4): 538-45.

Balci K, Erdogan F and Bayrak S (1989). "A practical test and its value in rheumatoid arthritis." Indian J Med Sci **43**(2): 32-5.

Barland P and Lipstein E (1996). "Selection and use of laboratory tests in the rheumatic diseases." Am J Med **100**(2A): 16S-23S.

Barton A, Bowes J, Eyre S, et al. (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**(4): 1117-21.

Barton A, Bowes J, Eyre S, Symmons D, Worthington J and Silman A (2005). "Investigation of polymorphisms in the PADI4 gene in determining severity of inflammatory polyarthritis." Ann Rheum Dis **64**(9): 1311-5.

Barton A, Eyre S, Bowes J, Ho P, John S and Worthington J (2005). "Investigation of the SLC22A4 gene (associated with rheumatoid arthritis in a Japanese population) in a United Kingdom population of rheumatoid arthritis patients." Arthritis Rheum **52**(3): 752-8.

Barton A, Thomson W, Ke X, et al. (2008). "Re-evaluation of putative rheumatoid arthritis susceptibility genes in the post-genome wide association study era and hypothesis of a key pathway underlying susceptibility." Hum. Mol. Genet. **17**(15): 2274-9.

Begovich AB, Carlton VE, Honigberg LA, et al. (2004). "A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis." Am J Hum Genet **75**: 330 - 7.

Berglin E, Padyukov L, Sundin U, et al. (2004). "A combination of autoantibodies to cyclic citrullinated peptide (CCP) and HLA-DRB1 locus antigens is strongly associated with future onset of rheumatoid arthritis." Arthritis Res Ther **6**: R303 - R8.

Bodman-Smith MD, Corrigall VM, Berglin E, et al. (2004). "Antibody response to the human stress protein BiP in rheumatoid arthritis." Rheumatology **43**(10): 1283-7.

Bongi SM, Manetti R, Melchiorre D, et al. (2004). "Anti-cyclic citrullinated peptide antibodies are highly associated with severe bone lesions in rheumatoid arthritis anti-CCP and bone damage in RA." Autoimmunity **37**(6-7): 495-501.

Bowes J and Barton A (2008). "Recent advances in the genetics of RA susceptibility." Rheumatology **47**(4): 399-402.

Brennan P, Bankhead C, Silman A and Symmons D (1997). "Oral contraceptives and rheumatoid arthritis: results from a primary care-based incident case-control study." Semin Arthritis Rheum **26**(6): 817-23.

Brighton SW, de la Harpe AL, van Staden DJ, Badenhorst JH and Myers OL (1988). "The prevalence of rheumatoid arthritis in a rural African population." J Rheumatol **15**(3): 405-8.

Burgos-Vargas R and Vazquez-Mellado J (1995). "The early clinical recognition of juvenile-onset ankylosing spondylitis and its differentiation from juvenile rheumatoid arthritis." Arthritis Rheum **38**(6): 835-44.

Burr ML, Naseem H, Hinks A, et al. (2010). "PADI4 genotype is not associated with rheumatoid arthritis in a large UK Caucasian Population." Ann Rheum Dis.

Carlton VE, Hu X, Chokkalingam AP, et al. (2005). "PTPN22 genetic variation: evidence for multiple variants associated with rheumatoid arthritis." American journal of human genetics **77**(4): 567-81.

Carmona L, Villaverde V, Hernandez-Garcia C, Ballina J, Gabriel R and Laffon A (2002). "The prevalence of rheumatoid arthritis in the general population of Spain." Rheumatology **41**(1): 88-95.

Chang M, Li Y, Yan C, et al. (2008). "Variants in the 5q31 cytokine gene cluster are associated with psoriasis." Genes Immun **9**(2): 176-81.

Chang X, Zhao Y, Sun S, Zhang Y and Zhu Y (2009). "The expression of PADI4 in synovium of rheumatoid arthritis." Rheumatol Int **29**(12): 1411-6.

Chen R, Wei Y, Cai Q, et al. (2010). "The PADI4 gene does not contribute to genetic susceptibility to rheumatoid arthritis in Chinese Han population." Rheumatol Int.

Chen X, Bhandari R, Vinkemeier U, Van Den Akker F, Darnell JE, Jr. and Kuriyan J (2003). "A reinterpretation of the dimerization interface of the N-terminal domains of STATs." Protein Sci **12**(2): 361-5.

Chibnik LB, Mandl LA, Costenbader KH, Schur PH and Karlson EW (2009). "Comparison of threshold cutpoints and continuous measures of anti-cyclic citrullinated peptide antibodies in predicting future rheumatoid arthritis." J Rheumatol **36**(4): 706-11.

Civilibal M, Canpolat N, Yurt A, et al. (2007). "A child with primary Sjogren syndrome and a review of the literature." Clin Pediatr (Phila) **46**(8): 738-42.

Cochran WG (1954). "Some methods for strengthening the common chi-square tests." Biometrics **10**: 417-51.

Colbaugh P and Stastny P (1978). "Antigens in human monocytes. III. Use of monocytes in typing for HLA-D related (DR) antigens." Transplant Proc **10**(4): 871-4.

Consortium TIH (2005). "A haplotype map of the human genome." Nature **437(7063)**(2005 October 27): 1299-320.

Consortium TIH (2007). "A second generation human haplotype map of over 3.1 million SNPs." Nature **449**: 851-61.

Corper AL, Sohi MK, Bonagura VR, et al. (1997). "Structure of human IgM rheumatoid factor Fab bound to its autoantigen IgG Fc reveals a novel topology of antibody-antigen interaction." Nat Struct Biol **4**(5): 374-81.

Costenbader K, Chang S-C, De Vivo I, Plenge R and Karlson E (2008). "Genetic polymorphisms in PTPN22, PADI-4, and CTLA-4 and risk for rheumatoid arthritis in two longitudinal cohort studies: evidence of gene-environment interactions with heavy cigarette smoking." Arthritis Res Ther **10**(3): R52.

Criswell LA, Pfeiffer KA, Lum RF, et al. (2005). "Analysis of families in the Multiple Autoimmune Disease Genetics Consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes." Am J of Hum Genet **76**: 561 - 71.

Cui J, Saevarsdottir S, Thomson B, et al. (2010). "PTPRC rheumatoid arthritis risk allele is also associated with response to anti-TNF therapy." Arthritis Rheum.

Daha NA, Kurreeman FA, Marques RB, et al. (2009). "Confirmation of STAT4, IL2/IL21, and CTLA4 polymorphisms in rheumatoid arthritis." Arthritis Rheum **60**(5): 1255-60.

de Vries RR, Huizinga TW and Toes RE (2005). "Redefining the HLA and RA association: to be or not to be anti-CCP positive." J Autoimmun **25 Suppl**: 21-5.

Dean AG, Dean JA, Coulombier D, et al. (1994). Epi Info Version 6: A Word Processing, Database, and Statistics Program for Epidemiology on Microcomputers. Atlanta, GA, Centers for Disease Control.

Deighton CM, Walker DJ, Griffiths ID and Roberts DF (1989). "The contribution of HLA to rheumatoid arthritis." Clin Genet **36**(3): 178-82.

Del Puente A, Knowler WC, Pettitt DJ and Bennett PH (1989). "High incidence and prevalence of rheumatoid arthritis in Pima Indians." Am J Epidemiol **129**(6): 1170-8.

Ding B, Padyukov L, Lundstrom E, et al. (2009). "Different patterns of associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis in the extended major histocompatibility complex region." Arthritis Rheum **60**(1): 30-8.

Dixey J, Solymossy C and Young A (2004). "Is it possible to predict radiological damage in early rheumatoid arthritis (RA)? A report on the occurrence, progression, and prognostic factors of radiological erosions over the first 3 years in 866 patients from the Early RA Study (ERAS)." J Rheumatol Suppl **69**: 48-54.

du Montcel ST, Michou L, Petit-Teixeira E, et al. (2005). "New classification of HLA–DRB1 alleles supports the shared epitope hypothesis of rheumatoid arthritis susceptibility." Arthritis Rheum **52**(4): 1063-8.

Dugowson CE, Koepsell TD, Voigt LF, Bley L, Nelson JL and Daling JR (1991). "Rheumatoid arthritis in women. Incidence rates in group health cooperative, Seattle, Washington, 1987-1989." Arthritis Rheum **34**(12): 1502-7.

Fan LY, Wang WJ, Wang Q, et al. (2008). "A functional haplotype and expression of the PADI4 gene associated with increased rheumatoid arthritis susceptibility in Chinese." Tissue Antigens **72**(5): 469-73.

Feng T and Zhu X (2010). "Genome-wide searching of rare genetic variants in WTCCC data." Hum Genet **128**(3): 269-80.

Fernandez-Suarez A, Reneses S, Wichmann I, Criado R and Nunez A (2005). "Efficacy of three ELISA measurements of anti-cyclic citrullinated peptide antibodies in the early diagnosis of rheumatoid arthritis." Clin Chem Lab Med **43**(11): 1234-9.

Fernando MM, Stevens CR, Walsh EC, et al. (2008). "Defining the role of the MHC in autoimmunity: a review and pooled analysis." PLoS Genet **4**(4): e1000024.

Fischel EE (1956). "The erythrocyte sedimentation rate and the C-reactive protein tests." Bull Rheum Dis **7**(Suppl 3): S 9-12.

Fisher BA, Plant D, Brode M, et al. (2011). "Antibodies to citrullinated α-enolase peptide 1 and clinical and radiological outcomes in rheumatoid arthritis." Annals of the Rheumatic Diseases.

Fonseca JE, Canhao H, Resende C, et al. (2000). "Histology of the synovial tissue: value of semiquantitative analysis for the prediction of joint erosions in rheumatoid arthritis." Clin Exp Rheumatol **18**(5): 559-64.

Forslind K, Ahlmen M, Eberhardt K, Hafstrom I and Svensson B (2004). "Prediction of radiological outcome in early rheumatoid arthritis in clinical practice: role of antibodies to citrullinated peptides (anti-CCP)." Ann Rheum Dis **63**(9): 1090-5.

Fries JF, Spitz PW and Young DY (1982). "The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales." J Rheumatol **9**(5): 789-93.

Fritsch-Stork R, Mullegger D, Skriner K, Jahn-Schmid B, Smolen JS and Steiner G (2006). "The spliceosomal autoantigen heterogeneous nuclear ribonucleoprotein A2 (hnRNP-A2) is a major T cell autoantigen in patients with systemic lupus erythematosus." Arthritis Res Ther **8**(4): R118.

Gauderman WJ (2002). "Sample size requirements for association studies of gene-gene interaction." Am J Epidemiol **155**(5): 478-84.

Germer S and Higuchi R (1999). "Single-tube genotyping without oligonucleotide probes." Genome Res **9**(1): 72-8.

Goeb V, Dieude P, Daveau R, et al. (2008). "Contribution of PTPN22 1858T, TNFRII

196R and HLA-shared epitope alleles with rheumatoid factor and anti-citrullinated protein antibodies to very early rheumatoid arthritis diagnosis." Rheumatology **47**(8): 1208-12.

Gorman JD, Lum RF, Chen JJ, Suarez-Almazor ME, Thomson G and Criswell LA (2004). "Impact of shared epitope genotype and ethnicity on erosive disease: a meta-analysis of 3,240 rheumatoid arthritis patients." Arthritis Rheum **50**(2): 400-12.

Gregersen PK, Goyert SM, Song QL and Silver J (1987). "Microheterogeneity of HLA-DR4 haplotypes: DNA sequence analysis of LD"KT2" and LD"TAS" haplotypes." Hum Immunol **19**(4): 287-92.

Gregersen PK, Moriuchi T, Karr RW, et al. (1986). "Polymorphism of HLA-DR beta chains in DR4, -7, and -9 haplotypes: implications for the mechanisms of allelic variation." Proc Natl Acad Sci U S A **83**(23): 9149-53.

Gregersen PK, Silver J and Winchester RJ (1987). "The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis." Arthritis Rheum **30**(11): 1205-13.

Guillemin F, Saraux A, Guggenbuhl P, et al. (2005). "Prevalence of rheumatoid arthritis in France: 2001." Ann Rheum Dis **64**(10): 1427-30.

Gyetvai A, Szekanecz Z, Soos L, et al. (2010). "New classification of the shared epitope in rheumatoid arthritis: impact on the production of various anti-citrullinated protein antibodies." Rheumatology (Oxford) **49**(1): 25-33.

Haas JP, Kimura A, Truckenbrodt H, et al. (1995). "Early-onset pauciarticular juvenile chronic arthritis is associated with a mutation in the Y-box of the HLA-DQA1 promoter." Tissue Antigens **45**(5): 317-21.

Hallmayer J, Faraco J, Lin L, et al. (2009). "Narcolepsy is strongly associated with the T-cell receptor alpha locus." Nature Genet **41**(6): 708-11.

Harris ED, Jr., Faulkner CS, 2nd and Brown FE (1975). "Collagenolytic systems in rheumatoid arthritis." Clin Orthop Relat Res(110): 303-16.

Harvey J, Lotze M, Stevens MB, Lambert G and Jacobson D (1981). "Rheumatoid arthritis in a Chippewa Band. I. Pilot screening study of disease prevalence." Arthritis Rheum **24**(5): 717-21.

Hassfeld W, Steiner G, Hartmuth K, et al. (1989). "Demonstration of a new antinuclear antibody (anti-RA33) that is highly specific for rheumatoid arthritis." Arthritis Rheum **32**(12): 1515-20.

Hassfeld W, Steiner G, Studnicka-Benke A, et al. (1995). "Autoimmune response to

the spliceosome. An immunologic link between rheumatoid arthritis, mixed connective tissue disease, and systemic lupus erythematosus." Arthritis Rheum **38**(6): 777-85.

Hay FC (1988). "Rheumatoid Factor: Structural and Genetic Studies Indicating Novel Binding Sites May Confer the Specificity for IgG Fc." Rheumatology **XXVII**(suppl 2): 157-61.

He H, Isnard A, Kouriba B, et al. (2008). "A STAT6 gene polymorphism is associated with high infection levels in urinary schistosomiasis." Genes Immun **9**(3): 195-206.

Heliovaara M, Aho K, Knekt P, Aromaa A, Maatela J and Reunanen A (1995). "Rheumatoid factor, chronic arthritis and mortality." Ann Rheum Dis **54**(10): 811-4.

Herring JA (1984). "Destructive arthritis of the hip in juvenile rheumatoid arthritis." J Pediatr Orthop **4**(2): 259-61.

HKSAR (2011). "Health facts of Hong Kong 2011." Department of Health.

Hodkinson B, Meyer PW, Musenge E, et al. (2010). "The diagnostic utility of the anti-CCP antibody test is no better than rheumatoid factor in South Africans with early rheumatoid arthritis." Clin Rheumatol **29**(6): 615-8.

Hoppe B, H 瓣 upl T, Egerer K, et al. (2009). "Influence of peptidylarginine deiminase type 4 genotype and shared epitope on clinical characteristics and autoantibody profile of rheumatoid arthritis." Annals of the Rheumatic Diseases **68**(6): 898-903.

Horiuchi K, Saito S, Sasaki R, Tomatsu T and Toyama Y (2003). "Expression of granzyme B in human articular chondrocytes." J Rheumatol **30**(8): 1799-810.

Huang CH, Cong L, Xie J, Qiao B, Lo SH and Zheng T (2009). "Rheumatoid arthritis-associated gene-gene interaction network for rheumatoid arthritis candidate genes." BMC Proc **3 Suppl 7**: S75.

Huang Y, Lei Y, Zhang H, Zhang M and Dayton A (2011). "Interleukin-12 treatment down-regulates STAT4 and induces apoptosis with increasing ROS production in human natural killer cells." J Leukoc Biol **90**(1): 87-97.

Imanaka T, Shichikawa K, Inoue K, Shimaoka Y, Takenaka Y and Wakitani S (1997). "Increase in age at onset of rheumatoid arthritis in Japan over a 30 year period." Ann Rheum Dis **56**(5): 313-6.

Inanc N, Dalkilic E, Kamali S, et al. (2007). "Anti-CCP antibodies in rheumatoid arthritis and psoriatic arthritis." Clin Rheumatol **26**(1): 17-23.

Jahrsdorfer B, Vollmer A, Blackwell SE, et al. (2010). "Granzyme B produced by human plasmacytoid dendritic cells suppresses T-cell expansion." Blood **115**(6): 1156-65.

James DY, A. Kulinskaya, E. Knight, E. Thompson, W. Ollier, W. Dixey, J. (2004). "Orthopaedic intervention in early rheumatoid arthritis. Occurrence and predictive factors in an inception cohort of 1064 patients followed for 5 years." Rheumatology (Oxford) **43**(3): 369-76.

Jiang J, Tang NL, Ohlsson C, et al. (2010). "Association of genetic variations in aromatase gene with serum estrogen and estrogen/testosterone ratio in Chinese elderly men." Clin Chim Acta **411**(1-2): 53-8.

Jimeez-Morales S, Velaquez-Cruz R, Ramfez-Bello J, et al. (2009). "Tumor necrosis factor-[alpha] is a common genetic risk factor for asthma, juvenile rheumatoid arthritis, and systemic lupus erythematosus in a Mexican pediatric population." Hum Immunol **70**(4): 251-6.

Johansson M, Arlestig L, Hallmans G and Rantapaa-Dahlqvist S (2006). "PTPN22 polymorphism and anti-cyclic citrullinated peptide antibodies in combination strongly predicts future onset of rheumatoid arthritis and has a specificity of 100% for the disease." Arthritis Res Ther **8**(1): R19.

John S, Myerscough A, Marlow A, et al. (1998). "Linkage of cytokine genes to rheumatoid arthritis. Evidence of genetic heterogeneity." Ann Rheum Dis **57**(6): 361-5.

Jones MA, Silman AJ, Whiting S, Barrett EM and Symmons DP (1996). "Occurrence of rheumatoid arthritis is not increased in the first degree relatives of a population based inception cohort of inflammatory polyarthritis." Ann Rheum Dis **55**(2): 89-93.

Jung J, Song JJ and Kwon D (2009). "Allelic based gene-gene interactions in rheumatoid arthritis." BMC Proc **3 Suppl 7**: S76.

Kalla AA and Tikly M (2003). "Rheumatoid arthritis in the developing world." Best Pract Res Clin Rheumatol **17**(5): 863-75.

Kallberg H, Padyukov L, Plenge RM, et al. (2007). "Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis." American journal of human genetics **80**(5): 867-75.

Kalliomaki L (1954). "Correlation of the erythrocyte sedimentation rate and gold complications in rheumatoid arthritis." Ann Rheum Dis **13**(4): 336-7.

Kapitany A, Szabo Z, Lakos G, et al. (2008). "Associations between serum anti-CCP antibody, rheumatoid factor levels and HLA-DR4 expression in Hungarian patients with rheumatoid arthritis." Isr Med Assoc J **10**(1): 32-6.

Karlson EW, Chibnik LB, Cui J, et al. (2008). "Associations between human leukocyte antigen, PTPN22, CTLA4 genotypes and rheumatoid arthritis phenotypes of autoantibody status, age at diagnosis and erosions in a large cohort study." Ann Rheum Dis **67**(3): 358-63.

Kastbom A, Strandberg G, Lindroos A and Skogh T (2004). "Anti-CCP antibody test predicts the disease course during 3 years in early rheumatoid arthritis (the Swedish TIRA project)." Ann Rheum Dis **63**(9): 1085-9.

Kawasaki Y, Kato Y, Sai Y and Tsuji A (2004). "Functional characterization of human organic cation transporter OCTN1 single nucleotide polymorphisms in the Japanese population." J Pharm Sci **93**(12): 2920-6.

Kawashiri SY, Kawakami A, Ueki Y, et al. (2010). "Decrement of serum cartilage oligomeric matrix protein (COMP) in rheumatoid arthritis (RA) patients achieving remission after 6 months of etanercept treatment: Comparison with CRP, IgM-RF, MMP-3 and anti-CCP Ab." Joint Bone Spine **77**(5): 418-20.

Kim HY, Kim TG, Park SH, Lee SH, Cho CS and Han H (1995). "Predominance of HLA-DRB1*0405 in Korean patients with rheumatoid arthritis." Ann Rheum Dis **54**(12): 988-90.

Kim YO, Kim HJ, Kim SK, Chung J-H and Hong S-J (2010). "Association of the CD28/CTLA4/ICOS polymorphisms with susceptibility to rheumatoid arthritis." Clin Chem Lab Med **48**(3): 345-53.

Kinloch A, Lundberg K, Wait R, et al. (2008). "Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis." Arthritis & Rheumatism **58**(8): 2287-95.

Kochi Y, Okada Y, Suzuki A, et al. (2010). "A regulatory variant in CCR6 is associated with rheumatoid arthritis susceptibility." Nature Genet **42**(6): 515-9.

Kokkonen H, Johansson M, Innala L, Eriksson C, Jidell E and Rantapaa Dahlqvist S (2007). "The PTPN22 1858C/T polymorphism is associated with anti-cyclic citrullinated peptide antibody positive early rheumatoid arthritis in northern Sweden." Arthritis Res Ther **9**: R56.

Kurreeman FA, Daha NA, Chang M, et al. (2009). "Association of IL2RA and IL2RB with rheumatoid arthritis: a replication study in a Dutch population." Ann Rheum Dis **68**(11): 1789-90.

Kuwahara M, Ikari K, Momohara S, et al. (2005). "Failure to confirm association between SLC22A4 polymorphism and rheumatoid arthritis in a Japanese population." Arthritis Rheum **52**(9): 2947-8.

Laasila K, Laasonen L and Leirisalo-Repo M (2003). "Antibiotic treatment and long

term prognosis of reactive arthritis." Ann Rheum Dis **62**(7): 655-8.

Lau E SD, Bankhead C, MacGregor A, Donnan S, Silman A (1993). "Low prevalence of rheumatoid arthritis in the urbanized Chinese of Hong Kong." J Rheumatol **20**(7): 1133-7.

Lau E, Symmons D, Bankhead C, Macgregor A, Donnan S and Silman A (1993). "Low prevalence of rheumatoid arthritis in the urbanized Chinese of Hong Kong." J Rheumatol **20**: 1133 - 7.

Lawrence JS (1967). "Genetics of rheumatoid factor and rheumatoid arthritis." Clin Exp Immunol **2**(2(Suppl)): 769–83.

Lee AT, Li W, Liew A, et al. (2005). "The PTPN22 R620W polymorphism associates with RF positive rheumatoid arthritis in a dose-dependent manner but not with HLA-SE status." Genes Immun **6**: 129 - 33.

Lee DM and Schur PH (2003). "Clinical utility of the anti-CCP assay in patients with rheumatic diseases." Ann Rheum Dis **62**(9): 870-4.

Lee HS, Lee AT, Criswell LA, et al. (2008). "Several regions in the major histocompatibility complex confer risk for anti-CCP-antibody positive rheumatoid arthritis, independent of the DRB1 locus." Mol Med **14**(5-6): 293-300.

Lee HS, Remmers EF, Le JM, Kastner DL, Bae SC and Gregersen PK (2007). "Association of STAT4 with rheumatoid arthritis in the Korean population." Mol Med **13**(9-10): 455-60.

Lettre G, Lange C and Hirschhorn JN (2007). "Genetic model testing and statistical power in population-based association studies of quantitative traits." Genetic epidemiology **31**(4): 358-62.

Li EK (2002). Introduction to Rheumatology. Principles and Practice of Clinical Medicine in Asia. Sung J, Wong L, Li P, Sanderson J and Kwok T. Hong Kong, Lippincott Williams & Wilkins.

Li P, Cao C, Luan H, et al. (2010). "Association of genetic variations in the STAT4 and IRF7/KIAA1542 regions with systemic lupus erythematosus in a Northern Han Chinese population." Hum Immunol.

Li Y, Chang M, Schrodi SJ, et al. (2008). "The 5q31 variants associated with psoriasis and Crohn's disease are distinct." Hum Mol Genet **17**(19): 2978-85.

Liao J, Ip WS, Cheung KY, et al. (2011). "Diagnostic utility of an anti-CCP point-of-care immunotest in Chinese patients with rheumatoid arthritis." Clinica chimica acta; international journal of clinical chemistry **412**(9-10): 778-81.

Liao J, Ip WS, Cheung KY, et al. (2011). "Diagnostic utility of an anti-CCP point-of-care immunotest in Chinese patients with rheumatoid arthritis." Clin Chim Acta.

Linn-Rasker SP, van der Helm-van Mil AH, van Gaalen FA, et al. (2006). "Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles." Ann Rheum Dis **65**(3): 366-71.

Lipsky PE, van der Heijde DM, St Clair EW, et al. (2000). "Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group." N Engl J Med **343**(22): 1594-602.

Liu C, Ackerman HH and Carulli JP (2011). "A genome-wide screen of gene-gene interactions for rheumatoid arthritis susceptibility." Hum Genet.

Liu C, Batliwalla F, Li W, et al. (2008). "Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in rheumatoid arthritis." Mol Med **14**(9-10): 575-81.

Luan H, Li P, Cao C, et al. (2011). "A single-nucleotide polymorphism of the STAT4 gene is associated with systemic lupus erythematosus (SLE) in female Chinese population." Rheumatol Int.

MacGregor AJ, Riste LK, Hazes JM and Silman AJ (1994). "Low prevalence of rheumatoid arthritis in black-Caribbeans compared with whites in inner city Manchester." Ann Rheum Dis **53**(5): 293-7.

Mackenzie AH (1988). "Differential diagnosis of rheumatoid arthritis." Am J Med **85**(4A): 2-11.

Maddison PJ and Huey P (2004). Serological profile. Oxford testbook of Rheumatology. Isenberg DA, Maddison PJ, Woo P, Glass D and Breedveld F. Oxford, Oxford University Press.

Maeda T, Hirayama M, Kobayashi D, Miyazawa K and Tamai I (2007). "Mechanism of the regulation of organic cation/carnitine transporter 1 (SLC22A4) by rheumatoid arthritis-associated transcriptional factor RUNX1 and inflammatory cytokines." Drug Metab Dispos **35**(3): 394-401.

Maksymowych WP, Gabriel CA, Luyrink L, et al. (1992). "Polymorphism in a T-cell receptor variable gene is associated with susceptibility to a juvenile rheumatoid arthritis subset." Immunogenetics **35**(4): 257-62.

Malaviya AN, Kapoor SK, Singh RR, Kumar A and Pande I (1993). "Prevalence of rheumatoid arthritis in the adult Indian population." Rheumatol Int **13**(4): 131-4.

Manning AK, Ngwa JS, Hendricks AE, et al. (2009). "Incorporating biological knowledge in the search for gene x gene interaction in genome-wide association studies." BMC Proc **3 Suppl 7**: S81.

Marsters SA, Ayres TM, Skubatch M, Gray CL, Rothe M and Ashkenazi A (1997). "Herpesvirus Entry Mediator, a Member of the Tumor Necrosis Factor Receptor (TNFR) Family, Interacts with Members of the TNFR-associated Factor Family and Activates the Transcription Factors NF-κB and AP-1." J Biol Chem **272**(22): 14029-32.

Martinez A, Martin MC, Mendoza JL, et al. (2006). "Association of the organic cation transporter OCTN genes with Crohn's disease in the Spanish population." Eur J Hum Genet **14**(2): 222-6.

Martinez A, Valdivia A, Pascual-Salcedo D, et al. (2006). "Role of SLC22A4, SLC22A5, and RUNX1 genes in rheumatoid arthritis." J Rheumatol **33**(5): 842-6.

Matesanz F, Fernandez O and Alcina A (2007). "Genomewide study of multiple sclerosis." N Engl J Med **357**(21): 2200; author reply -1.

Mathur AN, Chang HC, Zisoulis DG, et al. (2007). "Stat3 and Stat4 direct development of IL-17-secreting Th cells." Journal of immunology **178**(8): 4901-7.

Meyer JM, Evans TI, Small RE, et al. (1999). "HLA-DRB1 genotype influences risk for and severity of rheumatoid arthritis." J Rheumatol **26**(5): 1024-34.

Michou L, Croiseau P, Petit-Teixeira E, et al. (2006). "Validation of the reshaped shared epitope HLA-DRB1 classification in rheumatoid arthritis." Arthritis Res Ther **8**(3): R79.

Minami Y, Kono T, Miyazaki T and Taniguchi T (1993). "The IL-2 receptor complex: its structure, function, and target genes." Annu Rev Immunol **11**: 245-68.

Miriovsky BJ, Michaud K, Thiele GM, et al. (2010). "Anti-CCP antibody and rheumatoid factor concentrations predict greater disease activity in men with rheumatoid arthritis." Ann Rheum Dis **69**(7): 1292-7.

Miyagi T, Gil MP, Wang X, Louten J, Chu WM and Biron CA (2007). "High basal STAT4 balanced by STAT1 induction to control type 1 interferon effects in natural killer cells." The Journal of experimental medicine **204**(10): 2383-96.

Mor-Vaknin N, Punturieri A, Sitwala K, et al. (2006). "The DEK nuclear autoantigen is a secreted chemotactic factor." Mol Cell Biol **26**(24): 9484-96.

Mori M, Yamada R, Kobayashi K, Kawaida R and Yamamoto K (2005). "Ethnic differences in allele frequency of autoimmune-disease-associated SNPs." J Hum

Genet **50**(5): 264-6.

Mori S, Naito H, Ohtani S, Yamanaka T and Sugimoto M (2009). "Diagnostic utility of anti-cyclic citrullinated peptide antibodies for rheumatoid arthritis in patients with active lung tuberculosis." Clin Rheumatol **28**(3): 277-83.

Murray KJ, Szer W, Grom AA, et al. (1997). "Antibodies to the 45 kDa DEK nuclear antigen in pauciarticular onset juvenile rheumatoid arthritis and iridocyclitis: selective association with MHC gene." J Rheumatol **24**(3): 560-7.

Mutru O, Laakso M, Isomaki H and Koota K (1989). "Cardiovascular mortality in patients with rheumatoid arthritis." Cardiology **76**(1): 71-7.

Myllykangas-Luosujarvi R, Aho K, Kautiainen H and Isomaki H (1995). "Shortening of life span and causes of excess mortality in a population-based series of subjects with rheumatoid arthritis." Clin Exp Rheumatol **13**(2): 149-53.

Namjou B, Sestak AL, Armstrong DL, et al. (2009). "High-density genotyping of STAT4 reveals multiple haplotypic associations with systemic lupus erythematosus in different racial groups." Arthritis Rheum **60**(4): 1085-95.

Newman B, Wintle RF, van Oene M, et al. (2005). "SLC22A4 polymorphisms implicated in rheumatoid arthritis and Crohn's disease are not associated with rheumatoid arthritis in a Canadian Caucasian population." Arthritis Rheum **52**(2): 425-9.

Nienhuis RL and Mandema E (1964). "A New Serum Factor in Patients with Rheumatoid Arthritis; the Antiperinuclear Factor." Ann Rheum Dis **23**: 302-5.

Niewold TB, Harrison MJ and Paget SA (2007). "Anti-CCP antibody testing as a diagnostic and prognostic tool in rheumatoid arthritis." QJM **100**(4): 193-201.

Nishikomori R, Usui T, Wu CY, Morinobu A, O'Shea JJ and Strober W (2002). "Activated STAT4 has an essential role in Th1 differentiation and proliferation that is independent of its role in the maintenance of IL-12R beta 2 chain expression and signaling." Journal of immunology **169**(8): 4388-98.

Nyholt DR (2004). "A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other." Am J Hum Genet **74**(4): 765-9.

Ohmura K, Terao C, Maruya E, et al. (2010). "Anti-citrullinated peptide antibody-negative RA is a genetically distinct subset: a definitive study using only bone-erosive ACPA-negative rheumatoid arthritis." Rheumatology (Oxford) **49**(12): 2298-304.

Okada Y, Mori M, Yamada R, et al. (2008). "SLC22A4 polymorphism and rheumatoid arthritis susceptibility: a replication study in a Japanese population and a metaanalysis." J Rheumatol **35**(9): 1723-8.

Oliver JE and Silman AJ (2006). "Risk factors for the development of rheumatoid arthritis." Scand J Rheumatol **35**(3): 169-74.

Orces CHDR, I. Abel, M. P. Escalante, A. (2002). "The number of deformed joints as a surrogate measure of damage in rheumatoid arthritis." Arthritis Rheum **47**(1): 67-72.

Ostergaard M, Duer A, Moller U and Ejbjerg B (2004). "Magnetic resonance imaging of peripheral joints in rheumatic diseases." Best Pract Res Clin Rheumatol **18**(6): 861-79.

Ostergaard M and Szkudlarek M (2003). "Imaging in rheumatoid arthritis--why MRI and ultrasonography can no longer be ignored." Scand J Rheumatol **32**(2): 63-73.

Ota N, Brett TJ, Murphy TL, Fremont DH and Murphy KM (2004). "N-domain-dependent nonphosphorylated STAT4 dimers required for cytokine-driven activation." Nature immunology **5**(2): 208-15.

Panayi GS and Corrigall VM (2006). "BiP regulates autoimmune inflammation and tissue damage." Autoimmun Rev **5**(2): 140-2.

Papadopoulos IA, Katsimbri P, Alamanos Y, Voulgari PV and Drosos AA (2003). "Early rheumatoid arthritis patients: relationship of age." Rheumatol Int **23**(2): 70-4.

Patsopoulos NA and Ioannidis JP (2010). "Susceptibility variants for rheumatoid arthritis in the TRAF1-C5 and 6q23 loci: a meta-analysis." Ann Rheum Dis **69**(3): 561-6.

Pedersen JK, Kjaer NK, Svendsen AJ and Horslev-Petersen K (2009). "Incidence of rheumatoid arthritis from 1995 to 2001: impact of ascertainment from multiple sources." Rheumatol Int **29**(4): 411-5.

Perdigones N, Vigo AG, Lamas JR, et al. (2010). "Evidence of epistasis between TNFRSF14 and TNFRSF6B polymorphisms in patients with rheumatoid arthritis." Arthritis Rheum **62**(3): 705-10.

Pincus T and Callahan LF (1985). "Formal education as a marker for increased mortality and morbidity in rheumatoid arthritis." J Chronic Dis **38**(12): 973-84.

Pincus TS, C. Wolfe, F. (1999). "Toward a multidimensional Health Assessment Questionnaire (MDHAQ): assessment of advanced activities of daily living and psychological status in the patient-friendly health assessment questionnaire format."

Arthritis Rheum **42**(10): 2220-30.

Plenge RM, Cotsapas C, Davies L, et al. (2007). "Two independent alleles at 6q23 associated with risk of rheumatoid arthritis." Nature Genet **39**(12): 1477-82.

Plenge RM, Padyukov L, Remmers EF, et al. (2005). "Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4." Am J Hum Genet 77(6): 1044-60.

Plenge RM, Seielstad M, Padyukov L, et al. (2007). "TRAF1-C5 as a risk locus for rheumatoid arthritis--a genomewide study." The New England journal of medicine **357**(12): 1199-209.

Plenge RM, Seielstad M, Padyukov L, et al. (2007). "TRAF1-C5 as a Risk Locus for Rheumatoid Arthritis -- A Genomewide Study." N Engl J Med: NEJMoa073491.

Potter C, Hyrich KL, Tracey A, et al. (2009). "Association of rheumatoid factor and anti-cyclic citrullinated peptide positivity, but not carriage of shared epitope or PTPN22 susceptibility variants, with anti-tumour necrosis factor response in rheumatoid arthritis." Ann Rheum Dis **68**(1): 69-74.

Pritchard MH (1992). "An examination of the role of female hormones and pregnancy as risk factors for rheumatoid arthritis, using a male population as control group." Br J Rheumatol **31**(6): 395-9.

Purcell AW TA, Kinoshita G, Lynch TA, Keech CL, Gething MJ, Gordon TP. (2003). "Association of stress proteins with autoantigens: a possible mechanism for triggering autoimmunity?" Clin Exp Immunol **132**(2): 193-200.

Purcell S, Neale B, Todd-Brown K, et al. (2007). "PLINK: a tool set for whole-genome association and population-based linkage analyses." Am J Hum Genet **81**(3): 559-75.

Qiao B, Huang CH, Cong L, Xie J, Lo SH and Zheng T (2009). "Genome-wide gene-based analysis of rheumatoid arthritis-associated interaction with PTPN22 and HLA-DRB1." BMC Proc **3 Suppl** 7: S132.

Quinn MA, Gough AK, Green MJ, et al. (2006). "Anti-CCP antibodies measured at disease onset help identify seronegative rheumatoid arthritis and predict radiological and functional outcome." Rheumatology (Oxford) **45**(4): 478-80.

Ramos-Casals M, Brito-Zeron P, Perez-De-Lis M, et al. (2010). "Sjogren syndrome or sjogren disease? The histological and immunological bias caused by the 2002 criteria." Clin Rev Allergy Immunol **38**(2-3): 178-85.

Remmers EF, Plenge RM, Lee AT, et al. (2007). "STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus." N Engl J Med **357**(10): 977-86.

Rodriguez-Rodriguez L, Taib WR, Topless R, et al. (2011). "The PTPN22 R263Q polymorphism is a risk factor for rheumatoid arthritis in Caucasian case-control samples." Arthritis Rheum **63**(2): 365-72.

Roitt IM, Corbett M, Festenstein H, et al. (1978). "HLA-DRW4 AND PROGNOSIS IN RHEUMATOID ARTHRITIS." The Lancet **311**(8071): 990-.

Rommel C, Camps M and Ji H (2007). "PI3K delta and PI3K gamma: partners in crime in inflammation in rheumatoid arthritis and beyond?" Nat Rev Immunol **7**(3): 191-201.

Ronday HK, van der Laan WH, Tak PP, et al. (2001). "Human granzyme B mediates cartilage proteoglycan degradation and is expressed at the invasive front of the synovium in rheumatoid arthritis." Rheumatology (Oxford) **40**(1): 55-61.

Rotter JI and Landaw EM (1984). "Measuring the genetic contribution of a single locus to a multilocus disease." Clin Genet **26**(6): 529-42.

Samanci N, Ozdem S, Akbas H, et al. (2005). "Diagnostic value and clinical significance of anti-CCP in patients with advanced rheumatoid arthritis." J Natl Med Assoc **97**(8): 1120-6.

Sauerland U, Becker H, Seidel M, et al. (2005). "Clinical utility of the anti-CCP assay: experiences with 700 patients." Ann N Y Acad Sci **1050**: 314-8.

Schellekens GA, Visser H, de Jong BA, et al. (2000). "The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide." Arthritis Rheum **43**(1): 155-63.

Shibata Y, Tsukazaki T, Hirata K, Xin C and Yamaguchi A (2004). "Role of a new member of IGFBP superfamily, IGFBP-rP10, in proliferation and differentiation of osteoblastic cells." Biochem Biophys Res Commun **325**(4): 1194-200.

Shichikawa K, Inoue K, Hirota S, et al. (1999). "Changes in the incidence and prevalence of rheumatoid arthritis in Kamitonda, Wakayama, Japan, 1965-1996." Ann Rheum Dis **58**(12): 751-6.

Shimane K, Kochi Y, Horita T, et al. (2010). "The association of a nonsynonymous single-nucleotide polymorphism in TNFAIP3 with systemic lupus erythematosus and rheumatoid arthritis in the Japanese population." Arthritis Rheum **62**(2): 574-9.

Silveira IG, Burlingame RW, von Muhlen CA, Bender AL and Staub HL (2007). "Anti-CCP antibodies have more diagnostic impact than rheumatoid factor (RF) in a

population tested for RF." Clin Rheumatol **26**(11): 1883-9.

Simon M, Girbal E, Sebbag M, et al. (1993). "The cytokeratin filament-aggregating protein filaggrin is the target of the so-called "antikeratin antibodies," autoantibodies specific for rheumatoid arthritis." J Clin Invest **92**(3): 1387-93.

Simonsson M, Bergman S, Jacobsson LT, Petersson IF and Svensson B (1999). "The prevalence of rheumatoid arthritis in Sweden." Scand J Rheumatol **28**(6): 340-3.

Skriner K, Steiner G, Sommergruber WH, Sinski A and Smolen JS (1994). "Anti-RA33 autoantibodies may recognize epitopes in the N-terminal region of hnRNP-A2 (RA33)." Clin Exp Rheumatol **12 Suppl 11**: S79-82.

Stahl EA, Raychaudhuri S, Remmers EF, et al. (2010). "Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci." Nature Genet **42**(6): 508-14.

Stastny P (1978). "Association of the B-Cell Alloantigen DRw4 with Rheumatoid Arthritis." New England Journal of Medicine **298**(16): 869-71.

Stastny P (1978). "A complex of HLA-D specificities detected by HTC typing: Dw7, Dw11, and TMo." Transplant Proc **10**(4): 759-61.

Stastny P (1978). "HLA-D and Ia antigens in rheumatoid arthritis and systemic lupus erythematosus." Arthritis Rheum **21**(5 Suppl): S139-43.

Steinbrocker O, Traeger CH and Batterman RC (1949). "Therapeutic criteria in rheumatoid arthritis." JAMA **140**: 659-62.

Steiner G, Skriner K, Hassfeld W and JS. S (1996). "Clinical and immunological aspects of autoantibodies to RA33/hnRNP-A/B proteins--a link between RA, SLE and MCTD." Mol Bio Rep **23**(3-4): 167-71.

Steiner. G and JS S (2006). Autoantibodies in rheumatoid arthritis: prevalence and clinical significance. Rheumatoid Arthritis. Firestein GS, Pan G and Wollheim FA. Oxford, Oxford University Press.

Steuer A, Watkins J, Smith F, Day L, Demetriadi F and Chapel H (2008). "RF latex and anti-CCP antibodies: a combined strategy for diagnosing RA in primary care?" Rheumatology (Oxford) **47**(3): 375-6.

Sun KH, Tang SJ, Wang YS, Lin WJ and You RI (2003). "Autoantibodies to dsDNA cross-react with the arginine-glycine-rich domain of heterogeneous nuclear ribonucleoprotein A2 (hnRNP A2) and promote methylation of hnRNP A2." Rheumatology **42**(1): 154-61.

Suzuki A, Yamada R, Chang X, et al. (2003). "Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis." Nature Genet **34**: 395 - 402.

Symmons DPM (2002). "Epidemiology of rheumatoid arthritis: determinants of onset, persistence and outcome." Best Pract Res Clin Rheumatol **16**(5): 707-22.

Syversen SW, Goll GL, van der Heijde D, et al. (2010). "Prediction of radiographic progression in rheumatoid arthritis and the role of antibodies against mutated citrullinated vimentin: results from a 10-year prospective study." Ann Rheum Dis **69**(2): 345-51.

Takata Y, Inoue H, Sato A, et al. (2008). "Replication of reported genetic associations of PADI4, FCRL3, SLC22A4 and RUNX1 genes with rheumatoid arthritis: results of an independent Japanese population and evidence from meta-analysis of East Asian studies." J Hum Genet **53**(2): 163-73.

Tamai M, Kawakami A, Iwamoto N, Arima K, Aoyagi K and Eguchi K (2010). "Contribution of anti-CCP antibodies, proximal interphalangeal joint involvement, HLA-DRB1 shared epitope, and PADI4 as risk factors for the development of rheumatoid arthritis in palindromic rheumatism." Scand J Rheumatol **39**(4): 287-91.

Tanaka G, Matsushita I, Ohashi J, et al. (2005). "Evaluation of microsatellite markers in association studies: a search for an immune-related susceptibility gene in sarcoidosis." Immunogenetics **56**(12): 861-70.

Taneja V, Giphart MJ, Verduijn W, Naipal A, Malaviya AN and Mehra NK (1996). "Polymorphism of HLA-DRB, -DQA1, and -DQB1 in Rheumatoid Arthritis in Asian Indians : Association with DRB1*0405 and DRB1*1001." Human Immunology **46**(1): 35-41.

Tang NL, Liao CD, Ching JK, et al. (2010). "Sex-specific effect of Pirin gene on bone mineral density in a cohort of 4000 Chinese." Bone **46**(2): 543-50.

Thierfelder WE, van Deursen JM, Yamamoto K, et al. (1996). "Requirement for Stat4 in interleukin-12-mediated responses of natural killer and T cells." Nature **382**(6587): 171-4.

Thompson SD, Sudman M, Ramos PS, et al. (2010). "The susceptibility loci juvenile idiopathic arthritis shares with other autoimmune diseases extend to PTPN2, COG6, and ANGPT1." Arthritis and rheumatism **62**(11): 3265-76.

Tighe H, Silverman GJ, Kozin F, et al. (1990). "Autoantibody production by severe combined immunodeficient mice reconstituted with synovial cells from rheumatoid arthritis patients." Eur J Immunol **20**(8): 1843-8.

Tokuhiro S, Yamada R, Chang X, et al. (2003). "An intronic SNP in a RUNX1

binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis." Nat Genet **35**(4): 341-8.

Tokuhiro SY, R. Chang, X. Suzuki, A. Kochi, Y. Sawada, T. Suzuki, M. Nagasaki, M. Ohtsuki, M. Ono, M. Furukawa, H. Nagashima, M. Yoshino, S. Mabuchi, A. Sekine, A. Saito, S. Takahashi, A. Tsunoda, T. Nakamura, Y. Yamamoto, K. (2003). "An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis." Nat Genet **35**(4): 341-8.

Trembleau S, Hoffmann M, Meyer B, et al. (2010). "Immunodominant T-cell epitopes of hnRNP-A2 associated with disease activity in patients with rheumatoid arthritis." Eur J Immunol **40**(6): 1795-808.

Tsuchiya N, Ito I and Kawasaki A (2010). "Association of IRF5, STAT4 and BLK with systemic lupus erythematosus and other rheumatic diseases." Jpn J Clin Immunol **33**(2): 57-65.

Tyler DR, Persky ME, Matthews LA, Chan S and Farrar JD (2007). "Pre-assembly of STAT4 with the human IFN-alpha/beta receptor-2 subunit is mediated by the STAT4 N-domain." Molecular immunology **44**(8): 1864-72.

Uhlig T, Heiberg T, Mowinckel P and Kvien TK (2008). "Rheumatoid arthritis is milder in the new millennium: health status in patients with rheumatoid arthritis 1994-2004." Ann Rheum Dis **67**(12): 1710-5.

Ungureanu D and Silvennoinen O (2005). "SLIM trims STATs: ubiquitin E3 ligases provide insights for specificity in the regulation of cytokine signaling." Sci STKE **2005**(304): pe49.

van der Heijde DM, van't Hof MA, van Riel PL, van Leeuwen MA, van Rijswijk MH and van de Putte LB (1992). "Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis." Ann Rheum Dis **51**(2): 177-81.

van der Helm-van Mil AH, Toes REM and Huizinga TWJ (2010). "Genetic variants in the prediction of rheumatoid arthritis." Ann Rheum Dis **69**(9): 1694-6.

van Gaalen FA, Linn-Rasker SP, van Venrooij WJ, et al. (2004). "Autoantibodies to cyclic citrullinated peptides predict progression to rheumatoid arthritis in patients with undifferentiated arthritis: a prospective cohort study." Arthritis Rheum **50**(3): 709-15.

Vandenbroucke JP, Hazevoet HM and Cats A (1984). "Survival and cause of death in rheumatoid arthritis: a 25-year prospective followup." J Rheumatol **11**(2): 158-61.

Vella A, Cooper, J. D. Lowe, C. E. Walker, N. Nutland, S. Widmer, B. Jones, R. Ring, S. M. McArdle, W. Pembrey, M. E. Strachan, D. P. Dunger, D. B. Twells, R.C.J.

Clayton, D. G. Todd, J.A. (2005). "Localization of a Type 1 Diabetes Locus in the IL2RA/CD25 Region by Use of Tag Single-Nucleotide Polymorphisms." Am J Hum Genet **76**(5): 773-9.

Vinkemeier U, Moarefi I, Darnell JE, Jr. and Kuriyan J (1998). "Structure of the amino-terminal protein interaction domain of STAT-4." Science **279**(5353): 1048-52.

Vittecoq O, Incaurgarat B, Jouen-Beades F, et al. (2004). "Autoantibodies recognizing citrullinated rat filaggrin in an ELISA using citrullinated and non-citrullinated recombinant proteins as antigens are highly diagnostic for rheumatoid arthritis." Clin Exp Immunol **135**(1): 173-80.

Waaler E (1939). "On the occurrence of a factor in human serum activating the specific agglutination of sheep blood corpuscles." Acta Pathol Microbiol Scand C **17**(172-182).

Wang J, Chuang K, Ahluwalia M, et al. (2005). "High-throughput SNP genotyping by single-tube PCR with Tm-shift primers." Biotechniques **39**(6): 885-93.

Watford WT, Hissong BD, Bream JH, Kanno Y, Muul L and O'Shea JJ (2004). "Signaling by IL-12 and IL-23 and the immunoregulatory roles of STAT4." Immunological Reviews **202**: 139-56.

Wernhoff P, Olofsson, P. and Holmdahl, R (2003). "The genetic control of rheumatoid factor production in a rat model of rheumatoid arthritis." Arthritis Rheum **48**.

Wiell C, Szkudlarek M, Hasselquist M, et al. (2007). "Ultrasonography, magnetic resonance imaging, radiography, and clinical assessment of inflammatory and destructive changes in fingers and toes of patients with psoriatic arthritis." Arthritis Res Ther **9**(6): R119.

Witte T, Hartung K, Sachse C, et al. (2000). "Rheumatoid factors in systemic lupus erythematosus: association with clinical and laboratory parameters. SLE study group." Rheumatol Int **19**(3): 107-11.

Woo J, Suen EW, Leung JC, Tang NL and Ebrahim S (2009). "Older men with higher self-rated socioeconomic status have shorter telomeres." Age Ageing **38**(5): 553-8.

Woo J, Tang N and Leung J (2008). "No association between physical activity and telomere length in an elderly Chinese population 65 years and older." Arch Intern Med **168**(19): 2163-4.

WTCCC (2007). "Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls." Nature **447**: 661 - 78.

Yaich L, Dupont WD, Cavener DR and Parl FF (1992). "Analysis of the PvuII restriction fragment-length polymorphism and exon structure of the estrogen receptor gene in breast cancer and peripheral blood." Cancer Res **52**(1): 77-83.

Yamada R, Tokuhiro S, Chang X and Yamamoto K (2004). "SLC22A4 and RUNX1: identification of RA susceptible genes." J Mol Med **82**(9): 558-64.

Yamazaki K, Takazoe M, Tanaka T, et al. (2004). "Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn disease." J Hum Genet **49**(12): 664-8.

Yoshimoto T, Furuhata M, Kamiya S, et al. (2003). "Positive modulation of IL-12 signaling by sphingosine kinase 2 associating with the IL-12 receptor beta 1 cytoplasmic region." Journal of immunology **171**(3): 1352-9.

Young AD, J. Cox, N. Davies, P. Devlin, J. Emery, P. Gallivan, S. Gough, A. James, D. Prouse, P. Williams, P. Winfield, J. (2000). "How does functional disability in early rheumatoid arthritis (RA) affect patients and their lives? Results of 5 years of follow-up in 732 patients from the Early RA Study (ERAS)." Rheumatology (Oxford) **39**(6): 603-11.

Yukioka M, Wakitani S, Murata N, et al. (1998). "Elderly-onset rheumatoid arthritis and its association with HLA-DRB1 alleles in Japanese." Br J Rheumatol **37**(1): 98-101.

Zeng Q, Chen R, Darmawan J, et al. (2008). "Rheumatic Diseases in China." Arthritis Res Ther **10**(1): R17.

Zhang M, Lin Y, Wang L, Pungpapong V, Fleet JC and Zhang D (2009). "Case-control genome-wide association study of rheumatoid arthritis from Genetic Analysis Workshop 16 using penalized orthogonal-components regression-linear discriminant analysis." BMC Proc **3 Suppl 7**: S17.

Zhang SS, Welte T and Fu XY (2001). "Dysfunction of Stat4 leads to accelerated incidence of chemical-induced thymic lymphomas in mice." Exp Mol Pathol **70**(3): 231-8.

Zhao J, Liu X, Wang Z and Li Z (2010). "Significance of anti-CCP antibodies in modification of 1987 ACR classification criteria in diagnosis of rheumatoid arthritis." Clin Rheumatol **29**(1): 33-8.

Zhao J, Zhao YI, He J, Jia R and Li Z (2008). "Prevalence and Significance of Anti-Peptidylarginine Deiminase 4 Antibodies in Rheumatoid Arthritis." J Rheumatol **35**(6): 969-74.

Zhernakova A, Eerligh P, Wijmenga C, Barrera P, Roep BO and Koeleman BP (2005). "Differential association of the PTPN22 coding variant with autoimmune diseases in

a Dutch population." Genes Immun **6**: 459 - 61.

Zhernakova A, Stahl EA, Trynka G, et al. (2011). "Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci." PLoS genetics **7**(2): e1002004.

Zlabinger GJ, Haberhauer G, Dax K, Menzel EJ and Broll H (1990). "Rheumatoid factor isotypes and circulating immune complexes in rheumatoid arthritis." Clin Exp Rheumatol **8**(2): 113-9.