Presence of immunological markers preceding the onset of rheumatoid arthritis

Mikael Brink
Figure 1  Reproduction, based partly on Fig 2 in (Simmonds & Gough, 2005)
Figure 3  Reproduction, based partly on Fig 1A in (Shi et al., 2011)
Figure 4  Reproduction, based partly on Fig 1B in (Shi et al., 2011)
Figure 5  Reproduction, based partly on Fig 1 in (Strand et al., 2007)
To my family
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>v</td>
</tr>
<tr>
<td>List of papers</td>
<td>vi</td>
</tr>
<tr>
<td>Sammanfattning på svenska</td>
<td>vii</td>
</tr>
<tr>
<td><strong>Background</strong></td>
<td>1</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>1</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>1</td>
</tr>
<tr>
<td>Diagnosis and classification</td>
<td>1</td>
</tr>
<tr>
<td>Pharmacological treatment</td>
<td>3</td>
</tr>
<tr>
<td>Radiology</td>
<td>4</td>
</tr>
<tr>
<td>Aetiology and risk factors for the development of rheumatoid arthritis</td>
<td>4</td>
</tr>
<tr>
<td>Genetic influence in rheumatoid arthritis</td>
<td>4</td>
</tr>
<tr>
<td>Environmental factors</td>
<td>6</td>
</tr>
<tr>
<td>Occupation and occupational exposures</td>
<td>7</td>
</tr>
<tr>
<td>Infections and infectious agents</td>
<td>7</td>
</tr>
<tr>
<td>Hormonal factors</td>
<td>8</td>
</tr>
<tr>
<td>The immune system</td>
<td>8</td>
</tr>
<tr>
<td>Autoantibodies in rheumatoid arthritis</td>
<td>10</td>
</tr>
<tr>
<td>Rheumatoid factor</td>
<td>10</td>
</tr>
<tr>
<td>Anti-citrullinated protein/peptide antibodies and peptidylarginine deiminase</td>
<td>11</td>
</tr>
<tr>
<td>Anti- fibrinogen antibodies</td>
<td>11</td>
</tr>
<tr>
<td>Anti- vimentin antibodies</td>
<td>12</td>
</tr>
<tr>
<td>Anti-cyclic citrullinated peptide antibodies</td>
<td>12</td>
</tr>
<tr>
<td>Anti- collagen antibodies</td>
<td>12</td>
</tr>
<tr>
<td>Anti- α-enolase antibodies</td>
<td>13</td>
</tr>
<tr>
<td>Peptidylarginine deiminase</td>
<td>13</td>
</tr>
<tr>
<td>Anti- carbamylated peptides/protein antibodies</td>
<td>14</td>
</tr>
<tr>
<td>Anti- CarP antibodies</td>
<td>15</td>
</tr>
<tr>
<td>Timing of occurrence of antibodies in the pre-symptomatic period</td>
<td>16</td>
</tr>
<tr>
<td>Pathogenesis</td>
<td>16</td>
</tr>
<tr>
<td>The arthritic joint</td>
<td>17</td>
</tr>
<tr>
<td>Seropositive vs. Seronegative rheumatoid arthritis</td>
<td>18</td>
</tr>
<tr>
<td><strong>Aims</strong></td>
<td>19</td>
</tr>
<tr>
<td><strong>Study populations</strong></td>
<td>20</td>
</tr>
<tr>
<td>Patients with early and established RA</td>
<td>20</td>
</tr>
<tr>
<td>The Medical Biobank of Northern Sweden</td>
<td>20</td>
</tr>
<tr>
<td>Västerbotten Intervention Programme (VIP)</td>
<td>21</td>
</tr>
<tr>
<td>The Northern Sweden MONICA project</td>
<td>21</td>
</tr>
<tr>
<td>The Mammary Screening Programme</td>
<td>21</td>
</tr>
</tbody>
</table>
Collection of samples within the NSHDS 21
The Maternity Cohort 22
Identification of pre-symptomatic individuals 22
Identification of population control subjects 22
Paper I-III 22
Multi RA-case family material 23
Paper IV 23

**Methods**
- Autoantibody analyses 25
- Multiplex analysis of ten antibodies against citrullinated peptides 25
- Anti-CCP2 antibodies 26
- Anti-CarP antibodies 26
- Rheumatoid factor 26
- Radiology 26
- Detection of cell subsets and populations 27
- Flow cytometry 27
- Detection of cytokines and chemokines 27
- Genetic analyses 27
- Statistics 28

**Results and discussion**
- Paper I-III 30
- Paper IV 45

**Concluding remarks**
- Conclusions 49

**Future perspectives**
- Acknowledgements 53

**References**
Abstract

Rheumatoid arthritis (RA) is a chronic inflammatory disease with an unknown aetiology characterized by joint destruction. Both genetic and environmental factors contribute to the disease development with HLA-DRB1* alleles and smoking identified as most important. The disease is characterized by the presence of autoantibodies, originally by rheumatoid factor (RF) and more recently by anti citrullinated protein/peptide antibodies (ACPA) and antibodies against carbamylated peptides (CarP). These autoantibodies are present, not only after the onset of disease, but also prior to the onset of symptoms. The development of RA is a gradual process lasting several years before the onset of any joint symptom, but when and if there is a temporal difference in the development both between and within the different antibody systems is currently unknown. B-cells produce the antibodies, and a subset of B-cells, i.e., B-regulatory (Breg) cells, produces interleukin-10, and thus have the ability to down-regulate pro-inflammatory cytokines. Whether the Breg cells are involved in the pathogenesis of RA is, as yet, unknown.

The aim of this thesis was to increase knowledge of the pathophysiological processes in the development of RA through identification of factors involved. The analyses involved detection of autoantibodies to post-translationally modified peptides/proteins in addition to RF isotypes, cell surface markers on immune cells in asymptomatic individuals, who have an increased risk of developing RA. In a co-analysis of the registers of patients with RA attending the Department of Rheumatology, with the registers from population based screening programmes within the Biobank of Northern Sweden, blood samples collected from individuals prior to the onset of symptoms were identified, as were those from population control subjects. A cohort of pre-symptomatic individuals also donated samples at the time of receiving a diagnosis of RA. First-degree relatives (FDR) of patients with RA were also identified and included for analyses. The levels of ten different ACPAs, i.e., (fibrinogen (Fib) α563-583(573), Fibα580-600(591), Fibβ62-81a(72), Fibβ62-81b(74), Fibβ36-52, α-enolase (CEP-1), triple helical collagen type II (citC1III), filaggrin (Fil307-324), vimentin (Vim) 2-17, and Vim60-75) were measured using the ImmunoCAP ISAC system (Phadia/ThermoFischer, Uppsala, Sweden) in blood samples from individuals before the onset of symptoms and when diagnosed with RA in comparison with those in population based controls. In a subset of samples, the levels of anti-CarP antibodies were measured using ELISA coated with anti-CarP-FCS, as well as analysis of RF of IgM, IgG and IgA isotype using the EliA assay (Phadia, Uppsala, Sweden). Breg cells were analysed both with and without stimulation ex vivo along with other cell types using flow cytometry in samples from patients with RA, their first degree relatives (FDR) and healthy controls.
In paper I it was shown that levels of ACPA were initially restricted to a few antibodies but disseminated over time to involve additional different antibodies. The levels of antibodies to CEP-1, Fibβ36-52, and filaggrin were significantly increased. In paper II, anti-CarP antibodies were positive in 5-13% of the individuals negative for the various ACPA studied. The presence of anti-CarP antibodies was significantly related to radiological destruction of joints at baseline, at follow-up after 24 months and to the radiological progress between baseline and 24 months. In paper III, the relationships between the frequencies of RF isotypes, the ten different ACPA, anti-CCP2 and anti-CarP antibodies before the onset of any symptoms and the presence of certain combinations of antibodies were associated with a very high risk of developing RA. In paper IV, Breg cells from patients with RA are functionally impaired and FDR showed a similar pattern by responding less to stimulation ex vivo than cells from healthy controls.

In conclusion, individuals who subsequently develop RA have an increased number and amount of ACPAs, anti-CarP antibodies and RF of IgM, IgG and IgA isotype, several years before symptom onset. Most of the different antibodies analysed remain associated with disease development after adjustments for each separate antibody. In FDRs, Breg cells were functionally altered in that they produce less IL-10 and consequently contribute to a more inflammation-prone status, as in their relatives with RA. These findings contribute to information about the development of RA as well as a given individual’s risk(s) of developing RA and its progression.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACPA</td>
<td>Anti-citrullinated protein/peptide antibody</td>
</tr>
<tr>
<td>ACR</td>
<td>American College of Rheumatology</td>
</tr>
<tr>
<td>Anti-CCP</td>
<td>Anti-cyclic citrullinated peptide</td>
</tr>
<tr>
<td>BCR</td>
<td>B-cell receptor</td>
</tr>
<tr>
<td>Breg</td>
<td>B Regulatory cell</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CD</td>
<td>Cluster of differentiation</td>
</tr>
<tr>
<td>CRP</td>
<td>C reactive protein</td>
</tr>
<tr>
<td>DAS</td>
<td>Disease activity score</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>ESR</td>
<td>Erythrocyte sedimentation rate</td>
</tr>
<tr>
<td>EULAR</td>
<td>European league against rheumatism</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence Activated Cell Sorting</td>
</tr>
<tr>
<td>HLA-SE</td>
<td>Human leukocyte antigen-shared epitope allele</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>IQR</td>
<td>Inter quartile range</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>MCV</td>
<td>Mutated citrullinated vimentin</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MONICA</td>
<td>Monitoring of Trends and Determinants in Cardiovascular Disease</td>
</tr>
<tr>
<td>MS</td>
<td>Multiple sclerosis</td>
</tr>
<tr>
<td>NK</td>
<td>Natural killer</td>
</tr>
<tr>
<td>NSHDS</td>
<td>Northern Sweden Health and Disease Study</td>
</tr>
<tr>
<td>OR</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>PAD</td>
<td>Peptidyl arginine deiminase</td>
</tr>
<tr>
<td>PTPN22</td>
<td>Protein tyrosine phosphatase non-receptor type 22</td>
</tr>
<tr>
<td>RA</td>
<td>Rheumatoid arthritis</td>
</tr>
<tr>
<td>RF</td>
<td>Rheumatoid factor</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver operating characteristics</td>
</tr>
<tr>
<td>SLE</td>
<td>Systemic lupus erythematosus</td>
</tr>
<tr>
<td>TCR</td>
<td>T-cell receptor</td>
</tr>
<tr>
<td>Th</td>
<td>T helper cell</td>
</tr>
<tr>
<td>TNF</td>
<td>Tumour necrosis factor</td>
</tr>
<tr>
<td>Treg</td>
<td>Regulatory T-cell</td>
</tr>
<tr>
<td>VIP</td>
<td>Västerbotten intervention programme</td>
</tr>
</tbody>
</table>
List of papers


IV. **Brink M**, Ärlestig L, Rantapää-Dahlqvist S, Lejon K. B-regulatory cells are functionally impaired in patients with rheumatoid arthritis and in their first degree relatives. *Submitted.*

Papers I and II are reprinted with permission of the respective publisher.
Sammanfattning på svenska

Reumatoid artrit (RA), ledgångsreumatism, är en kronisk autoimmun sjukdom som drabbar strax under 1 % av den vuxna befolkningen, huvudsakligen kvinnor. Av ännu okänd anledning angriper immunförsvaret framförallt ledvävnad i händer och fötter, men alla leder kan angrivas och även andra organ såsom lungsäck, hjärtsäck eller blodkärl. Sjukdomens svårighetsgrad varierar. Vissa individer har ett snabbt sjukdomsförlopp med aggressiv sjukdom som bryter ner brosk och ben, vissa har ett lugnare förlopp. En tidig diagnos och snabbt insatt behandling är ett viktigt steg för att förhindra permanenta skador.

Vad som orsakar sjukdomen är till stor del okänt. Ärfliga faktorer, liksom miljöfaktorer (t.ex. rökning) bidrar till ökad risk att insjukna i RA.


Syftet med denna avhandling var att försöka identifiera faktorer av betydelse för utvecklandet av RA.

I delarbete I analyserades blodprover från Medicinska Biobanken i Umeå från 386 individer som donerat 717 prov innan insjuknandet i RA (pre-symptomatiska individer), för 204 av dem fanns även blodprov från diagnostillfället tillgängligt. Utöver dessa identifierades även 1305 kontrollindivider från Medicinska Biobanken i Umeå med liknande köns- och åldersfördelning som hos de pre-symptomatiska individerna. Alla prover analyserades för nivåer av tio olika ACPA samt anti-CCP2. Resultaten visade att fem av dessa ACPA var signifikant förhöjda redan innan symptomdebut jämfört med kontroller. Resultaten visade också att antalet ACPA ökar ju närmre symptomdebuten provet är taget.

I delarbete II analyserades en subgrupp av de individer som ingick i delarbete I. Blodproven analyserades för nivåer av en ny antikropp, antikroppar mot karbamylerade peptider/ proteiner (anti-CarP). Anti-CarP antikroppar påvisades i signifikant högre nivåer både innan sjukdomsdebut och vid diagnostillfället, jämfört med kontroller. För patienterna fanns även röntgenbedömningar av
hand- och fotleder. De individer där anti-CarP antikroppar påvisats innan symptomdebut visade högre grad av skadade leder vid diagnosstillfället.

I delarbete III analyserades en subgrupp av de individer som ingått i delarbete I där huvuddelen även ingått i delarbete II. Blodproven analyserades för RF av tre olika isotyper, varianter av immunglobuliner (Ig), IgA, IgG och IgM. Förekomsten av de tre RF-isotyperna relaterades till förekomsten av de tio ACPA och anti-CarP antikropparna. Samförekomst av flera av dessa innan symptomdebut var associerat med mycket hög risk att senare insjukna i RA.

I delarbete IV analyserades immunceller i blod från patienter med RA, deras förstagradsssläktingar (syskon, barn eller förälder) och friska kontroller. En del av cellerna stimulerades, den andra stimulerades inte. Därefter analyserades en särskild celltyp, kallad B-regulatorisk cell. Skillnaden mellan före och efter stimulering räknades ut. Friska kontroller visade sig reagera starkast på stimulering, medan cellerna från både patienter och deras förstagradsssläktingar reagerade mindre. Denna lägre responsgrad skulle kunna ge en ökad mottaglighet för utvecklandet av RA.

Slutsatserna från dessa studier är att många år innan symptom uppträder finns mätbara förändringar i blodet hos individer som kommer att utveckla RA. Både högre nivåer av och antal av ACPA, anti-CarP- antikroppar och RF av IgA, IgG och IgM isotyp kan påvisas flera år innan symptomdebut. Förekomst av kombinationer av dessa antikroppar är förenat med hög risk att insjukna i RA senare. B-regulatoriska celler från förstagradsssläktingar till patienter med RA har visats vara funktionellt sämre än hos kontroller, vilket skulle kunna bidra till att de är mer inflammationsbenägna, precis som deras släktingar med RA. Sammantaget bidrar dessa fynd till information om hur sjukdomen utvecklas och även hur stor risk en individ har att insjukna i RA och efterföljande sjukdomsförlopp.
Background

Rheumatoid arthritis
Rheumatoid arthritis (RA) is a chronic progressive autoimmune disease affecting primarily the joints, leading to disabilities, fatigue, pain and also a shorter life expectancy as a result of associated co-morbidities. RA is the most common inflammatory disease of the joints and is characterized by inflamed synovial tissue which, with time, extends and erodes cartilage and bone. In some cases extra-articular symptoms of disease are manifest in organs other than joints, such as pericarditis, pleuritis, rheumatic nodules and pulmonary fibrosis among many others. (Turesson et al., 2003). If untreated, 20 to 30 percent of persons afflicted with RA will be permanently work-disabled within two to three years of diagnosis (Rindfleisch & Muller, 2005). Furthermore, cardiovascular disease (i.e., stroke and myocardial infarction) is not only more common in patients with RA compared with the general population, but RA patients affected by, e.g., myocardial infarction have a worse prognosis compared with the general population (Södergren et al., 2007).

Epidemiology
In developed countries approximately 0.5-1.0% the adult population are affected by RA (Alamanos et al., 2006; Neovius et al., 2011). The overall incidence of RA in Sweden is 41 per 100000 and increases with age in both sexes with lower incidence rates being observed in population-dense areas and among more highly educated individuals (Eriksson et al., 2013). In Sweden, the mean age at the onset of disease is 53-61 years, with a clear female dominance with about 70 percent of all affected individuals being women (Söderlin et al., 2002; Rantapää-Dahlqvist et al., 2003; Eriksson et al., 2013).

Diagnosis and classification
The diagnosis of RA is reliant on clinical features, laboratory tests and/or radiological findings. To ensure that RA is classified similarly between studies and populations, and to differentiate RA to other arthritides, a set of criteria was set by the American Rheumatism Association (ARA) to define RA in 1987 (Table 1) (Arnett et al., 1988). The 1987 criteria for RA were well suited for established disease but had a poor performance in identifying RA cases in the early stages of disease (Harrison et al., 1998). In 2010, the American College of Rheumatology (ACR) together with the European League of Arthritis and Rheumatism (EULAR) published an updated, score-based system with four categories (Table 2) (Aletaha et al., 2010). The 2010 criteria includes RF, as previously, but also ACPA, erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP), all of which indicate an ongoing inflammation. A total score of 10 is possible and a score ≥6 is defined as being definitive of RA (Aletaha et al., 2010).
How well the 2010 ACR/EULAR criteria perform is yet to be determined, but a few studies have been conducted and it would appear that the newer criteria have, by and large, the same sensitivity but less effective in identifying erosive disease compared with the older criteria, possibly owing to the effects of intense and effective treatment (Berglin & Rantapää-Dahlqvist, 2013; Mäkinen et al., 2013).

**Table 1.** 1987 revised ARA criteria for the classification of RA

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Morning stiffness</td>
<td>Stiffness in and around joints lasting at least 1 hour before maximal improvement.</td>
</tr>
<tr>
<td>2. Arthritis of 3 or more joints</td>
<td>At least 3 joint areas simultaneously have had soft tissue swelling of fluid (not bone overgrowth alone) observed by a physician. The 14 possible areas are right or left proximal interphalangeal (PIP), metacarpophalangeal (MCP), wrist, elbow, knee, ankle, and metatarsophalangeal (MTP) joints.</td>
</tr>
<tr>
<td>3. Arthritis of hand joints</td>
<td>At least one area swollen in wrist, MCP or PIP joints</td>
</tr>
<tr>
<td>4. Symmetric arthritis</td>
<td>Simultaneous involvement of the same joint areas (as defined in point 2) on both sides of the body (bilateral involvement of PIPs, MCPs, or MTPs is acceptable without absolute symmetry).</td>
</tr>
<tr>
<td>5. Rheumatoid nodules</td>
<td>Subcutaneous nodules, over bone prominences, or extensor surfaces, or in juxtaarticular regions, observed by a physician.</td>
</tr>
<tr>
<td>6. Serum rheumatoid factor</td>
<td>Demonstration of abnormal amounts of serum rheumatoid factor by any method for which the result has been positive in &lt;5% of normal control subjects.</td>
</tr>
<tr>
<td>7. Radiographic changes</td>
<td>Radiographic changes typical for RA on posteroanterior hand and wrist radiographs, which must include erosions or unequivocal bony decalcification localized in or most marked adjacent to the involved joints (osteoarthritis changes alone do not qualify).</td>
</tr>
</tbody>
</table>

A patient is diagnosed with RA if ≥4 of 7 criteria are fulfilled and criteria 1-4 must have been present ≥6 weeks.
Pharmacological treatment

Therapies for RA are aimed at reducing the inflammation and, thereby, also reducing joint damage and loss of function and, if possible, inducing clinical remission. Current best practice is early administration of a disease modifying anti-rheumatic drug (DMARD), methotrexate or sulfasalazine, combining tight controls with prompt adjustments of dosage and drug if the effects of treatment are not satisfying (Suarez-Almazor et al., 2000; Visser et al., 2009). If the effects of treatment remain unsatisfactory the prescription of a TNF-inhibitor, a so-called biological drug, is recommended (Moreland et al., 1999; Lipsky et al., 2000; Weinblatt et al., 2003). Both national (Baecklund et al., 2014) and European guidelines (Smolen et al., 2014) are available for the treatment of RA.

Table 2. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for RA

<table>
<thead>
<tr>
<th>Category A – Joint Involvement</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 large joint* (*refers to shoulder, elbow, hip, knee and ankle)</td>
<td>0</td>
</tr>
<tr>
<td>2–10 large joints</td>
<td>1</td>
</tr>
<tr>
<td>1–3 small joints* (with or without involvement of large joints) (*refers to MCP and PIP, second to fifth MTP, thumb interphalangeal joint and wrist)</td>
<td>2</td>
</tr>
<tr>
<td>4–10 small joints (with or without involvement of large joints)</td>
<td>3</td>
</tr>
<tr>
<td>&gt;10 joints (at least 1 small joint)</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category B – Serology (at least 1 test result is needed for classification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative RF and negative ACPA</td>
</tr>
<tr>
<td>Low-positive RF or low-positive ACPA</td>
</tr>
<tr>
<td>High-positive RF or high-positive ACPA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category C – Acute-phase reactants (at least 1 test result is needed for classification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal CRP and normal ESR</td>
</tr>
<tr>
<td>Abnormal CRP or normal ESR</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category D – Duration of symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;6 weeks duration of symptoms</td>
</tr>
<tr>
<td>≥6 weeks duration of symptoms</td>
</tr>
</tbody>
</table>

A score of ≥6 is needed for a definite classification of RA.
**Radiology**
Plain radiography (X-ray) is routinely used to assess joint destruction, to follow disease development and to determine the severity of disease by visual evaluation of cartilage destruction and joint erosions. One of the most used scoring system to interpret the X-ray images of hands and feet is the modified Larsen method based on a global score for both erosions and joint space narrowing of 32-joint areas (Larsen, 1995).

**Aetiology and risk factors for the development of rheumatoid arthritis**
Albeit extensive research has been undertaken, the aetiology of RA is largely unknown. However, many risk factors have been identified: genetic, environmental, hormonal, infectious agents and certain occupations have all been suggested to increase the risk for developing RA. These factors, or exposures, are also thought to interact, sometimes during several years, making it a complex challenge to elucidate the aetiology of RA.

**Genetic influence in rheumatoid arthritis**
It has long been known that there is a genetic, i.e., an inheritable, component involved in the development of RA. This is supported by familial clustering seen by the greater disease occurrence in first degree relatives (FDR) of probands than of healthy controls (Lawrence, 1970), and by an increased relative risk of two or more for FDR (Frisell et al., 2013). Furthermore, in several studies of twins, monozygotic twins (MZ) (twins sharing the exact same genetic background) have a higher probability of inheriting the same disease (in this case RA) compared with dizygotic twins (DZ) (not sharing the exact same genetic background), i.e. larger disease concordance in MZ compared with DZ, estimated to be approximately 60% (Aho et al., 1986; Silman et al., 1993; MacGregor et al., 2000). In a Swedish study conducted in 2013, the heritability was estimated to be 41% for ACPA positive RA, 10% for ACPA without RA and 39% for RA (Hensvold et al., 2015). New, advanced techniques for genotyping allow analyses of more than a million genetic variations in one single experiment. In 2014, 98 genes at 101 risk loci was reported for Asian and Caucasian populations by Okada and co-workers to be associated with RA (Okada et al., 2014). The list of gene associations continues to increase and will increase as study groups are expanded (Kim et al., 2015). The strongest and most replicated genetic risk factors are the human leukocyte antigen (HLA) alleles, identified in a series of studies of leukocytes from patients with RA and healthy controls during the 1970’s (Stastny, 1976, 1978). The human HLAreon is located on chromosome 6, and comprises three clusters, class I, class II coding for MHC class I and II, respectively, both having a central role in antigen presenting (Figure 1). The third cluster in the HLA-region codes for immune related components, e.g., cytokines and complement factors (Figure 1). Five main loci are located in the HLA class II
cluster: HLA-DP, -DQ, -DR, -DM and -DO. The HLA-DR is the locus having the strongest associated with RA disease, and in particular the so called “shared epitope” (SE) alleles (i.e. DRB1*04 and DRB1*01 alleles) (Gregersen et al., 1987). These HLA-SE alleles have similar sequences within the third hypervariable region (aa70-74, QKRAA, QRRAA, RRRAA) (Gregersen et al., 1987; Raychaudhuri et al., 2012). Carriers of HLA-SE allele have, as stated above, a higher risk of developing RA; an association also shown to be related to the higher risk of having sero-positive disease (ACPA or rheumatoid factor [RF] positive) (Padyukov et al., 2004; Huizinga et al., 2005) and might even be the only reason this association is observed (van der Helm-van Mil et al., 2006). This strong relationship between MHC class II alleles and RA indicate that adaptive immunity might be involved in the development of RA.

Figure 1. Simplified map of the HLA class I, class II and class III gene with a more detailed map of the HLA class II and the HLA-DRB1 ‘shared epitope’ alleles associated with RA.

The second strongest gene association with RA disease has been shown for the PTPN22 gene (protein tyrosine phosphatase non-receptor type 22) (Begovich et al., 2004). The PTPN22 gene encodes for the lymphoid tyrosine phosphatase (Lyp) protein, involved in the negative regulation of T-cell activation (Hasegawa et al., 2004). A substitution of two amino acids (Arginine to Tryptophan) at position 1858 (rs2476601) in the PTPN22 gene has been shown to cause a gain-of-function of the Lyp protein, leading to a stronger suppression of the T-cell activation (Vang et al., 2005).

In ACPA-negative RA, many fewer genes have been identified. TNFAIP3, GIN1/C5orf30, STAT4, ANKRD55/IL6ST, BLK and PTPN22 have been shown to be associated with ACPA-negative RA, different from the susceptibility loci in seropositive RA, adding further evidence that there may be two genetically different subsets (Viatte et al., 2012).
**Environmental factors**
The most established environmental risk factor associated with RA, and in particular with sero-positive RA, is cigarette smoking this association remains evident after adjustment for related factors such as body mass index, social class, age and self-perceived health (Vessey et al., 1987; Heliövaara et al., 1993; Stolt et al., 2003). Smoking has also been shown to increase the risk of sero-positive RA in a dose dependent manner in combination with the strongest genetic risk factor (HLA-SE), and also showing that more citrullinated protein was found in bronchoalveolar lavage from smokers compared with non-smokers (Klareskog et al., 2006). The association between smoking and RA has also been shown to be associated with sero-negative RA, but with a higher risk in the sero-positive group (Bang et al., 2010). The gene-environment interaction between HLA-SE and smoking has been shown to be associated with a variety of antibodies against citrullinated peptides, indicating that smoking is associated with non-specific citrullination of several antigens rather than citrullination to only a few specific antigens (Willemze et al., 2011). Interestingly, it has recently been shown that patients with chronic pulmonary disease (COPD), when investigating both current- and ex-smokers, have an increased citrullination in lung tissue while no difference was found between current- and never-smokers without COPD, indicating that it might not be the smoking per se promoting the citrullination but rather the inflammatory response (Lugli et al., 2015). Furthermore, inflammation as a risk factor for ACPA-positive RA, has previously been shown for bronchiectasis (one of the most infectious diseases of the lung) irrespective of smoking status, (Walker, 1967; Perry et al., 2014).

In smokers, but not in non-smoking individuals, the intake of dietary salt has been shown to double the risk of developing RA (Sundström et al., 2014). This finding is interesting since salt intake has been suggested to play a role in the development of several autoimmune disorders, and that a study both in vitro and in vivo, increased salt levels enhanced Th17 cell differentiation and promoted IL-23R expression, promoting tissue inflammation (Wu et al., 2013).

Alcohol consumption, even in low quantities, has been shown to be protective for the development of RA, with some studies showing that this effect is only present in ACPA-positive RA individuals (Scott et al., 2013; Jin et al., 2014). At present, the mechanism by which alcohol can protect against RA development is not known albeit a few potential mechanisms have been proposed. One potential mechanism is through down-regulation of innate inflammatory processes, analysed in collagen induced inflammatory arthritis (CIA) mice, resulting in less leukocyte migration, an up-regulation of testosterone, leading to decreased NF-κB activation (Jonsson et al., 2007). Down regulation of NF-κB has also been shown after an intake of alcohol in man (Mandrekar et al., 2006). Also, this effect of alcohol could also be explained by individuals with a low to moderate alcohol
intake have a healthier lifestyle compared with abstainers; alcohol consumption could, therefore, be a result of confounding factors. This explanation has been proposed for the J-shaped relationship between alcohol and cardiovascular disease, and overall mortality (Reynolds et al., 2003).

**Occupation and occupational exposures**

It has been shown that blue-collar workers, independent of smoking status (Bergström et al., 2011), farmers and possibly hairdressers have an increased risk of developing RA (Reckner Olsson et al., 2001). Whether it is exposure to the particular occupation or a factor associated with the labour is, as yet, unknown (Khuder et al., 2002). Of the compounds that are associated with an occupational exposure, silica is the most established, being present in stone dust, rock drilling debris and stone crushing and associated with an increased risk of developing ACPA-positive RA, especially in current smokers (Khuder et al., 2002; Stolt et al., 2010).

**Infections and infectious agents**

It has long been hypothesized that infections may act as a trigger of autoimmune diseases. In RA, frequent infections have been associated with an increased risk of developing RA. Furthermore, infections during the first year of life have been associated with the development of sero-negative RA (Carlens et al., 2009; Rogers et al., 2012). Several viral species, such as the Epstein-Barr virus, parvovirus B19 and retroviruses, have been hypothesized to be involved in the aetiology of RA, although it has remained difficult to prove (Woolf & Cohen, 1995). Several of these viruses have also been shown to be present in specimens in early arthritis irrespective of the clinical diagnosis, e.g., osteoarthritis, undifferentiated arthritis or rheumatoid arthritis (Stahl et al., 2000). One of the more interesting links between infectious agents and RA is *Porphyromonas (P.) gingivalis*, a bacterium present in periodontitis (Rosenstein et al., 2004). Periodontitis has been found to be more common and more severe in patients with RA in comparison to patients with osteoarthritis, independent of age, ethnicity, sex, smoking habits and diabetes mellitus (Dissick et al., 2010). In the same study it was also reported that patients positive for anti-CCP antibodies were more likely to have moderate to severe periodontitis compared with patients who were anti-CCP negative. Several oral pathogens are involved in periodontitis but *P. gingivalis* has been the organism of interest due to its unique ability among prokaryotes to express the PAD enzyme, i.e., the enzyme responsible for the conversion of arginine to citrulline (McGraw et al., 1999; Rosenstein et al., 2004). Antibodies against alpha enolase peptide 1 (CEP-1) have been shown to be specific for RA and also to be cross reactive with the bacterial enolase (Kinloch et al., 2005).
**Hormonal factors**

Rheumatoid arthritis is 2-4 times more common in women than men but the mechanisms behind this are largely unknown. The use of oral contraceptives has been shown to be protective against future RA in several studies (Vandenbroucke et al., 1982; Jorgensen et al., 1996; Berglin et al., 2010). Additionally women have a lower risk of developing RA during pregnancy but they do have an increased risk during the first year postpartum (Peschken et al., 2012). Other related factors have been shown with conflicting results, e.g., breast-feeding (Berglin et al., 2010; Lahiri et al., 2014), age at menarche, age at menopause and post-menopausal hormone use (Carette et al., 1989; Hernández Avila et al., 1990; Deighton et al., 1993; Karlson et al., 2004; Pikwer et al., 2012).

**The immune system**

To protect the human body from external (e.g., bacteria, viruses and parasites) and internal threats (e.g., malignant cells and auto-reactive cells) the immune system upholds homeostasis via a sophisticated defence system, i.e., the immune system. The immune system is divided into two parts, the innate immune system and the adaptive immune system. In theory these two systems are described as separate, while in reality they function concomitantly and interactively rather than as two separate entities. The innate immune system upholds the first line of defence, able to combat a wide range of pathogens through physical barriers (e.g., skin and mucosal membranes), chemical barriers (e.g., acidic gastric juice), plasma proteins (e.g., C-reactive protein), signalling molecules (e.g., cytokines) and phagocytic cells with receptors recognizing conserved patterns of pathogens (e.g., pathogen associated molecular pattern receptors or Toll-like receptors).

Apart from the innate immune system, the adaptive immune system offers more sophisticated memory functions, i.e., “remembering” previously encountered pathogens and, thereby, enhancing the response to better protect its “host”. Within the adaptive immune system T- and B-lymphocytes (T- and B-cells respectively) have a key role in that they are the effector cells. The immune system has a number of tasks to achieve, particularly maintaining the delicate balance between protecting its host against pathogens while not reacting to self:

- Recognition of foreign and tolerance towards self
- Self-control and regulation of the immune response
- Effector functions – neutralization and elimination
- Immunologic memory (specific for the adaptive immune system)
B-cells originate from the bone marrow, where they develop and mature. Activation of B-cells occurs when an antigen is bound to the B-cell receptor (BCR), a cell surface immunoglobulin (Ig). The naïve B-cell becomes activated and starts to produce antibodies (immunoglobulins) with the same specificity as the BCR, secreted into, e.g., the blood stream. The BCR (and the immunoglobulins) is able to recognize a variety of molecules with different conformations of, i.e., lipids, proteins, carbohydrates and nucleic acids, present on small chemical groups or large macromolecules. There are five different groups (isotypes) of immunoglobulins that are present in man: IgM, IgG, IgA, IgE and IgD (Figure 2). Each isotype has its own properties: IgM is mainly found in plasma, it is a good activator of the complement cascade, expressed as a naïve B-cell antigen receptor and, with its pentameric structure and ten antigen binding sites it has a high avidity and low affinity. Unlike IgM, the other immunoglobulins are smaller, and diffuse into other tissues more easy. IgA can be found as monomers, dimers and trimers and is mainly present in mucosal surfaces while IgE is mainly associated with mast cell activation and allergic reactions. IgG antibodies have a higher affinity than IgM and can be divided into four subclasses IgG1-4, in the order of their abundance in plasma, with IgG1 being found in the highest concentration. Other functions of the B-cells have been described, the B-

![Figure 2. Cartoon of three different immunoglobulin isotypes](image)

regulatory cells (Breg) with their ability to produce the anti-inflammatory interleukin 10 (IL-10) have been shown to suppress inflammatory responses in mouse models of colitis, collagen induced arthritis, and experimental autoimmune encephalitis (Fillatreau et al., 2002; Mizoguchi et al., 2002; Mauri et al., 2003). In patients with systemic lupus erythematosus (SLE), Breg cells, characterized as cluster of differentiation (CD)19+CD24hiCD38hi were shown to produce less IL-10 in comparison to the same cell type from healthy controls.
T-cells, as well as B-cells, originate in the bone marrow and mature in the thymus. T-cells are involved in the cell mediated immune response. The T-cell receptor (TCR) is, unlike the BCR, only able to detect small linear peptides, peptides that must be presented by an antigen presenting cell (APC) via their major histocompatibility complex (MHC) products. For correct binding of the peptide a co-receptor is required to be present. Either CD4, present on T-helper cells or CD8, present on cytotoxic T-cells specific to MHC class II or MHC class I, respectively. Also, to elicit a response a co-stimulatory molecule (CD80 or CD86) on the APC has to bind to the T-cell. CD4+ T-cells can be differentiated into T helper 1 (Th1), T helper 2 (Th2) or Th17, cells differentiated and named after their respective cytokine production profile. Th1 cells stimulate phagocytes to ingest and kill microbes, and they mainly produce IFN-γ. Th2 cells stimulate eosinophils via IL-4 production. Th17 cells secrete IL-17 and IL-22, with the ability to induce and mediate a pro-inflammatory response. Regulatory T-cells (Treg) are a sub-group of CD4+ cells, also expressing the IL-2 receptor α (CD25) and the transcription factor FoxP3 that is crucial for Treg development. The function of Tregs in rheumatoid arthritis in man, has been shown to be impaired (Ehrenstein et al., 2004). Also, in mice models of arthritis depletion of Tregs and adoptive transfer experiments have shown worsening of and relief from symptoms respectively (Morgan et al., 2003; Morgan et al., 2005).

Autoantibodies in rheumatoid arthritis
Understanding of the disease associated antibodies in RA is a continuously developing field, with new autoantibody systems and an increased knowledge about the epitopes recognized by these antibodies.

Rheumatoid factor
Autoantibodies directed against the Fc portion of IgG antibodies were discovered more than 50 years ago and are designated rheumatoid factor (RF). RF was first described by Waaler and Rose and the definitive test can be, and still is, referred to as the Waaler-Rose test – it is included as one of the ACR/EULAR criteria for RA (Waaler, 1940; Rose & Ragan, 1948; Arnett et al., 1988; Aletaha et al., 2010). RF is primarily associated with RA, being present in between 70 and 80% of the patients, but can also be found in patients with other autoimmune diseases, e.g., Sjögren’s syndrome (van Boekel et al., 2002), infectious diseases (e.g., tuberculosis and hepatitis) (Lovy et al., 1996; Elkayam et al., 2006) and occurs in about 5% of the healthy population (van Boekel et al., 2002).

The presence of RF in RA patients has been associated with a more severe disease including extra-articular manifestations, bone erosions and increased mortality, compared with RF negative patients (van der Heijde et al., 1992; Combe et al., 1995; Sihvonen et al., 2005; Berglin et al., 2006). RF can be of several immunoglobulin isotypes (i.e., IgE, IgM, IgA and IgG) (Gioud-Paquet et al., 1987) with RF of the IgM isotype as the most prevalent at the time of disease onset.
(Rantapää-Dahlqvist et al., 2003). RF of IgM isotype have been shown to precede the onset of RA (Aho et al., 1985). Also, several isotypes of RF, particularly IgA, but also IgM and IgG have been shown to precede the onset of symptoms in RA (Rantapää-Dahlqvist et al., 2003).

The presence of RF of the IgM and IgA isotypes has been associated with exposure of tobacco smoke in both patients with RA and non-RA individuals (Jónsson et al., 1998); a positive correlation has also been found between smoking and levels of IgA and IgM rheumatoid factors (Masdottir et al., 2000; Padyukov et al., 2004).

Genetic susceptibility to the development of RF antibodies has been shown for carriage of the PTPN22 R620W allele. Carriers with two risk alleles in PTPN22 R620W have a higher risk of developing RF antibodies compared with carriers with only one risk allele (Lee et al., 2005). HLA-SE, also a risk factor for RA, has also been described as a risk factor for the development of RF antibodies in individuals with RA (Padyukov et al., 2004).

**Anti-citrullinated protein/peptide antibodies and peptidylarginine deiminase**

There are a number of autoantibodies to citrullinated protein/peptide antigens (ACPA) identified as having a high specificity for RA. The first ACPA to be identified was anti-perinuclear factor (APF) - described for the first time more than 30 years ago by indirect immunofluorescence on buccal mucosal cells and was found to have high specificity for RA (Nienhuis & Mandema, 1964). Anti-keratin antibodies (AKA) were found in 1979 to be co-existent using the sheep-agglutination test (later known as RF) but with a better specificity and a lower sensitivity (Young et al., 1979). After some 15 years of investigation, both of these antibodies, APF and AKA, were shown to share specificity for filaggrin, a cytokeratin-filament aggregating protein (Simon et al., 1993; Sebbag et al., 1995). Also, these antibodies were found to be dependent on the post transcriptional modification citrullination, i.e., the conversion of arginine residues to citrulline by peptidylarginine deiminase enzyme (PAD) (Schellekens et al., 1998; Girbal-Neuhauser et al., 1999).

**Anti-fibrinogen antibodies**

When analysing synovial membranes from patients with RA, Masson-Bessièrre and colleagues found citrullinated forms of both the α- and β-chain of fibrin (Masson-Bessière et al., 2001). Later it was shown that antibodies against two different citrullinated fibrinogen α-chain (at position 16 or 252) peptides, was present in synovial fluid from patients with RA (Takizawa et al., 2006). Using high-resolution mass spectrometry on synovial tissue from RA patients Hermansson and co-workers found two peptides from citrullinated fibrinogen, within amino acids 559-575 of the α-chain and 52-77 within the β-chain (Hermansson et al., 2010).
Anti- vimentin antibodies
The clinically available test for anti-MCV (modified citrullinated vimentin) antibodies has its origin in the anti-Sa antibodies, highly specific for RA (Després et al., 1994). Anti-Sa antibodies were later shown to be directed against vimentin and use an altered form of vimentin in which arginine is replaced by glycine (Vossenaar et al., 2004). The anti-MCV-test has been shown to have a sensitivity of 64-84% and a specificity of 79-96%, not as specific as anti-CCP2 test but may contribute to an improved classification in patients negative for anti-CCP2 (Dejaco et al., 2006; Luime et al., 2010). Anti-MCV antibodies have also been associated with a higher disease activity in RA, as measured by ESR, DAS28 and swollen joint count and when compared with anti-CCP2, CCP3 and CCP3.1 (Innala et al., 2008). Presence of anti-MCV antibodies have also been associated with worse progression of RA disease (Mathsson et al., 2008; Syversen et al., 2010).

Anti- cyclic citrullinated peptide antibodies
As mentioned above, antibodies to keratin and perinuclear factor were, in 1998, found to share specificity for the citrullinated form of filaggrin (Schellekens et al., 1998). The same research group later developed what is known as the first generation anti CCP-test (anti-CCP-1) using a cyclic (instead of the linear) form of filaggrin peptide to increase the tests' sensitivity (Schellekens et al., 2000). The specificity of the anti-CCP-1 test was better when compared with RF while the sensitivity was still lower than RF (Schellekens et al., 2000; Nishimura et al., 2007). The second generation of CCP tests (anti-CCP2) was improved by screening cyclic citrullinated peptides with sera from RA patient, the high specificity (>95%) was retained and the sensitivity increased to comparable to RF (~70%) (Nishimura et al., 2007). Later, a third generation anti-CCP (anti-CCP-3) test was developed for, again, improved specificity and sensitivity, detecting IgG isotype antibodies, and later still an anti-CCP3.1 test capable of detecting both IgG and IgA reactivities was developed (Innala et al., 2008; Santiago et al., 2008; dos Anjos et al., 2009).

Unlike many other clinical tests, the identity of the peptides used in the CCP2 and CCP3 tests are kept confidential while the peptide sequence for CCP1 has been published (Schellekens et al., 2000). For this reason researchers should be aware that these kits may not contain autoantigens present in vivo in man.

Anti- collagen antibodies
The idea that collagen and other joint cartilage proteins have a role in the pathogenesis of RA has been around for many years. Also, collagen induced arthritis (CIA) is the most commonly used animal model of RA, induced by immunisation with cartilage-derived type II collagen resulting in a chronic relapsing disease, or a severe but self-limiting disease depending on animal model used.
Antibodies to collagen type II have been shown in both synovial fluid and serum from patients with RA, with a relatively large discrepancy in the frequency between different findings, from a few percent up to 50 percent of RA patients (Morgan et al., 1987; Clague et al., 1994; Cook et al., 1996; Mullazehi et al., 2012). These antibodies are present in patients with early RA and their levels apparently decrease as the disease progresses, and may also predict a more severe disease progress (Cook et al., 1996). Several epitopes of the type II collagen molecule have been described, the most frequent in synovial fluid and CCP-negative serum is the U1 epitope whilst the CitC1 epitope is the most frequent in CCP-positive sera (Lindh et al., 2014). For collagen type II C1 epitope, both the native (i.e., the arginine containing peptide) and the citrullinated peptide have been shown to be autoantigens (Uysal et al., 2009).

**Anti-α-enolase antibodies**

Antibodies against citrullinated α-enolase (CEP-1) have been found to be reactive with 37-62% of sera from patients with RA and 2% of sera from healthy control subjects (Lundberg et al., 2008). To identify candidate peptide(s) specific for joints, CEP-1 was discovered by probing untreated or citrullinated lysates of monocytic HL-60 cells with RA patient sera to identify reactive polypeptides which were later identified using mass spectrometry. Of the RA sera studied 46% reacted with the citrullinated α-enolase whilst only 13% reacted to the non-citrullinated form (Kinloch et al., 2005).

**Peptidylarginine deiminase**

The enzymes responsible for citrullination is, as stated above, the peptidylarginine deiminase (PAD), a family of calcium binding enzymes which converts the strongly basic arginine to the more neutral amino acid citrulline, by a process known as deimination (Figure 3). To date, five iso-enzymes have been identified, all located in one cluster at 1p36.13 (Vossenaar et al., 2003; Chavanas et al., 2004). The PAD iso-enzymes are expressed in many mammalian tissues: PAD1 in the epidermis and uterus, PAD3 in hair follicles, PAD4 in neutrophils and eosinophils, PAD6 in eggs, ovaries and in early embryos and PAD2 in skeletal muscle, brain, spleen and secretory glands (Vossenaar et al., 2003). The key regulator of PAD enzymes are calcium ions which bind and activate the enzyme (Arita et al., 2004). The deimination can lead to a change in the protein structure and is implicated in several physiological processes, such as terminal differentiation of the epidermis and apoptosis (György et al., 2006). A European genome-wide linkage study of RA sibpairs demonstrated a linkage in the 1p36 locus (i.e., the locus where all PADI genes are located) (Cornélis et al., 1998). It was later shown in a Japanese population that a certain haplotype in PADI4 had a significant association with RA whereas the neighbouring PADI1, PADI2 or PADI3 did not (Suzuki et al., 2003). Furthermore, the expression of PADI4 was found to correlate with levels of antibodies to citrullinated filaggrin (Suzuki et al.,
This is, however, not the case for RA patients of European ancestry as was shown in two English and one French study (Barton et al., 2004; Caponi et al., 2005; Burr et al., 2010), while a Korean study confirms the PADI4 gene susceptibility for patients with an Asian descent (Kang et al., 2006). PADI4 has also been shown to be a target for autoantibodies in RA and to correlate with disease severity and with anti-CCP2 sero-positivity (Takizawa et al., 2005; Halvorsen et al., 2008; Harris et al., 2008; Kolfenbach et al., 2010). It has also been demonstrated that PAD2 expression is higher in the joints of RA patients than in control subjects (de Rycke et al., 2005).

A prokaryotic PAD enzyme has been found in the bacterium *P. gingivalis*, with little sequence similarity with the human homologue, but nevertheless is efficient in deiminating both soluble and peptidyl arginine (McGraw et al., 1999). *P. gingivalis* is a key player in periodontitis, a common infection of the alveolar bone around the teeth caused by, amongst others, this bacterium (McGraw et al., 1999).

![Figure 3. The process of citrullination](image)

**Anti-carbamylated peptides/protein antibodies**

Carbamylation resembles citrullination, both are posttranslational modifications of amino acids with only a methylene groups’ difference (Mydel et al., 2010). Carbamylation is the conversion of Lysine to the non-standard amino acid Homocitrulline, a non-enzymatic reaction dependent on cyanate, which in the body is in equilibrium with urea that can be converted from thiocyanate to Cyanate via the Myeloperoxidase (MPO) enzyme (Figure 4) (Mydel et al., 2010). Carbamylation can occur at times of higher urea concentration *e.g.*, kidney failure.
but the most important pathway in inflammation is when MPO is released from neutrophils (Sirpal, 2009).

**Anti-CarP antibodies**

Based on the high specificity for antibodies to citrullinated peptides in RA patients and the two post-translational modifications, antibodies against carbamylated peptides were investigated and shown to be in significantly higher concentrations in sera from RA patients compared with those from controls (Shi et al., 2011). Since then, anti-CarP antibodies has been shown to be present before patients receive a diagnosis of RA (Shi et al., 2013a), to predict the development of RA (Shi et al., 2013d) and that presence of anti-CarP antibodies is not associated with the HLA-DRB1 or PTPN22 genotype (Jiang et al., 2014). Smoking (and possibly also to [chronic] inflammation) has been shown to enhance carbamylation whilst the development of anti-CarP antibodies has been shown not to be related to smoking (Wang et al., 2007; Jiang et al., 2014). Anti-CarP antibodies have been detected in several animal models of arthritis, but whether ACPA is present or not in these models is currently a subject of debate (Stoop et al., 2015).

The potential cross-reactivity between anti-CarP antibodies and anti-CCP antibodies has been evaluated and next to cross-reactive antibodies, antibodies

![Figure 4](image.png)

**Figure 4.** The process of carbamylation
only reacting with either citrullinated or carbamylated proteins were also present in RA (Shi et al., 2011; Shi et al., 2013e)

**Timing of occurrence of antibodies in the pre-symptomatic period**

It is now well established that antibodies to several different autoantibody systems are present before the onset of the symptoms of RA, but at which time and order they are detectable is still uncertain with conflicting results published. Between RF and ACPA there have been publications stating RF positivity before ACPA positivity but also the opposite (Bos et al., 2008; Majka et al., 2008; Gan et al., 2015). Regarding the time for first positivity to anti-CarP-antibodies directed against either Fibrinogen (Fib) or foetal calf serum (FCS), anti-CCP2 antibodies and IgM-RF it has been reported that they were detectable in (median (IQR)) 7(4-10), 5(3-7), 6(3-10) and 2(1-5) years before diagnosis, respectively (Shi et al., 2013a).

**Pathogenesis**

Rheumatoid arthritis is considered a complex disease in which environmental factors in combination with genetic factors and stochastic (chance) factors contribute to its development. As described previously, the contribution of genetic factors is estimated to be 50-60% (at least for ACPA positive disease) with the remaining 40-50% left for environmental and stochastic factors (see also, Genetic influence in rheumatoid arthritis, page 4). The most established environmental risk factor is smoking (see also, Environmental factors, page 6), which is related to the presence of ACPA and RF. One theory, presented by Klareskog and colleagues is that smoke (and possibly other irritants, e.g. silica) activates macrophages in the lung and some cells go into apoptosis, necrosis or both (Löfdahl et al., 2006; Klareskog et al., 2009). This environment could then lead to an increased synthesis and activity of PAD, leading to citrullination (Klareskog et al., 2006; Bongartz et al., 2007; Makrygiannakis et al., 2008). These citrullinated proteins bind to the HLA-DR/MHC class II on APC with higher affinity in transgenic mice carrying the HLA-SE (Hill et al., 2003; Hill et al., 2008). Continuous cigarette smoking may further increase these effects by stimulating APC and promoting T-and B-cell activation with subsequent ACPA production. Since ACPA may be present years before any symptoms occur, the homing to joints could possibly be due to citrullination in joints after non-specific synovial inflammation (Vossenaar et al., 2004; Makrygiannakis et al., 2006). Also, in (sero-positive) early arthritis patients an increase in activated CD8 and CD19Bcells in inguinal lymph nodes compared with healthy controls has been shown, possibly preceding the earliest changes in synovium (van Baarsen et al., 2013). It has also been shown that antibodies against vimentin may cause osteoclast activation, and potentially spread and perpetuate the inflammation in bone marrow and synovium (Harre et al., 2012). Transferring patient sera containing ACPA or antibodies against citrullinated fibrinogen does not seem, at
least in mice models, to initiate arthritis, although in mice disease induced with anti-collagen antibodies can be enhanced (Kuhn et al., 2006; Klareskog et al., 2014). Explanatory models such as this are needed to account for the presence of ACPA without the development of disease. Another possible ‘second hit’ is the ability of ACPA to activate the formation of neutrophil extra-cellular traps (NETs) from neutrophils (Khandpur et al., 2013), shown to contain citrullinated molecules (Li et al., 2010) and LL37 which is involved in the innate anti-microbial defence system. It is, therefore, possible that anti-microbial defence mechanisms inside the joint actually trigger an autoimmune reaction against citrulline via NETs. Another possible link between citrullinated antigens and inflammation has been proposed by Sokolove et al. showing that citrullinated fibrinogen is able to bind to fibroblasts, and synovial fibroblasts via TLRs with an enhanced activation in the presence of antibodies against citrullinated fibrinogen (Sokolove et al., 2011).

**The arthritic joint**

After the onset of symptoms and when clinical arthritis can be observed, the synovial membrane, which normally has few cells, becomes, with time, hypercellular with a lining of synovial fibroblasts and macrophages. This synovium contains a range of immune cells: synovial fibroblasts, macrophages, mast cells, CD4+ and CD8+ T-cells, NK- and NKT-cells, B-cells and plasma cells. With time the inflamed synovium protrudes into the nearby cartilage and leads to the destruction of the joint lining. The joint is subsequently destroyed by osteoclasts, chondrocytes and synovial fibroblasts (Figure 5). Also within the bone marrow signs of inflammation can be detected. As time passes and joint destruction progresses it is likely that new antigens are created from the destroyed joint and may prolong the reactions with further autoimmune reactivity. Within this arthritic joint, a hypoxic environment is created, promoting vascular growth, a characteristic of the rheumatoid joints (McInnes & Schett, 2007).
Seropositive vs. Seronegative rheumatoid arthritis

As mentioned previously RA is a heterogeneous disease. These differences are to a large extent enlightened by the differences between sero-positive and sero-negative disease. The differences are several: sero-positive disease is related to smoking (Padyukov et al., 2004; Klareskog et al., 2006), with HLA-DRB1 alleles (Stastny, 1976; Wellcome Trust Case Control Consortium, 2007) and PTPN22 (Lee et al., 2005). Furthermore, ACPA positive and negative disease have found to share the positions associated with risk in HLA-DRB1 and HLA-B (11 and 9 respectively) but the amino acids at those positions differed (Han et al., 2014). Also, disease activity, the presence of extra-articular manifestations and co-morbidity (mainly due to cardiovascular disease) is higher in sero-positive patients (Wallberg-Jonsson et al., 1997; Roman et al., 2005; Farragher et al., 2008).

Figure 5. Cartoon of a normal joint and a rheumatoid arthritis joint
**Aims**

The overall aim of this thesis was to identify potential predictors and markers of an ongoing immunological process as part of the early events that subsequently lead to the development and progression of rheumatoid arthritis.

The specific aims with this thesis in pre-symptomatic individuals:

- To assay antibodies against ten citrullinated autoantigen-derived peptides, to evaluate the significance of each antibody separately and in combinations, their time course and possible pathogenic significance for disease development.

- To analyse anti-carbamylated protein antibodies and relate them to the presence of the ten previously analysed ACPA, anti-CCP2 antibodies and disease progression.

- To investigate the interplay between three isotypes of RF and ten different ACPA fine specificities, anti-CCP2 and anti-CarP antibodies analysed previously in samples collected from individuals prior to the onset of symptoms and following a diagnosis of RA in comparison to population controls.

- To compare subsets of T-cells and the suppressive abilities of CD19⁺CD24⁺CD38⁺ regulatory B-cells in individuals at risk of developing RA, patients with RA and control subjects upon stimulation *ex vivo.*
Study populations

Samples from two different patient populations were used, one based on the Biobank of Northern Sweden, including matched population controls and individuals sampled before onset of RA and at the time of receiving a diagnosis of RA. The other included families with several members affected with RA including both probands and their unaffected relatives. The use of blood samples from these individuals were approved by the regional ethics committee in Umeå and all patients gave their informed consent.

Patients with early and established RA
The blood samples from the patients had been collected since December 1995 from a cohort of patients with early RA, i.e., having symptoms of RA ≤12 months at the time of inclusion into the study, and who met the 1987 ARA criteria for RA (Arnett et al., 1988). All patients were residents of the four northern-most counties of Sweden (Norrbotten, Västerbotten, Jämtland or Västernorrland). Patients are continuously recruited into the early RA project that currently involves approximately 1500 individuals. Patients are also registered in the nationwide Swedish Rheumatoid Arthritis Registry.

At the time of diagnosis, a number of variables were recorded, e.g., the 28-joint count of tender and swollen joints; a visual analogue scale for pain and the patient’s global assessment, a health assessment questionnaire, and inflammatory markers (C-reactive protein and erythrocyte sedimentation rate) measured. Apart from these variables, blood was sampled from the patients and divided into aliquots and stored at -80°C.

A register has been created with all patients fulfilling the 1987 ARA criteria ever visited the Rheumatology department in Umeå, Skellefteå or Lycksele since 1995 and with date of symptoms registered.

The Medical Biobank of Northern Sweden
The Medical Biobank of Northern Sweden comprises several different sub-cohorts:

A. The Northern Sweden Health and Disease Study (NSHDS) cohort in which three sub-cohorts are identified:
   i. The Västerbotten Intervention Programme (VIP);
   ii. The Northern Sweden part of the World Health Organization (WHO) study for Monitoring the trends and Determinants in Cardiovascular Disease (MONICA);
   iii. The Mammary Screening Programme.
B. The Northern Sweden Maternity cohort belonging to the Serological Biobank.
Västerbotten Intervention Programme (VIP)
The VIP cohort was initiated in Norsjö, Västerbotten in 1985 as a long-term health promotion project and was gradually extended, until 1991, to include the total County of Västerbotten (Weinehall et al., 2001). All inhabitants aged 40, 50 and 60 (and also 30 year-olds before 1996) in the population of Västerbotten are invited to attend their local health care centre to: complete a diet and lifestyle questionnaire, undergo a medical examination (i.e., measurement of blood pressure, an oral glucose tolerance test, and serum cholesterol and triglycerides) anthropometric measurements and to donate a separate blood sample for later research purposes. A follow up was undertaken after a 10 year interval when a subsequent blood sample was collected and a questionnaire completed.

In October 2013 the VIP included 135000 samples from 95000 individuals ("Enheten för biobanksforskning - NSHDS - Umeå universitet," 2015). The rate of participation has been 57% and differences between participants and non-participants in terms of social characteristics have been described as being marginal (Weinehall et al., 1998).

The Northern Sweden MONICA project
The MONICA project was initiated in 1985 in the two northernmost Counties (Västerbotten and Norrbotten) of Sweden as a part of a multi-national project initiated by the WHO with the main purpose of monitoring trends in cardiovascular disease. Every four or five years, 2000-2500 individuals from these two counties are invited to participate in a health survey involving a health questionnaire, anthropometric measurements, blood pressure and, also, the donation of a separate blood sample for later research purposes, in a similar manner to that in the VIP. These surveys were done in 1986, 1990, 1994, 1999 and 2004, 2009 and 2014. The rate of participation has been high (73-83%) and included, as of October 2013, 14000 samples ("Enheten för biobanksforskning - NSHDS - Umeå universitet," 2015).

The Mammary Screening Programme
Within the county of Västerbotten, women, 50-69 years of age, are invited to undergo mammary radiographs; approximately 7000 women are involved annually. The invitation to participate is sent out every second year. A separate blood sample can be donated for later research purposes. Currently, 29000 individuals are included with 54000 samples being available ("Enheten för biobanksforskning - NSHDS - Umeå universitet," 2015).

Collection of samples within the NSHDS
The patient is sampled after 4 hours of fasting or in the morning following an over-night fast. A twenty millilitre sample is sub-divided into 10 aliquots
consisting of 6 plasma-, 2 leucocyte (buffy coat cells) and 2 erythrocyte samples. All material is frozen at -80ºC within 1 hour following collection.

**The Maternity Cohort**
In 1975 the maternity cohort was initiated when collecting samples from pregnant women being screened for immunity to rubella (i.e., German measles) within the general health care system. The maternity cohort covers the four northern-most counties in Sweden (Norrbotten, Västerbotten, Jämtland and Västernorrland) and includes approximately 126,000 samples from 91,000 women as of August 2013 ("Enheten för biobanksforskning - Maternitetskohorten - Umeå universitet," 2014). The samples collected within the maternity cohort are stored as serum at -20ºC. From the beginning of the cohort until 1987 all sera were heat treated to inactivate any potential contagious agents. After 1988 the sera have not been heat inactivated.

**Identification of pre-symptomatic individuals**
The registries of patients with early and established RA at the Department of Rheumatology at the University Hospital, Umeå, Sweden, was co-analysed with those of the Biobank of Northern Sweden, in order to identify samples pre-dating the onset of symptoms in these patients. All patients fulfilling the 1987 ACR criteria (Arnett *et al.*, 1988) and with a known date of symptom onset were included in the co-analysis.

Altogether, 409 pre-symptomatic individuals were identified as having contributed 742 blood samples ranging from 1 to 7 samples per individual (481 samples from the NSHDS cohorts and 261 from the Maternity cohort).

**Identification of population control subjects**
When the pre-symptomatic individuals were identified, control subjects were randomly selected from the same cohorts as the pre-symptomatic individuals with a three to one ratio, matched for sex and age at the time of sampling. In all, 1308 individuals represented by one sample each were identified with similar age, sex and date of sampling and residence distribution whenever possible.

**Paper I-III**
In Paper I, 409 pre-symptomatic individuals with 742 samples were initially identified from the Biobank of Northern Sweden. Samples from 23 individuals (9 from the Biobank cohorts and 14 from the Maternity cohort) was not available for these studies (25 samples from 18 individuals could not be located; in 3 cases, the storage tubes were empty; and 2 individuals, each of whom donated 2 samples, were excluded as being misdiagnosed). At the end of the selection process, 386 individuals with 717 samples were included for further analysis. Of the 1308
control subjects initially identified three were excluded due to their development of RA identified after the first co-analysis, leaving 1305 population controls for analysis. In Paper I, 204 samples from patients with RA were included from the time point of diagnosis from the “Early RA” cohort. Based on the data derived in Paper I, a subset of individuals was used for further analyses in Papers II and III; the number of individuals included in the separate studies is shown in Table 3.

In Paper II, only pre-symptomatic individuals whose samples had been donated <13.0 years from the onset of symptoms were included, which in turn reduced the number of samples from RA patients, since all samples from RA patients used were required to have at least one sample before symptom onset.

In Paper III, all 717 samples from pre-symptomatic individuals were included for analysis of RF isotypes. Initial analyses of the concentrations of RF isotypes revealed increased amounts in samples from the Maternity cohort donated before 1988, both in cases and control subjects compared with equally old samples from the NSHDS cohorts, probably because of heat inactivation of the samples from that period of time and cohort. Therefore, samples from the maternity cohort collected before 1988 were excluded, reducing the pre-symptomatic material to 598 samples from 321 individuals.

Samples from 495 control subjects were analysed for RF isotypes; three individuals were excluded due to the development of RA after the co-analysis of the registers, leaving 492 control samples. Of the 321 pre-symptomatic individuals, 187 was also sampled at the time of diagnosis.

Multi RA- case family material
From a previous study (Ärlestig et al., 2012) undertaken using samples from the same cohorts, forty-seven multi-case families were identified, including 134 RA patients and 216 unaffected relatives. The number of affected individuals varied between 2 to 8 individuals per family.

Paper IV
From the previously identified 47 -multi-case families, the families living closest to the University Hospital Umeå were invited to participate (Ärlestig et al., 2012). Twenty-one families were included in the study with 27 patients diagnosed with RA according to the 1987 ACR classification criteria (Arnett et al., 1988) and 23 first-degree relatives (FDR), i.e., parent, child or sibling, from the 21 families. One patient could not be phlebotomized and was, therefore, excluded from the study. Eleven control individuals were also included for flow cytometric analysis (FACS). Controls were not allowed to have any rheumatic disease themselves or in their families. For the multiplex cytokine analysis, 17 samples (five men, 12
women) from the Medical Biobank of Northern Sweden (NSHDS cohorts only) were used as additional control material for that analysis (Table 4).

Table 3. Demographic data for the subjects included in Paper I-III

<table>
<thead>
<tr>
<th></th>
<th>Paper I</th>
<th>Paper II</th>
<th>Paper III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-symptomatic individuals</td>
<td>386/717</td>
<td>252/423</td>
<td>321/598</td>
</tr>
<tr>
<td>(n individuals)/(n samples)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females %</td>
<td>82</td>
<td>77</td>
<td>78</td>
</tr>
<tr>
<td>Age at sampling, median (IQR), years</td>
<td>49.9 (28.7)†</td>
<td>52 (17.4)†</td>
<td>50.3 (20.0)†</td>
</tr>
<tr>
<td>Predating time before onset of symptoms, median (IQR), years</td>
<td>7.4 (9.3)</td>
<td>5.2 (6.3)</td>
<td>6.2 (7.2)</td>
</tr>
<tr>
<td>RA patients (n)</td>
<td>204</td>
<td>192</td>
<td>187</td>
</tr>
<tr>
<td>Females %</td>
<td>77</td>
<td>75</td>
<td>74</td>
</tr>
<tr>
<td>Age at disease onset, median (IQR), years</td>
<td>56.5 (15)</td>
<td>57.5 (14.3)</td>
<td>57.1 (15.3)</td>
</tr>
<tr>
<td>Symptom duration prior to diagnosis, median (IQR), months</td>
<td>7.2 (5.9)</td>
<td>7 (5.8)</td>
<td>7.4 (5.9)</td>
</tr>
<tr>
<td>Controls (n)</td>
<td>1305</td>
<td>197</td>
<td>492</td>
</tr>
<tr>
<td>Females %</td>
<td>72</td>
<td>84</td>
<td>72</td>
</tr>
<tr>
<td>Age at sampling, median (IQR), months</td>
<td>51.0 (19.9)</td>
<td>50.1 (20.2)</td>
<td>51.1 (20.1)</td>
</tr>
</tbody>
</table>

†Calculated median age for all samplings

Table 4, Demographic data for the subjects included in Paper IV

<table>
<thead>
<tr>
<th></th>
<th>Paper IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA patients (n)</td>
<td>26</td>
</tr>
<tr>
<td>Females %</td>
<td>77</td>
</tr>
<tr>
<td>Age at the time of sampling, median (IQR), years</td>
<td>64 (22)</td>
</tr>
<tr>
<td>Symptom duration prior to diagnosis, median (IQR), months</td>
<td>14 (18)</td>
</tr>
<tr>
<td>First degree relatives (n)</td>
<td>22</td>
</tr>
<tr>
<td>Female %</td>
<td>41</td>
</tr>
<tr>
<td>Age at sampling, median (IQR), months</td>
<td>53 (23)</td>
</tr>
<tr>
<td>Healthy Controls (n)</td>
<td>11</td>
</tr>
<tr>
<td>Female %</td>
<td>73</td>
</tr>
<tr>
<td>Age at sampling, median (IQR), months</td>
<td>59 (15)</td>
</tr>
<tr>
<td>Population Controls (n)</td>
<td>17</td>
</tr>
<tr>
<td>Female %</td>
<td>71</td>
</tr>
<tr>
<td>Age at sampling, median (IQR), months</td>
<td>52 (10)</td>
</tr>
</tbody>
</table>
Methods

Autoantibody analyses

**Multiplex analysis of ten antibodies against citrullinated peptides**
The presence of ten different ACPA of the IgG isotype was detected using a custom made microarray based on the ImmunoCAP ISAC system (Phadia, Uppsala, Sweden). A full description of the method, including its validation and the diagnostic performance of this technique in comparison with ELISA-based technology, has been published (Hansson et al., 2012). The ten autoantibody antigens investigated are shown in Table 5.

**Table 5. Amino acid sequences of the ten analysed citrullinated antigens**

<table>
<thead>
<tr>
<th>Citrullinated antigen</th>
<th>Citrullinated position</th>
<th>Amino acid sequence</th>
<th>Ref.</th>
<th>Specificity (%)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibrinogen α563-583</td>
<td>573</td>
<td>HHPGIAEFPS(cit)GKSSSYSKQF</td>
<td>[1]</td>
<td>96</td>
</tr>
<tr>
<td>Fibrinogen α580-600</td>
<td>591</td>
<td>SKQFTSSTSYN(cit)GDSTFESKS</td>
<td>[1]</td>
<td>95</td>
</tr>
<tr>
<td>Fibrinogen β62-81a</td>
<td>72</td>
<td>APPPISGGGY(cit)ARPAAAAT</td>
<td>[1]</td>
<td>97</td>
</tr>
<tr>
<td>Fibrinogen β62-81b</td>
<td>74</td>
<td>APPPISGGGYYRA(cit)PAAAAT</td>
<td>[1]</td>
<td>98</td>
</tr>
<tr>
<td>Fibrinogen β36-52</td>
<td>44</td>
<td>NEEGFFS(cit)HGRPLD KK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha-enolase</td>
<td>10 and 16</td>
<td>CKIHA(cit)EIIFDS(cit)GNPTVEC (cyclic)</td>
<td>[3,4]</td>
<td>95</td>
</tr>
<tr>
<td>Triple-helical collagen type II C1 (359-369) (citC1III)</td>
<td>360 and 365</td>
<td>(GPO)5-GA(cit)GLTG(cit) PGDA(GPO)2-GKKYG</td>
<td>[5,6]</td>
<td>96</td>
</tr>
<tr>
<td>Filaggrin (CCP-1/Fil307-324)</td>
<td>13</td>
<td>SHQEST(cit)GRSRRGRSGRSGS (cyclic)</td>
<td>[7,8]</td>
<td>97</td>
</tr>
<tr>
<td>Vimentin 2-17</td>
<td>4, 12 and 13</td>
<td>ST(cit)SVSSSY(cit)(cit)MFGG</td>
<td>[9]</td>
<td>97</td>
</tr>
<tr>
<td>Vimentin 60-75</td>
<td>64, 69 and 71</td>
<td>VYAT(cit)SSAV(cit)I(cit)SSVP</td>
<td>[2,10]</td>
<td>95</td>
</tr>
</tbody>
</table>

*Values of specificity obtained in Paper I, calculated using ROC curves.


The analyses were performed using ten citrullinated peptides and their corresponding arginine-containing peptides using the difference in fluorescence intensity for all peptides except one in subsequent calculations. The exception was citC1III as it has been shown that the ArgC1 peptide is an autoantigen in its own right with conformational epitopes differing from those of the citC1III peptide (Uysal et al., 2009). No comparisons have been made regarding the diagnostic performance using circular or linear peptides (Fil307-324 and CEP-1 are circular peptides, all of the others are linear). Cut-off limits were set using receiver operating characteristic (ROC)-curves giving the optimal sensitivity and a
specificity ≥95%. The specificity for each antibody achieved is presented in Table 5.

**Anti-CCP2 antibodies**
The levels of anti-CCP2 were measured using an ELISA (Euro-Diagnostica, Malmö, Sweden), with the cut-off for positivity set at >25 AU/ml, according to the manufacturers’ instructions. This corresponded, in Paper I, to a specificity of 98% and a sensitivity of 74.5% for RA patients.

**Anti-CarP antibodies**
Antibodies to carbamylated peptides were measured using an ELISA as previously described (Shi et al., 2011). Maxisorp plates (Nunc, Roskilde, Denmark) were coated overnight with either carbamylated foetal calf serum (Ca-FCS) or non-modified FCS. After washing and blocking, samples were added and allowed to bind overnight. Bound antibodies were determined using horseradish peroxidase (HRP)-conjugated rabbit-anti-human IgG (DAKO) in combination with ABTS (2,2’-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid). The cut-off level for anti-CarP antibodies was defined as the antibody reactivity expressed as the sum of optimal sensitivity and specificity, using a ROC-curve, based on the concentrations in samples from the patients with RA and control subjects. The cut-off for positivity was set at 256.07 arbitrary units/mL (AU/mL) giving a specificity of 97%.

**Rheumatoid factor**
Rheumatoid factor of IgA, IgG and IgM isotypes were analyzed using the commercially available EliA assays (Phadia GmbH, Freiburg, Germany) analyzed in an ImmunoCAP 2500-system (Phadia, Uppsala, Sweden) according to the manufacturers’ instructions. The EliA assay is designed as a sandwich immunoassay with the antigen of interest bound to the solid phase, to which secondary enzyme-labelled isotype specific antibodies are added. Unbound antibodies are washed away and a developing agent added, and after stopping the enzymatic reaction, the fluorescence is measured. The fluorescence level is proportional to levels of, in this assay, RF antibodies. Clinical grade assays at the time of analysis were available for IgM-RF and IgA-RF, and the assay for IgG-RF were of research grade. For this reason IgA-RF and IgM-RF are expressed in AU/mL whilst IgG-RF is measured in mg/mL. Cut-off values for all three RF-isotypes were set at 96% specificity using ROC-curves.

**Radiology**
Radiographs of the hands and feet of patients with RA were available at baseline and after 24 months; these radiographs were graded according to the Larsen score and read in chronological order by two clinicians, who were blind to the antibody data as previously described (Berglin et al., 2006). Radiological
progression was defined as the increase of the Larsen score between baseline and 24 months, with the smallest detectable change less than four (calculated according to (Bruynesteyn et al., 2005)).

**Detection of cell subsets and populations**

*Flow cytometry*

In Paper IV, flow cytometry was used to distinguish between several different immune system cells in human plasma, accomplished by using fluorochrome labelled monoclonal antibodies against certain proteins specific for the cell types of interest. Flow cytometry was performed on the LSRII instrument (Becton-Dickinson, San Jose, CA) that uses three lasers beamed through a single cell stream able to detect forward- and side scatter (as a measure of cell size and the inner complexity of the cell, respectively) in addition to the emissions of the fluorochromes on the monoclonal antibodies. The cell types analysed were three subsets of B-cells (CD19+CD20+), namely Breg (CD24+CD38+), CD24intCD38int and CD24hiCD38- B-cells measured for STAT-3 both with anti-CD40 antibody stimulation and without. T-cells were also analysed, both CD4+ and CD8+ populations for the level of the activity markers CD25 and CD69 and also T-regulatory cells (CD4+CD25+FoxP3+).

**Detection of cytokines and chemokines**

Seven cytokines/chemokines were analysed, namely: eotaxin, interleukin (IL)-1β, IL-10, IL-12(p70), IL-6, chemokine (C-C motif) ligand 2 (CCL2) and IL-17A. The concentrations of the selected cytokines were measured in plasma samples as detected by the Milliplex Human Cytokine kit (Millipore Corp, St. Charles, Missouri, USA) following the manufacturers protocol together with the addition of HeteroBlock (Omega Biologicals) to all samples in a concentration of 150 µg/ml to prevent heterophilic antibody interference (Todd et al., 2011). Analyses were made using a Luminex 200 Labmap system (Luminex, Austin, TX, USA). Samples were measured in duplicate using two 96-well plates with first degree relatives, patients and control subjects distributed evenly between the two plates to prevent erroneous interpretations due to plate-to-plate variability.

**Genetic analyses**

Human leucocyte antigen

Genotyping of the HLA-DRB1 shared epitope (SE) alleles was performed using a polymerase chain reaction (PCR) kit with sequence specific primers from a DR low-resolution kit and a DRB01*01 and *04 subtyping kit from Dunal (Oslo, Norway) and Olerup SSP AB (Saltsjöbaden, Sweden). SE-alleles were defined as DRB1*0101*, 0401, *0404, *0405 and *0408. PTPN22 1858C/T genotyping
The PTPN22 1858 C/T polymorphism (rs2476601) was determined by using two methods: firstly, the 5´nuclease assay using ABI PRISM 7900HT Sequence Detector System with the data processed using the SDS 2.1 software (Applied Biosystems, Foster City, CA, USA) as previously described (Johansson et al., 2006); and secondly, genotyping of the PTPN22 1858 C/T (rs2476601) performed using the Immunochip microchip of patients with RA (Eyre et al., 2012).

Statistics
To test categorical data between groups the either Chi-square test was used or the Fischer's exact test when appropriate. Continuous data were compared (pre-patients versus controls or versus patients) by non-parametric analysis using the Mann-Whitney U test for 2 groups, the Kruskal-Wallis test for comparing several groups, and the Friedman’s test for matched pairs (i.e., several samples from the same individual). Correlation analyses were performed using Spearman’s rank correlation. Stratified data were compared with simple logistic regression with adjustments when appropriate. Sensitivities and specificities with 95% confidence intervals (CI) were calculated using the XLSTAT software with the “XLSTAT-Life” add-on (Addinsoft). All P values were presented as two-sided, and P values ≤0.05 considered significant.

In Papers I, II and III calculations for all ACPAs were performed using the delta values (i.e., the value for the respective reactivity against the arginine peptide was subtracted from the value for reactivity against the citrullinated peptide). For each antibody ROC curves were used to determine the optimal cut-off level for positivity using patients and controls, from the largest material available, i.e., cut-offs’ for ACPAs was determined in Paper I, anti-CarP antibodies in Paper II and cut-offs’ for the RF isotypes in Paper III. The time point of the first appearance between individual ACPAs was compared using a Mann-Whitney U test for 2 groups.

In Paper III, hierarchical cluster analysis was used for the investigation of different patterns and groupings of the different antibodies between patients, controls and pre-symptomatic individuals (R Development Core Team, 2014). Conditional inference tree analysis was used to identify the most discriminative sequence order of antibody analysis using “ever positivity” for all analysed antibodies in pre-symptomatic individuals and control subjects (R Development Core Team, 2014).

In Paper IV, data from the flow cytometry analysis were analysed blind by using BD FACSDiva (version 6.1.3) by making a standard set gating for each staining procedure, which was then applied to all experiments, with only minor adjustments allowed within each experiment. Differences in the levels of STAT-3, with and without stimulation, were compared on an individual level using Wilcoxon signed rank test.
Statistical tests were performed using the SPSS for Windows software versions 20-22 (SPSS, Chicago, IL, USA), XLSTAT software (versions 2011-2015) in Microsoft Excel 2010-2013 (Addinsoft) and software R (R Development Core Team, 2014).
Results and discussion

Paper I-III

In order to investigate the presence of autoantibodies in individuals before they presented with any symptoms of a subsequent disease a custom-made microarray, based on the ImmunoCAP ISAC system, of ten different ACPA of IgG isotype was analysed. These antibodies, and anti-CCP2 antibodies, were measured in 386 pre-symptomatic individuals contributing with a total of 717 samples, of whom 204 of these individuals also donated samples at the time of diagnosis, and in 1305 population control subjects matched to the pre-patients as presented in Paper I (see also Table 3, page 24).

Using all available samples, the levels of five out of the ten analysed ACPAs (Fibβ36–52, Fibβ62-81b(74), CEP-1, citC1III, and Fil307-324) were already significantly increased in pre-symptomatic individuals compared with controls (P<0.001); all antibodies differed significantly from controls at the time of diagnose (P<0.011).

The individual specificity values of the ten ACPA: was 95% for Fibα580–600(591), Fibβ36-52, CEP-1 and Vim60-75; 96% for Fibα563-583(573), CitC1III; 97% for Fibβ62-81a(72), Fil307-324, Vim2-17; and 98% for Fibβ62-81b(74). The frequency of antibodies above cut-off levels in samples from pre-symptomatic individuals was highest for antibodies against Fibβ36–52, CEP-1, and Fil307-324 (30.2%, 29.4%, and 26.2%, respectively) using “ever-positive” values, i.e., one value for each pre-symptomatic individual to reduce potential bias from individuals with several samples.

When, instead of stratifying all pre-symptomatic samples into quartiles, the quartile closest to the onset of symptoms (<3.35 years), was associated with the most prevalent antibodies showing a high OR (% frequency) for developing RA were 9.7 (33.9%), 12.5(39.7%), and 14.7 (31%) for Fibβ36–52, CEP-1, and Fil307-324, respectively. Frequencies for all analysed ACPA fine specificities in this time period are presented in Figure 6. For the same time period, combining the presence of anti-CCP2, Fibα72 and/or CEP-1 could be increased to 59.2% sensitivity and a specificity of 91.2%.
Figure 6. Presence of ten ACPA fine specificities in samples from controls, pre-symptomatic individuals with <3.35 years until onset of symptoms and patients with RA. *p<0.05, **p<0.01, ***p<0.001, n.s= not significant

Analysing the presence of ACPA specificities over the quartiles of the whole pre-symptomatic time-span, frequencies increased closer to the onset of symptoms for antibodies against Fil307-324, Fibβ62-81b(74), Fibβ36–52, CEP-1, citC1III, Fibα563-583(573), and Vim2–17 (P<0.05–0.001; chi-square test for trend).

The development of antibody responses before the onset of symptoms of RA was evaluated using the number of ACPA detected positive. One or more ACPA fine specificities was found positive in 50.1% of all pre-symptomatic samples and in 81.9% of all patient samples and 23.2% of all control samples. The number of positive ACPAs increased significantly over time, both when only using samples from the pre-symptomatic period but also when including the samples collected at the time of diagnosis (P<0.001 for both). The increase in the number positive ACPAs was also statistically significant when data where calculated on an individual level (P<0.001). The ability to discriminate between controls and patients using the number of ACPA showed an area under the ROC curve (ROC AUC) and standard error (SE) of 0.877 (SE 0.017) in controls and patients with...
RA and 0.663 (SE 0.013) for controls and samples from pre-symptomatic individuals. In comparison, anti-CCP2 antibodies displayed a slightly higher AUC in patients with RA, i.e., 0.885 (SE 0.017) whilst a slightly lower AUC 0.651 (SE 0.013) was found in pre-symptomatic individuals. The cumulative percentage of positivity for the number of ACPA fine specificities over time is presented in Figure 7.

![Figure 7](image)

**Figure 7.** Cumulative percentage of positivity for different numbers of antibodies against citrullinated peptides present before the onset of symptoms and at diagnosis of rheumatoid arthritis (RA).

In order to evaluate how the levels of the ten ACPAs changed during the pre-symptomatic period, data for individuals with multiple samples were analysed with Friedman’s test and showed a significant increase for antibodies against CEP-1, Fibβ36-52, Fibβ62-81b(74), Fibα563-583(573) and citC1III (P<0.05–0.001). The antibodies not reaching statistical significance (i.e., Fil307-324, Fibβ62-81a(72), Vim2–17, Vim60–75 and Fibα580-600(591)) were also the antibodies with the highest degree of variability in terms of stable positive results in subsequent samples. Cumulative positivity for the ten ACPAs and anti-CCP2 antibodies over time during the whole pre-symptomatic period is shown in Figure 8 and the cumulative percent positive for the ten years closest to onset of symptoms and patients are shown in Figure 9.
Figure 8. Cumulative percentage of positivity for antibodies against citrullinated peptides in blood samples from individuals (n=375) obtained during the time period predating the onset of symptoms of rheumatoid arthritis (RA) and at the time of diagnosis (Patients).

We also evaluated the time point of the first detected positivity for each antibody (Figure 10), with Fibα580-600(591) appearing significantly or close to significantly earlier compared with antibodies against CEP-1 (P<0.05), CCP-2 (P<0.05), Fibα563-583(573) (P<0.05), and Vim2–17 (P=0.076). Also anti-Vim60-75 appeared significantly earlier compared with antibodies against CEP-1 (P<0.05), Fibα563-583(573) (P<0.05), and CCP-2 (P<0.05).

Figure 9. Cumulative percentage of positivity for antibodies against citrullinated peptides in blood samples from individuals (n=375) ≤10 years predating the onset of symptoms of rheumatoid arthritis (RA) and at the time of diagnosis (Patients).
Many of the ACPAs are co-existent in the same samples during the pre-symptomatic period, e.g., covariance between antibodies against CEP-1 and Fib36-52 or CEP-1 and Fil307-324 was shown to be between 50 and 59%. Despite that, combinations of the presence of several ACPAs may provide more powerful estimations of the risk associated with disease development compared with an antibody alone. The OR for the development of RA in individuals expressing the combination of antibodies to CEP-1 and Fibβ36–52 was 40.4 (95% CI 19.8–82.3) for the time period <3.35 years to onset of symptoms, and was fully comparable with the OR of anti-CCP2 antibodies 45.5 (95% CI 28.1–73.8). Adding Fil307-324 to the combination for samples of the same time period, triple positivity yielded an OR of 69.7 (95% CI 25.8–188.4) compared with being negative for all three antibodies.

The presence of antibodies against carbamylated proteins (anti-CarP) was evaluated in samples, collected before the onset of symptoms of RA and previously analysed for the ten different ACPAs (presented in Paper II). In a subset of the samples analysed for Paper I, anti-CarP antibodies were analysed in 252 pre-symptomatic individuals contributing 423 samples when 192 of these individuals were also sampled at the time of diagnosis, and in addition 197 population controls matched to the pre-patients (see also Table 3, page 24). The levels of anti-CarP antibodies were significantly higher in pre-symptomatic individuals compared with the population controls (P<0.0001), with further increased levels in patients with RA (P<0.0001) (Figure 11).

**Figure 10.** Mean (SD) time for the first appearance antibody before the onset of symptoms of the first appearance autoantibody. P-values for comparison between the antibodies are presented to the right.

The presence of antibodies against carbamylated proteins (anti-CarP) was evaluated in samples, collected before the onset of symptoms of RA and previously analysed for the ten different ACPAs (presented in Paper II). In a subset of the samples analysed for Paper I, anti-CarP antibodies were analysed in 252 pre-symptomatic individuals contributing 423 samples when 192 of these individuals were also sampled at the time of diagnosis, and in addition 197 population controls matched to the pre-patients (see also Table 3, page 24). The levels of anti-CarP antibodies were significantly higher in pre-symptomatic individuals compared with the population controls (P<0.0001), with further increased levels in patients with RA (P<0.0001) (Figure 11).
The frequency of positivity in all samples from pre-symptomatic individuals was 13.9% (95% CI 11.0% to 17.6%), in RA patients 42.2% (95% CI 35.4% to 49.3%) and in controls 3% (95%CI 1.3% to 6.7%), significantly higher in both pre-symptomatic and patients compared with controls (P<0.001). Ever being positive for anti-CarP antibodies was associated with the development of RA in pre-symptomatic individuals (OR = 7.1 95%CI 3.0 to 17.0). Adjustments for age, sex or ever being a smoker, did not affect the association between anti-CarP antibodies and disease development.

The frequency of samples positive for anti-CarP antibodies during the pre-symptomatic period increased closer to onset of symptoms, although not significantly (Figure 12). The association between anti-CarP antibodies and disease development was independent of all ten ACPA specificities separately, but not independent when adjusted for anti-CCP2 antibodies (OR = 2.4, 95% CI 0.9 to 6.7). However, selecting pre-symptomatic individuals negative for anti-CCP2 antibodies, anti-CarP antibodies was associated significantly with disease development (OR=7.1 95% CI 3.0 to 17.0).

**Figure 11.** Concentrations of anti-CarP antibodies in pre-symptomatic individuals closest to symptom onset, controls and patients with RA.
One method of investigating the relationship between the ten ACPA specificities and anti-CarP antibodies was by comparing the presence of anti-CarP antibodies in samples where any ACPA was absent, this was found to be between 8.1 and 13.5%. In samples negative for all ACPAs, 5.8% of the samples from pre-symptomatic individuals remained positive for anti-CarP antibodies, the positivity being 13.3% in patients and 2.8% in controls. Furthermore, the relationship between the different ACPA specificities and anti-CarP antibodies was also evaluated in terms of correlation between the different antibody levels. This showed that the levels of anti-CarP antibodies correlated with the levels of all ACPA specificities ($r_s = 0.18$ to 0.34, $P<0.0001$ to 0.04) with the exceptions of Vim2-17, Fibα580-600(591) and Fibβ62-81a(72) in the pre-symptomatic individuals and Vim2-17 and Fibα580-600(591) in RA patients ($r_s = 0.18$ to 0.34, $P<0.0001$ to 0.04).

No difference was found in the time-point for the first occurrence between anti-CarP antibodies and the different ACPA specificities or anti-CCP2 in pre-symptomatic individuals.

In order to investigate the relationship between the presence of anti-CarP antibodies and the radiological progression of joint disease, the joint status at baseline and at 24 months was assessed, and the progression between those time points calculated using Larsen score. Being positive for anti-CarP antibodies during the pre-symptomatic period showed a significant association with
increased radiological damage at baseline compared with anti-CarP negative individuals (P=0.003) remaining statistical significance after adjusting for anti-CCP2 antibodies. The presence of anti-CarP antibodies at the time of RA diagnosis was significantly associated with radiological damage at baseline (β = 2.15, 95% CI 0.40 to 3.90, P=0.017), 24 months (β = 4.49, 95% CI 1.67 to 7.32, P=0.002) and also to disease progression (β = 2.44, 95% CI 0.53 to 4.35, P=0.013). When stratifying data for presence of anti-CCP2 antibodies, individuals’ positive at baseline did not display any association between anti-CarP and the three radiological outcome measures whereas a strengthened association was shown in anti-CCP2 negative individuals with increased damage at baseline (β = 5.38, 95% CI 0.94 to 39.82, P = 0.019), 24 months (β= 8.57, 95% CI 3.36 to 13.79, P = 0.002) and also the radiological progression (β = 2.3.02, 95% CI 0.17 to 5.87, P = 0.038). When analysing the association of anti-CarP antibodies to the three radiological parameters adjusting for antibodies against CCP2, or the three most prevalent analysed ACPA fine specificities (i.e., CEP-1, Vim60-75 or Fil307-324), the only significant association was found for anti-CarP (data not shown).

The presence of RF antibodies of IgA, IgG and IgM isotype was evaluated in pre-symptomatic individuals using samples previously analysed for the ten different ACPA specificities and for anti-CarP antibodies (presented in Paper III).

The levels of all three RF isotypes were significantly increased in pre-symptomatic individuals compared with controls (P<0.001, for all). The levels were further increased at the time of diagnosis, and were significantly higher compared with pre-symptomatic individuals and controls (Figure 13). Concentrations of IgM-RF was increased significantly during the pre-symptomatic period (P<0.001) and also at the individual level (P<0.001) (data not shown).

**Figure 13.** Concentrations of RF isotypes in control subjects, pre-symptomatic individuals and patients after onset of RA disease. Significances are displayed for comparison between all groups (Kruskal Wallis test) and between controls and pre-symptomatic individuals and between pre-symptomatic individuals and RA patients (Mann-Whitney test).
The frequency of RF positivity of IgA, IgG and IgM isotype in all samples from the pre-symptomatic individuals was 25%, 18% and 26%, respectively, and from patients at the time of diagnosis was 64%, 57% and 79%, respectively. The frequencies and OR (95%CI) for all ten of the analysed ACPA, anti-CCP2, anti-CarP-antibodies are presented in Table 6.

Table 6. Diagnostic accuracy of ten ACPA fine specificities (presented in alphabetic order), anti-CCP2 antibodies, anti-CarP antibodies and RF of IgA, IgG and IgM isotype, analysed separately and in combination with the three RF isotypes, comparing pre-symptomatic individuals with population controls.

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Sensitivity (95%CI)</th>
<th>OR (95%CI)</th>
<th>+ IgA-RF OR (95%CI)</th>
<th>+ IgG-RF OR (95%CI)</th>
<th>+ IgM-RF OR (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCP-2</td>
<td>29.6 (26-33)</td>
<td>20.3 (11-39)</td>
<td>50.7 (12.4-207.5)</td>
<td>30.2 (7-125)</td>
<td>67.6 (17-276)</td>
</tr>
<tr>
<td>CEP-1</td>
<td>23.9 (21-28)</td>
<td>4.8 (3-7)</td>
<td>19.6 (7-54)</td>
<td>13.3 (5-37)</td>
<td>48.1 (12-197)</td>
</tr>
<tr>
<td>CitCII</td>
<td>12.7 (10-16)</td>
<td>3.8 (2-6)</td>
<td>36.9 (5-272)†</td>
<td>25.1 (3-187)</td>
<td>45.7 (6-333)</td>
</tr>
<tr>
<td>Fibβ36-52</td>
<td>7.6 (6-10)</td>
<td>1.7 (0.99-3)</td>
<td>21.7 (3-162)</td>
<td>11.8 (2-92)</td>
<td>13.9 (3-59)</td>
</tr>
<tr>
<td>Fibβ36-600</td>
<td>7.0 (5-9)</td>
<td>1.1 (0.7-2)</td>
<td>6.2 (2-22)</td>
<td>6.3 (2-22)</td>
<td>5.2 (2-18)</td>
</tr>
<tr>
<td>Fibβ62-81a</td>
<td>22.5 (19-26)</td>
<td>4 (3-6)</td>
<td>17.5 (6-49)</td>
<td>15 (5-49)</td>
<td>45.2 (11-186)</td>
</tr>
<tr>
<td>Fibβ62-81b</td>
<td>8.6 (7-11)</td>
<td>2.7 (2-5)</td>
<td>17.5 (2-131)</td>
<td>9.1 (1-73)</td>
<td>18.1 (2-136)</td>
</tr>
<tr>
<td>Fil307-324</td>
<td>13.4 (11-16)</td>
<td>9.2 (4-19)</td>
<td>42.6 (6-312)†</td>
<td>24.8 (3-184)†</td>
<td>50.2 (7-366)†</td>
</tr>
<tr>
<td>Vim2-17</td>
<td>5.6 (4-8)</td>
<td>1.3 (0.7-2)</td>
<td>7.6 (2-34)</td>
<td>10.1 (1-80)</td>
<td>13.6 (2-104)</td>
</tr>
<tr>
<td>Vim60-75</td>
<td>9.5 (7-12)</td>
<td>1.9 (1-3)</td>
<td>8.1 (2-27)</td>
<td>9.1 (2-40)</td>
<td>9.3 (3-31)</td>
</tr>
<tr>
<td>CarP</td>
<td>13.8 (11-17)</td>
<td>5.8 (2-16)</td>
<td>12.9 (2-96)†</td>
<td>10.6 (2-79)</td>
<td>16.2 (2-119)†</td>
</tr>
<tr>
<td>IgA-RF</td>
<td>24.8 (21-28)</td>
<td>9.1 (6-15)</td>
<td>23.9 (7-77)</td>
<td>23.9 (7-77)</td>
<td>21.9 (9-55)</td>
</tr>
<tr>
<td>IgG-RF</td>
<td>17.6 (15-21)</td>
<td>5.9 (4-10)</td>
<td>23.9 (7-77)</td>
<td>-</td>
<td>34.5 (8-142)</td>
</tr>
<tr>
<td>IgM-RF</td>
<td>26 (23-30)</td>
<td>11.1 (7-19)</td>
<td>21.9 (9-55)</td>
<td>34.5 (8-142)</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 6. Diagnostic accuracy of ten ACPA fine specificities (presented in alphabetic order), anti-CCP2 antibodies, anti-CarP antibodies and RF of IgA, IgG and IgM isotype, analysed separately and in combination with the three RF isotypes, comparing pre-symptomatic individuals with population controls.

When evaluating the relationships between RF isotypes and ACPA samples from pre-symptomatic individuals, IgA-RF was present in between 16.7% and 23%, IgG-RF in 11.6% to 17.3% and IgM-RF in 15.6 to 25.0% of samples in which ACPA was negative for the ten different ACPA fine specificities. All RF isotypes correlated positively with an increased number of ACPAs in pre-symptomatic individuals (rs range 0.328-0.338, P<0.001).

Ever being positive for all analysed antibodies was analysed in controls and pre-symptomatic individuals using hierarchical cluster analysis. All RF isotypes and anti-CarP antibodies clustered together and antibodies against CEP-1, Fibβ36-52
and CCP2 formed another cluster, also closely related to antibodies against Vim60-75 and Fibβ62-81b(74). Antibodies against Fil307-324, Collagen CitC1, Fibα563-583 and Fibα580-600 constituted another cluster compared with controls where the clustering was overall hardly visible (data not shown).

To find the sequential order of which autoantibody test(s) that provided highest probability in discriminating between controls and pre-symptomatic individuals, conditional inference tree analysis was used (using ever positivity for samples collected <13 years). This analysis revealed that IgM-RF was the most discriminative starting point (p<0.001) (Figure 14). The second antibody to be detectable if individuals were positive for IgM-RF was anti-CEP-1 (p<0.001). Positivity for both of these antibodies identified cases from the controls with a probability of 95.8%. Negativity for anti-CEP-1 antibodies yielded a probability to identify cases of 69.4%. In IgM-RF negative cases the second antibody to appear was against Fil307-324 (p<0.001) as being positive resulting in identification of cases with a probability of 76% (Figure 14). These were followed anti-CCP2 antibodies and IgA-RF.

![Conditional Inference Tree](image)

**Figure 14.** Illustrative conditional interference tree for developing RA in pre-symptomatic individuals compared with controls. n= number of individuals, Pr= probability, + = positive test result, - = negative test result.

The first appearance of antibodies in the pre-symptomatic individuals was evaluated in samples collected ≥15 years before the onset of symptoms and showed that RF of the IgA and IgG isotype were significantly more frequent (and borderline significant for IgM-RF) compared with the most frequent ACPA of that time period, *i.e.*, Fibβ36-52 (P<0.05, for IgA-RF and IgG-RF and P=0.063 for IgM-RF). Analysing at which time point each individual antibody occurred did not confirm any differences since the primary antibody appearing was either any
of the three RF isotypes or ACPA (evaluated as positivity for anti-CEP-1, -Fibβ36-52 or –Fib307-324 antibodies). The development of positivity during the pre-symptomatic time did not show any significant increase for any of the three RF isotypes (Chi-square trend test non-significant).

The combination of having all three RF isotypes compared with having none of them was associated with a very high OR (83, 95%CI 11.3 to 611.3) for developing RA during the five years closest to symptom onset with only minor variations in the four year-groups closer to the onset of symptoms (OR range 92.1 to 101.8), compared with controls.

There was a strong relationship between smoking and the presence of any of the three RF isotypes in pre-symptomatic individuals (P<0.001), whilst no significant relationship was found among control subjects or patients with RA. The association between smoking and the RF isotypes in pre-symptomatic individuals remained when adjusting the data for the presence of the ten different ACPAs (p<0.05, for all) but not after adjustments for anti-CCP2 antibodies. No relationship was found between smoking and anti-CarP antibodies.

Antibodies against ten different ACPA specificities, anti-CarP antibodies and RF of IgA, IgG and IgM isotypes, as well as anti-CCP2 antibodies, have been analysed in the same samples from the largest collection of pre-symptomatic individuals in the world, together with controls and patients with RA. Cut-off levels were determined for the respective antibody using ROC curves yielding specificities between 95% and 98%. The levels for five of the ten analysed ACPAs, anti-CarP antibodies and all analysed RF isotypes were shown to be significantly increased already in the samples from pre-symptomatic individuals compared with controls. In patients, all antibodies analysed were significantly increased compared with controls (P<0.001). Also, in Paper I patterns between the ten ACPAs could be visualised, three antibodies (CEP-1, Fibβ36-52 and Fil307-324) increased significantly during the pre-dating time and were more stable in terms of positivity in subsequent samples, whilst other antibodies (Fibα580-600(591) and Vim60-75) fluctuated in positivity over time but were instead the earliest detectable antibodies. During the pre-symptomatic period an increase of the numbers of ACPA could be observed, with significantly more ACPA reactivities the closer to disease onset, indicating an epitope spreading; there was also an increase at the individual level.

Being able to perform analyses in samples acquired before the onset of symptoms is rare. Therefore, comparisons with other studies of the autoantibodies analysed in this study is not possible. However, in samples from patients with RA, the levels of antibodies in these studies are largely comparable to those published previously. Comparison of frequencies of antibodies between different studies
can often be difficult due to different methods of analysis, different ways of determining the cut-off limit, as well as the cut-off limit itself. With that in mind, the reported sensitivity in RA patients for the ACPAs presented in Paper I showed similar results using ELISAs: for example, antibodies against Fibβ36–52, the observed frequency of 65% is consistent with the published frequencies of 68% (Lundberg et al., 2013) or 73% (Verpoort et al., 2007; van der Woude et al., 2010) reported by other investigators. For anti–CEP-1, the cyclic peptide, a frequency of 67% was noted, which was reasonably similar to the 61% reported for citrullinated α-enolase peptide, a linear peptide (van de Stadt et al., 2011). Furthermore, the previously reported frequency of antibodies against Fil307-324 was 58% (van de Stadt et al., 2011) compared with 46% in the present study, and anti-citC1III was reported to be 40.4% (Burkhardt et al., 2005) compared to 32% in this analysis. The largest disparity for the ACPA analysis was for antibodies against Vim60–75, which was reported to be 49% (Verpoort et al., 2007; van der Woude et al., 2010) or 59% in patients with RA (van de Stadt et al., 2011), whereas the current frequency was found to be 29%. Regarding frequencies of positivity for anti-CarP we reported 42.2% positive at the time of RA diagnosis whereas 26.6% was reported by a Dutch study, although that study also reported lower sensitivity (41.8%) for anti-CCP2 antibodies compared with 74.6% found here, possibly indicating a somewhat different group of RA patients (Shi et al., 2013a). In another, more recent, study involving both Dutch and Swedish patients with early arthritis anti-CarP positivity among the patients was reported to be 44.9% and 35.6%, respectively, which is in line with the findings of this study (Jiang et al., 2014). Frequencies of RF IgA, IgG and IgM isotypes (64%, 57% and 79%, respectively) at the time of a diagnosis of RA were similar to those presented previously (70%, 46% and 73%, respectively) (Rantapää-Dahlqvist et al., 2003). These antibody analyses are considered to be robust, reliable and appropriate for the present studies.

The levels of five ACPA fine specificities (Fibβ36–52, Fibβ62-81b(74), CEP-1, citC1III and Fil307-324), anti-CarP antibodies and three RF isotypes were significantly elevated in samples from pre-symptomatic individuals compared with controls. Levels of all antibodies analysed were further increased at the time of diagnosis and significantly higher compared with controls. Among the ACPA fine specificities three clusters of antibodies were observed to behave similarly during the pre-symptomatic period. The first cluster consisted of antibodies against CEP-1, Fibβ36–52, and Fil307-324, displaying an expanding number of individuals converting to sero-positivity over time with increasing levels of these antibodies present the closer to the time of disease onset. These three antibodies also showed the strongest association to subsequent disease development. A second group of ACPAs consisted of the earliest detected antibodies, Fibα580-600(591) and Vim60–75. These two antibodies and Fibβ62-81a(72) were also the most fluctuating antibodies in positivity in subsequent samples during the pre-
symptomatic time. The increase in antibody levels was modest in samples collected close to the time of disease onset, although the levels of these antibodies, in particular antibodies against Vim60–75, increased further after the onset of clinical symptoms. It is speculated that these results indicate a lesser pathogenic importance as compared with other antibodies that vary in intensity more dramatically during the pre-symptomatic period. Additionally a third cluster of antibodies including citCl11, Fibα563-583(573), and Fibβ62-81b(74) was identified in which the antibodies occurred at low frequencies before the onset of symptoms but were more consistently positive over time and with a significant increase in concentration after disease onset (P<0.001). It is suggested that these antibodies may not be involved in the initial phase of autoimmunity to citrullinated proteins occurring systemically, but may be related to the clinical onset of joint disease.

A conditional inference tree analysis was performed on all analysed autoantibodies sampled <13 years before symptom onset (10 ACPA fine specificities, anti-CarP antibodies and 3 RF isotypes) including anti-CCP2, in order to find the combination(s) of autoantibody tests providing the highest probability of disease development by discriminating between pre-symptomatic individuals and population controls. This analysis revealed, unexpectedly, that positivity for IgM-RF followed by CEP-1 positivity provided the highest probability of RA disease development and not anti-CCP2 antibodies. The presence of anti-CCP2 antibodies had the highest OR when evaluated as a separate factor or in combination(s) with other antibodies using logistic regression analysis. The discrepancies between the results of inference tree analysis and logistic regression is due to the fact that in a logistic regression analysis the highest OR is given when controlled for other variables present in the model, reflecting a more general effect of the variable on the response variable concerned, and also that the variables in a conditional inference tree analysis are selected consecutively so that the variables can show the highest probability of developing RA.

The first occurrence of autoantibody positivity was evaluated between the three RF isotypes, anti-CarP antibodies and the ten ACPA fine specificities, stratified for sample collection ≥15 years before the onset of symptoms (i.e., making comparisons with anti-CarP impossible because only samples <13 years were analysed). Stratified for that time period, IgA-RF and IgG-RF were present in 26 and 13% of the samples, respectively, both being significantly more frequent in comparison to the most frequent ACPA for that time period, i.e., Fibβ36-52 (P<0.05). The discussion about whether RF or ACPA occurs first, and also at which time, is under current debate, with the results being conflicting; both with ACPA occurring first and also the opposite as well as a non-significant difference
between them (Nielen et al., 2004; Bos et al., 2008; Majka et al., 2008; Gan et al., 2015).

The presence of several autoantibodies before the onset of RA is associated with an increased risk of developing the disease. In Paper I, comparison of population controls with pre-symptomatic individuals with less than 3.35 years before the onset of symptoms, the relative risk of developing RA if triple sero-positive for Fibβ36-52, CEP1 and Fil307-324 compared with being negative for all three was 70 times higher in individuals that will develop RA. Similarly, having all three RF isotypes compared with any other number of RF (none, one or two RF isotypes) yielded a 61 times higher relative risk of developing RA compared with controls. Alternatively, the combination of anti-CCP2 antibody presence with positivity for one or more RF isotypes and anti-CarP antibodies showed a sensitivity of 5.5% but with 100% specificity (presented in Paper III). The information about these types of combinations may provide useful information in the future for predicting disease development in individuals without symptoms but still at risk.

The presence of the ten ACPA, anti-CCP2 antibodies and anti-CarP antibodies was evaluated both in samples from the pre-symptomatic period and at the time of diagnosis, and related to radiological findings at baseline and at a follow-up after 24 months. It was then possible to conclude that the presence of anti-CarP antibodies before the onset of symptoms predicted the radiological findings at baseline. Additionally, confirmation of previously presented results regarding the presence of anti-CarP antibodies to be associated with the radiological progression was achieved (Shi et al., 2011). Furthermore, this association was particularly evident in anti-CCP2 or anti-Fil307-324 sero-negative individuals. The previous presented result by Mathsson and colleagues indicated an association between antibodies against citrullinated vimentin and poor radiographic progression was also found here for the Vim60-75 antibody, although not significant when adjusting for presence of anti-CarP antibodies (Mathsson et al., 2008).

None of the studies performed address the question of cross-reactivity against the different citrullinated and/or carbamylated peptides, but the fact is that different autoantibodies occurred in different patterns, indicating a limited degree of cross-reactivity. There are a few studies that have addressed the issue of cross-reactivity, demonstrated by cross-absorption experiments that provided evidence that there is some cross-reactivity, but also that a substantial proportion of the antibodies against α-enolase, vimentin, fibrinogen, and type II collagen do not cross-react when their respective citrullinated target antigens are used as a target in assays (Ioan-Facsinay et al., 2011; Snir et al., 2011; Lundberg et al., 2013). In the present study there was an evident covariance with antibodies against CEP-1, Fibβ36–52, and Fil307-324, which may essentially be a real covariance rather
than cross-reactivity. For anti-CarP antibodies, cross-reactive antibodies, as well as antibodies reactive to only citrullinated or carbamylated antigens have been shown to be present in patients with RA (Shi et al., 2013e). Recently a study showed that HLA-associations with anti-CarP antibodies to be completely different from what was previously shown to be associated with ACPA fine specificities, and also that smoking was not related to the development of anti-CarP antibodies (Jiang et al., 2014). Collectively, these findings do not favour the idea that anti-CarP antibodies only represent cross-reactivity to citrullinated antigen. It is likely that both cross-reactive and non-cross-reactive antibodies co-exist.

Although being able to analyse samples from the largest patient cohorts available, with samples identified before the ascertained date for the onset of disease symptoms, identification of any relationships between the different antibodies and statistical calculations remained limited. The samples were not, at the individual level, collected at regular intervals, and by stratifying the data the number of values for statistical calculation was reduced. The number of data points was also reduced as the treatment of samples before 1988 affected the analyses of RFs, and the age of the individuals also affected the concentrations of RFs, which was also taken into account.
In order to investigate the response to stimulation of B regulatory cells from patients with RA mononuclear cells were isolated from patients, their first-degree relatives (FDR), i.e., a known group at-risk of developing RA, and healthy control subjects. The isolated mononuclear cells were incubated with or without anti-CD40 stimulation and were, following stimulation, divided into either CD24^{hi}CD38^{hi}, CD24^{int}CD38^{int} or CD24^{hi}CD38^{-} B-cells (CD19^{+}CD20^{+}). First, the percentage of CD24^{hi}CD38^{hi}, CD24^{int}CD38^{int} and CD24^{hi}CD38^{-} B-cells of total number of CD19^{+}CD20^{+} B-cells, both with and without stimulation were analysed, however no significant differences in the percentage of these cells were found when comparing the patient, FDR and control groups. Secondly, the levels of STAT-3 at baseline in the three B-cell subsets were compared in patients, FDR and controls with no significant differences being found.

In order to investigate whether patients with RA, and potentially their FDR, have altered functionality of their CD24^{hi}CD38^{hi} B-cells, STAT-3 expression in the three different B-cell populations was analysed. Comparing patients and their FDR as well as controls at a group level, no significant difference could be found in any of the three B-cell populations in terms of STAT-3 expression at baseline or in the relative cell number whether with or without anti-CD40-stimulation.

Comparing the STAT-3 expression with and without anti-CD40 stimulation for each individual separately, significantly increased levels were found in healthy controls for all of the three cell populations analysed (p<0.05 for all, Wilcoxon signed rank test), whilst no difference was found in patients or FDR.

By calculating the STAT-3 difference (with versus without stimulation) as the relative change for each individual and B-cell subset, only the relative change of STAT-3 in CD24^{hi}CD38^{-} B-cells was found to be significantly different between patients, FDR and controls, with the lowest response in patients (p=0.008, Kruskal-Wallis test). These differences were also found to be significant when comparing FDR versus controls and patients versus controls (p=0.01 and p=0.003, respectively, Mann Whitney U-test), whilst no difference was found between FDR and patients (p=0.32) (Figure 15). In a subsequent analysis, the phosphorylated (Y705) variant of STAT-3 (pSTAT-3) was analysed in a sub-group of patients, FDRs and controls. This analysis showed similar levels of pSTAT-3 both with and without stimulation in all three groups (data not shown).

Comparing the STAT-3 difference in any of the three cell subsets stratified for HLA-SE, PTPN22 1858T, ACPA positivity, smoking status or by which family an individual belonged to, did not show any statistical significance.
The levels of cytokines could potentially contribute to the development of Breg cells, and/or reflect the presence the functionality of these and other related cell subsets. Levels of eotaxin, IL-1β, IL-10, IL-12(p70), IL-6, CCL2 and IL-17A were measured in patients, FDR and controls; all of these analytes had previously been shown to be altered in pre-symptomatic individuals or patients with RA (Kokkonen et al., 2010). The cytokine measurements showed elevated levels of IL-6 in RA patients compared with both controls and FDR (p<0.05 and p<0.01, respectively), and both IL-1β and IL-10 were found to be significantly higher in patients compared with FDR (p<0.05). IL-12 levels were significantly higher in both the patient and FDR groups compared with controls (p<0.05, for both). CCL2 was found in the lowest concentrations in FDR, being significantly lower compared with both control subjects and patients (p<0.01 for both). The level of IL-17 in the patients was significantly higher compared with controls (p<0.01). No statistical difference was found between the groups in terms of eotaxin levels.

Factors known to contribute to disease, i.e., HLA-SE, PTPN22 1858T-carriage, ACPA positivity or ever being a smoker were analysed against the different cytokine levels for the three groups separately (HLA-SE and PTPN22-1858T genotype data were only available for RA patients and FDR) with higher levels of IL-12 being found in ACPA positive RA patients compared with their ACPA

![Graph](image)

**Figure 15.** Median relative change (Q1-Q3) in STAT-3 expression, with versus without CD40-stimulation in three different CD19+CD20+ B-cell subsets; CD24hiCD38hi B cell, CD24intCD38int and CD24hiCD38+ cells in controls, first degree relatives (FDR) and patients, measured by flow cytometry. Comparisons were made for two groups using the Mann-Whitney U-test; 1 Patients vs. controls. 2 FDR vs. controls and 3 Patients vs. FDR. *P<0.05, **P<0.01, n.s= no significant difference
negative counterparts (p<0.05). None of the other risk factors showed any alterations in cytokine levels in patients with RA or their FDR.

A consequence of functionally defective Breg cells could be a reduced conversion of CD4+ cells into Treg cells resulting in less suppressive capacity, possibly promoting a more autoimmune prone state. The proportion of Treg cells, identified as CD4+CD25+FoxP3+ cells, to the total number of CD4+ cells, and their activation status, measured by expression of CD69, were analysed in the three subject groups. The relative number of Treg cells in relation to CD4+ cells and to lymphocytes was similarly distributed between the patients with RA, their FDR, and control subjects (data not shown). Furthermore, no significant relationships could be found between the relative change of STAT-3 in CD24hiCD38hi B-cells and association to the relative number of Treg cells or levels of CD69 expression. Whether the CD4+CD25+FoxP3+ population was influenced by the presence of HLA-SE, PTPN22 1858T-carriage, ACPA or to which family an individual belongs to was also investigated. All of these parameters remained non-significant, except that “ever smokers” had a slightly higher percentage of CD4+CD25+FoxP3+ in relation to CD4+ cells compared with non-smokers (4.6 vs. 3.2%, p<0.05, Mann Whitney U-test).

Analysis of CD4+ and CD8+ cells with and without expression of the activation markers CD25 and CD69 in patients with RA, their FDR and controls, showed a significant difference across the three groups in CD25 expression on CD4+ cells (p<0.05, Kruskal-Wallis test). The lowest CD25 expression was found in the FDR group, resulting in the only significant difference being between FDR and patients in a two subject group analysis (p=0.019, Mann Whitney U-test).

Three different B-cell populations (CD24hiCD38hi (B-regulatory cells), CD24intCD38int and CD24hiCD38- cells), previously described by Blair et al. were analysed (Blair et al., 2010). The only significant response to stimulation was found in the cells originating from healthy control individuals (p<0.05), analysed at an individual level, whilst patients with RA and their FDR did not show any changes, measured as intracellular STAT-3 levels using flow cytometry. Conversely, patients with RA showed a decrease in the level of available total STAT-3 with stimulation. This observation, together with no difference in the proportion of phosphorylated STAT-3 (pSTAT-3), was found when stimulated, implies that the available amount of the biologically active form of pSTAT-3 is lower in patients, and thus this could lead to functional consequences.

These findings are in concordance with a previous report that Breg cells (CD19+CD24hiCD38hi) from patients with SLE respond with an impaired IL-10/STAT-3 response following stimulation (Blair et al., 2010). In a recent report, CD19+CD24hiCD38hi Breg cells from RA patients have been shown to be
functionally defective in terms of converting naive CD4+ cells into functional Treg cells with a suppressive capacity and a failed prevention of the development of Th17 cells, compared with Breg cells (CD19+CD24hiCD38hi) from healthy individuals (Flores-Borja et al., 2013). Next, whether a decreased function of CD24hiCD38hi B cells would show a numerical deficit of Treg cells was investigated, however no such relationship could be confirmed. Additionally, Flores-Borja et al. also found that in RA patients with an active disease fewer CD19+CD24hiCD38hi Breg cells were found compared with both healthy control subjects and patients with inactive disease (Flores-Borja et al., 2013). Numerical deficits of Breg cells have also been described in other inflammatory/autoimmune diseases such as multiple sclerosis (CD19+ IL-10 producing cells), SLE (CD19+CD24hiCD38hi), and, recently, also in ANCA associated vasculitis (CD19+CD24hiCD38hi) (Duddy et al., 2007; Correale et al., 2008; Blair et al., 2010; Todd et al., 2014). However, in this study no difference was found in the number of CD24hiCD38hi Breg cells between RA patients, their FDR and healthy controls.

Of note is the finding that the CD24hiCD38hi Breg cells from RA patients and their FDR were similarly deficient in responding to stimulation, possibly indicating shared factors, e.g., genetic and/or environmental, between FDR and the patients, either alone or in combination, have an influence on CD24hiCD38hi Breg cell function. The familial aspects of this finding remain to be further elucidated within larger multi-case families, optimally using non-genetically related family members as controls, e.g., a spouse.

Based on these results, it is hypothesised that in FDR, CD24hiCD38hi B-cells are functionally impaired as well as in RA patients, but the FDR have not developed an active inflammatory disease as measured by the cytokine levels compared with the RA patients. The FDR express lower CCL2 levels (as a potential sign of a lower activity in macrophages), reduced Th17 cell activity (reflected by less pro-inflammatory IL-17) as well as a lower expression of activation markers on CD4+ cells. Furthermore, the RA patients show features of an activated immune system through elevated levels of IL-1β, IL-6 and IL-12, possibly due to also carrying defective CD24hiCD38hi Breg cells, and as Flores-Borja and colleagues recently presented, not being able to suppress the conversion of CD4+ into Th17 cells; a situation that is indicated in this study by elevated IL-17 levels (Flores-Borja et al., 2013).

It is recognised that this study has limitations, e.g., the number of participants was relatively low with a shift to a greater age in the patient group. However, this study does include more individuals than many of the previous studies analysing Breg cells in autoimmune diseases (Duddy et al., 2007; Correale et al., 2008; Blair et al., 2010). Another limitation is that the stimulation procedure, involved
all mononuclear cells with a subsequent analysis of STAT-3 expression in three selected B-cell populations, which could have been influenced by factors other than those inherently B-cell related. Furthermore, it is recognised that the cytokine levels expressed by the RA patients reported in this study are lower than those in many other studies. However, cytokine measurements can be influenced by many factors, such as the storage of samples, sensitivity of the chosen method and influence of heterophilic antibody interference, e.g., RF-antibodies.

In conclusion, the findings of this study suggest, not only a decreased CD24+CD38+ regulatory B-cell function in patients with RA, but also in unaffected FDR from multi-case families, with both groups differing significantly from healthy controls. As CD19+CD24+CD38+ Breg cells are involved in the conversion of CD4+ T cells into functional Treg cells, the numbers of both CD4+ T-cells and Treg cells in peripheral blood but no difference was found. No associations could be identified when analysing both Breg and Treg cells flow cytometry parameters in relation to known environmental (smoking) and genetic risk factors (HLA-SE, PTPN22 1858T variant) for the development of RA. Also analysed were cytokine levels in patients, FDR and controls, and based on these results, it is proposed that FDR do not progress to disease due to alternate cytokine pattern(s).

**Concluding remarks**

In this thesis several autoantibodies and immune cells have been analysed in individuals sampled before the onset of symptoms of RA or individuals at risk of developing RA. The largest part of the autoantibodies analysed are directed against post-translationally modified peptides and proteins. The majority of these autoantibodies showed increasing levels the closer to onset of symptoms with even higher levels at the time of diagnosis. This pattern varies between individuals and it is not known whether these antibodies are the result of an already active, although silent, underlying immune process or if they play a role in the pathogenesis themselves, successively driving the disease with more and more autoantibody reactivities. A shared decrease in immune cell function has been found between patients with RA and their first degree relatives, i.e., individuals with a higher risk of developing RA, a situation that was not present in controls compared with RA patients.
Conclusions

The main conclusions from this thesis, based on individuals from Northern Sweden who donated samples prior to the onset of RA, patients with RA and their first degree relatives, can be summarized as follows:

- Antibodies against citrullinated peptides can be detected several years before onset of symptoms of rheumatoid arthritis.

- The concentrations and number of autoantibodies increased with time towards the onset of symptoms.

- The most frequent antibodies against citrullinated peptides detected in pre-symptomatic individuals were those against CEP-1, Fibβ36-52 and Fil307-324.

- Autoantibodies against carbamylated peptides can be also be detected several years before onset of symptoms.

- The presence of antibodies against carbamylated peptides before onset of symptoms was associated with a worse progression of disease assessed radiologically after the onset of symptoms.

- Rheumatoid factor of the IgA, IgG and IgM isotype can be detected several years before onset of symptoms.

- Rheumatoid factor of the IgG and IgA isotypes were significantly more frequent in samples collected more than 15 years before the onset of symptoms compared with the most frequent anti-citrullinated protein/peptide antibodies (ACPA).

- Rheumatoid factor of the IgM isotype provided the highest probability value, in combination with CEP-1 antibodies, for disease development.

- A significantly decreased regulatory function of CD24hiCD38hi B-cells in patients with RA and their unaffected first degree relatives compared with healthy control subjects was found.
Future perspectives

Further investigation of the disease process in RA is needed to gain information about how to construct better predictions of disease, which immunological and molecular pathways that are involved in the pathogenesis of disease, how to treat patients, and how to, ultimately, prevent the disease. As RA is a complex disease further investigations are needed in several areas, for example in genetic associations, immunological responses, the role of environmental factors and the interplay between them. For example, more information is needed regarding the genetic influence of the autoantibody repertoire concerning the relationship between certain native and citrullinated peptides and certain HLA-types. In parallel, the environmental role on the development of autoantibodies requires investigation since periodontitis, smoking, and recently, bronchiolitis have been proposed as potential contributing factors in the development of ACPA. This should focus on separating the effects of smoking and those of, e.g., chronic inflammation of gum or lung.

It would be of interest to analyse several factors, other than autoantibodies, in individuals before onset of disease. A previous publication, consistent with those published by others, contained information about an altered gene expression of IFN-related genes during the pre-symptomatic period (Lubbers et al., 2013). This could be incorporated in a hypothesis-generating setting with, e.g., RNA sequencing analysis, investigating all available RNA present including micro RNA. This information could then be paired with information from the Medical Biobank, e.g., salt intake, smoking or the autoantibody profile presented here. Together these could be used as an indicator as to which molecular processes and pathways are active before the onset of disease, possibly providing information about the pathogenesis, and also used in constructing a model of disease risk.

Since the analyses of antibody detection are based on the presence of antibodies in patients with RA and not before the onset of RA it would be interesting to analyse the antibody reactivities present before the onset of disease in order to detect possible reactivities that may not be present at the time of diagnosis. This could be accomplished by analysing antibody reactivity in sera using a protein microarray with thousands of proteins present. Furthermore, regarding antibody reactivities, investigating the presence of antibodies against post-translational modifications other than citrullination and carbamylation, e.g., deamidation, would be of interest.

Further analyses of Breg cells in patients with RA, their unaffected FDR and using other family members as controls should prove interesting. Breg cells could be sorted, counted, isolated, and stimulated separately to further elucidate their
involvement in RA. Furthermore investigation of the relationship of Breg cells and disease activity is needed.
Acknowledgements

A lot of people have contributed to this work and in particular I would like to thank the following people:

My supervisor Solbritt Rantapää Dahlqvist, always there for questions and two steps ahead. I am continuously impressed by your enthusiasm and energy.

My colleagues in the lab Lisbeth Ä, Heidi K, Solveig L, Antonia B and Linda J for enriching my everyday life at the office. Lisbeth Ä and Heidi K thanks for the discussions about most things in life and for all help in the lab. The students: Linda for all laughs, crazy animal pictures and the word-of-the-week, Antonia for supporting crazy my ideas of cooling down our office during the warmest summer of mankind. Thanks Solveig L, for always being a cheerful spirit in the lab and office.

The people at the immunology lab for helping out a confused soul (i.e., me!) in the FACS-room; Kristina L, Radha T, Viqar B and Ann I. With a special thanks to Radha for our fun and inspiring adventures and discussions during travel and also for being the best co-host of the Umeå conference 😊

Thanks to all the people around the past and present fika and lunch-table, Malin O, Eva B, Agneta Ö, Åsa Lu, Clara O, Clas W, Karin B, Carina A, Susanne H, Terry P, Monica H, Maria L, Maria H, Urban H, Mikael L, Åsa La, Nina F P, Emma A E, Elin L, Lelle and Nina G.

To all people around the lab, Elin C, Kotryna Š and the pediatrics people.

Thank you Nina N, for sharing the same ideas about computers and being eager to learn new stuff.

Thanks to all the people in the National Clinical Research School in Chronic Inflammatory Diseases. Jan-Alvar L for being THE organizer and for Helena H and others starting this amazing research school. Thanks also to all the students, we’ve had great fun and I’m looking forward to collaborate with you in the future.

Thanks Ingegerd S, for introducing me to the Rheumatology Lab.

Thanks to Lotta L who together with Solbritt convinced me to start as a PhD-student, before all thoughts were on ´galonisar´ and other kids-stuff.
Thanks to the people over at the clinic of Rheumatology: Solveig W-J, Lena I, Anna S, Gerd-Marie A, Kristina J, Björn S, Stefan E with a special thanks to Sonja O and Inger B who always supports if blood is to be drawn and for seeing the possibilities and not the obstacles.

Thank you to Svensk Reumatologisk Förening and Reumatikerdistriktet Västerbotten who made it possible for me to visit and present my work at international conferences.

Hasse S and Priyantha W for statistical help.

Brian E for your prompt and skilful corrections and comments on my papers and thesis.

All my co-authors, for fruitful discussions and collaboration.

Fredrik N, for being my sparring-partner through elementary- and high school and friend.

My parents, Eva and Christer, for always supporting my decisions in life and for being the best parents.

My sister Sofie with family, thank you for always being there.

Last but not least, my family. Therese for your constant support and patience with me. Carl for being the best son ever and the sunshine of my life. I love you both.
References


rheumatoid arthritis and has a specificity of 100% for the disease. *Arthritis Research & Therapy, 8*(1), R19. doi: 10.1186/ar1868


Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature, 447*(7145), 661-678. doi: 10.1038/nature05911


http://www.nature.com/nature/journal/v496/n7446/abs/nature11984.html#supplementary-information
