ARTHRITIS & RHEUMATISM
Vol. 62, No. 2, February 2010, pp 383–391
DOI 10.1002/art.27186

Up-Regulation of Cytokines and Chemokines Predates the Onset of Rheumatoid Arthritis

Heidi Kokkonen,[1] Ingegerd Söderström,[1] Joacim Rocklöv,[2] Göran Hallmans,[2] Kristina Lejon,[2] and Solbritt Rantapää Dahlqvist[1]

Objective. **To identify whether cytokines, cytokine-related factors, and chemokines are up-regulated prior to the development of rheumatoid arthritis (RA).**

Methods. **A nested case–control study was performed in 86 individuals who had donated blood samples before experiencing any symptoms of disease (pre-patients) and 256 matched control subjects (1:3 ratio). In 69 of the pre-patients, blood samples were also obtained at the time of the diagnosis of RA. The plasma levels of 30 cytokines, related factors, and chemokines were measured using a multiplex system.**

Results. **The levels of several of the cytokines, cytokine receptors, and chemokines were significantly increased in individuals before disease onset compared with the levels in control subjects; i.e., those representing signs of general immune activation (interleukin-1β [IL-1β], IL-2, IL-6, IL-1 receptor antagonist, and tumor necrosis factor), activation of Th1 cells (interferon-γ, IL-12), Th2 cells (IL-4, eotaxin), Treg cells (IL-10), bone marrow–derived factors (IL-7, granulocyte–macrophage colony-stimulating factor, and granulocyte colony-stimulating factor), as well as chemokines (monocyte chemotactic protein 1 and macrophage inflammatory protein 1α). The levels were particularly increased in anti–cyclic citrullinated peptide antibody– and rheumatoid factor–positive individuals, and the concentration of most of these increased further after disease onset. The concentration of IL-17 in individuals before disease onset was significantly higher than that in patients after disease onset. Individuals in whom RA subsequently developed were discriminated from control subjects mainly by the presence of Th1 cells, Th2 cells, and Treg cell–related cytokines, while chemokines, stromal cell–derived cytokines, and angiogenic-related markers separated patients after the development of RA from individuals before the onset of RA.**

Conclusion. **Individuals in whom RA later developed had significantly increased levels of several cytokines, cytokine-related factors, and chemokines representing the adaptive immune system (Th1, Th2, and Treg cell–related factors); after disease onset, the involvement and activation of the immune system was more general and widespread.**

Rheumatoid arthritis (RA) is a chronic autoimmune disease characterized by joint inflammation involving the synovial tissue and eventually leading to destruction of cartilage and bone. The etiopathogenic process leading to disease development and progression is not completely understood, although various cells of the immune system and of synovial origin are suggested to be involved (1,2). Numerous cytokines are expressed and are functionally active in the synovial tissue once the disease has developed (3). In samples of synovial fluid, Raza et al (4) observed increased levels of the Th2 cytokines interleukin-4 (IL-4) and IL-13, but not interferon-γ (IFNγ), during the first months of development of RA. IL-17, a proinflammatory cytokine produced by Th17 cells, was also detected at higher levels in early disease compared with late disease (4). An extensive analysis of cytokines and cytokine-related markers in individuals up to 5 years before the diagnosis showed odds ratios (ORs) >2 for 7 of the analytes (IL-1α, IL-1β, IL-4, IL-10, IL-1 receptor antagonist [IL-1Ra], tumor

Supported by the Swedish Research Council (grant K2007-52X-20307-01-3), King Gustaf V's 80-Year Fund, the Swedish Rheumatism Association, the University of Umeå, and in part by the European Community Sixth Framework Programme (project AutoCure).

[1]Heidi Kokkonen, MSc, Ingegerd Söderström, PhD, Solbritt Rantapää Dahlqvist, MD, PhD: Umeå University Hospital, Umeå, Sweden; [2]Joacim Rocklöv, MSc, Göran Hallmans, MD, Kristina Lejon, BSc, PhD: Umeå University, Umeå, Sweden.

Address correspondence and reprint requests to Solbritt Rantapää Dahlqvist, MD, PhD, Department of Rheumatology, University Hospital, SE-901 85 Umeå, Sweden. E-mail: solbritt.rantapaa.dahlqvist@medicin.umu.se.

Submitted for publication March 17, 2009; accepted in revised form October 2, 2009.

necrosis factor α [TNFα], and soluble TNF receptor type I [sTNFRI]) (5). Interestingly, these factors demonstrated the involvement of Th1, Th2, and Treg cells as well as signs of more general immune activation (5). A longer predating time period for anti–cyclic citrullinated peptide (anti-CCP) antibodies and rheumatoid factor (RF) than for increased concentrations of the soluble factors was suggested. However, it was recently reported that the level of sTNFRII was elevated prior to disease onset, predating even the presence of anti-CCP antibodies (6).

We and other investigators have shown that the presence of anti-CCP antibodies precedes the development of RA by several years (7,8). The levels of monocyte chemoattractant protein 1 (MCP-1) in anti-CCP antibody–positive or IgM-RF–positive individuals were significantly increased in those in whom RA developed years later, although a time relationship was not evident (9). Individuals having the combination of anti-CCP antibodies and the shared epitope (SE; HLA–DRB1*0401 and 0404) or the combination of anti-CCP antibodies and the *PTPN22* 1858T variant had a highly increased risk of the development of RA (10,11).

In this study, we analyzed several proinflammatory and antiinflammatory cytokines, cytokine-related factors, and chemokines in blood samples obtained from individuals before the appearance of symptoms of RA, from individuals after the onset of RA, and from control subjects. A consecutive time-dependent involvement of the immune system in disease development and progression was evaluated.

SUBJECTS AND METHODS

Patients and control subjects. A nested case–control study, designed with a 1:3 ratio, was conducted within the Medical Biobank of Northern Sweden. All adult individuals in the county of Västerbotten are continuously invited to participate; consequently, the cohort is population-based. The conditions for recruitment and the collection and storage of blood samples have previously been described (7). The register of patients fulfilling the American College of Rheumatology (formerly, the American Rheumatism Association) classification criteria for RA (12) and with a known date of the onset of symptoms of joint disease was coanalyzed with the registers of the Medical Biobank. Eighty-six individuals (65 women and 21 men) were identified as having donated blood samples before the onset of any symptoms of joint disease. The median (interquartile range) period of time predating the onset of symptoms was 3.3 years (1.1–5.0 years). For every case ("prepatient"), 3 control subjects were randomly selected from the Biobank registers and matched for sex, age at the time of blood sampling, and area of residence. A total of 256 control subjects (194 women and 62 men) were selected. All donors were classified as nonsmoker, always smoker, or former smoker. Sixty-nine of the pre-patients had also provided blood samples while attending the clinic at the time of the diagnosis of RA; the mean $\pm$ SD time to diagnosis was 7.7 $\pm$ 3.6 months after the onset of symptoms. The Regional Ethics Committee at the University Hospital, Umeå, Sweden approved this study, and all participants gave their written informed consent.

Analyses of cytokines, cytokine receptors, and chemokines. Thirty cytokines and chemokines were measured in plasma samples, using multiplex detection kits from Bio-Rad (Hercules, CA). A 27-plex kit was used to measure the concentrations of IL-1β, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8 (CXCL8), IL-9, IL-10, IL-12, IL-13, IL-15, IL-17, eotaxin (CCL11), IL-1Ra, basic fibroblast growth factor (bFGF), granulocyte colony-stimulating factor (G-CSF), granulocyte–macrophage colony-stimulating factor (GM-CSF), IFNγ, IFN-inducible protein 10 (IP-10; CXCL10), MCP-1 (CCL2), macrophage inflammatory protein 1α (MIP-1α; CCL3), MIP-1β (CCL4), platelet-derived growth factor BB (PDGF-BB), RANTES (CCL5), TNFα, and vascular endothelial growth factor (VEGF), and a 3-plex kit was used to measure the concentrations of monokine induced by IFNγ (Mig; CXCL9), macrophage migration inhibitory factor (MIF), and IL-2Rα (CD25). The assay was performed according to the protocol, except that all samples were centrifuged for 20 seconds at 14,300 revolutions per minute to remove debris, and 50 μl of each sample was diluted at a ratio of 1:3 in sample diluent. All samples were assayed in duplicate and analyzed with a Luminex 200 Labmap system (Luminex, Austin, TX).

Data analyses were performed using Bio-Plex Manager software version 4.1.1 (Bio-Rad). Cytokine/chemokine concentrations were interpolated from an appropriate standard curve. In all analyses, an internal control was incorporated to evaluate interplate reproducibility. The internal control consisted of plasma from an admixture of blood samples from 3 patients with established RA. The mean value from the 2 analyses for each factor was used. In samples in which the cytokine/chemokine concentration was below the lowest point on the standard curve, we used the lowest value. One pre-patient and 2 control subjects were excluded from all calculations of cytokine and chemokine concentrations due to being outliers, with extremely high values in all analyses. Measurements of RANTES were unsuccessful, because almost all values calculated were out of range above the standard curve. The median coefficient of variation (CV) was $<$10% for all analyzed markers except IL-6, IL-8, IL-9, IL-15, GM-CSF, MCP-1, and VEGF, for which the CVs were 11–14.5%; for TNF, IL-2, and bFGF, the CVs were 16.2–17.4%.

It is known that samples containing RF of the IgM isotype (IgM-RF) may cause false-positive results in immunoassays by crosslinking the capture and detection antibodies (13). Therefore, we tested several different strategies to block such antibodies. A combination of 40% mouse serum, 20% goat serum, and 20% rabbit serum, which was shown by Raza et al (4) to be a good approach to block RF, was used. We also attempted to eliminate the effect of RF with HeterBlock (Omega Biologicals, Bozeman, MT) and/or protein L (Pierce, Rockford, IL). From the results of the different approaches to blocking IgM-RF, we concluded that none produced reliable results in our studies (data not shown). We were unable to reproduce the results from the various blocking strategies in

different Luminex reader sessions, and because we used duplicate samples in each run, we were able to determine that the CV values were not within an acceptable range. Consequently, no blocking agent was added prior to analysis of the samples. In addition, we analyzed the potential RF cross-binding activity by performing "mismatch simplex sandwich" tests, i.e., we used different specificities of the capture and detection antibodies. More specifically, 50 μl of anti–IL-2–coated beads was incubated with 50 μl of plasma from patients known to display high RF levels. Subsequently, after washing steps, the beads were incubated with 25 μl of phycoerythrin-labeled antieotaxin antibodies before analysis in the Luminex reader. In none of these analyzed mismatch assays did we detect any signal whatsoever, strongly suggesting that RF cannot act as a nonspecific bridge between the capture and detection antibodies in the assay.

Analyses of genetic factors and autoantibodies. HLA–DRB1 genotyping for 0404 and 0401 was performed using polymerase chain reaction (PCR) sequence-specific primers from an HLA—DR low resolution kit and an HLA–DRB1*04 subtyping kit (Olerup SSP, Saltsjöbaden, Sweden). The *PTPN22* 1858C/T polymorphism (rs2476601) was determined using the 5′-nuclease assay, as described previously (10,11). The PCRs were performed according to the manufacturer's instructions, and detection of the different genotypes was performed using an ABI PRISM 7900HT Sequence Detector System (Applied Biosystems, Foster City, CA). The levels of anti-CCP2 antibodies and IgM-RF were determined using enzyme-linked immunoassays, as previously described (7).

Statistical analysis. Statistical calculations were performed using SPSS for Windows, version 14.0 (SPSS, Chicago, IL). Continuous data were compared (pre-patients versus patients) by nonparametric analyses with Wilcoxon's signed rank test for matched pairs and conditional regression analyses (pre-patients versus control subjects). When data were stratified, the subgroups were compared with simple logistic regression analyses adjusted for age and sex. Because most factors were not normally distributed, calculations were performed on the logarithm of the measured factors. Correlation analyses were performed using Pearson's correlation analyses. Cutoff values, defining positivity used in logistic regression analyses and for calculations of sensitivity and specificity, were defined as positive if concentrations were above the 95th percentile of the levels of control subjects, as previously described (5). Regarding the study as explorative, P values less than or equal to 0.05 were considered significant. Relationships between categorical data (positive versus negative) were compared using chi-square analyses or Fisher's exact test, when appropriate.

To identify groups of cytokines, cytokine-related factors, and chemokines that allow the distinction between control subjects, individuals before the onset of symptoms (pre-patients), and individuals after the onset of symptoms (patients), we used a multivariate classification algorithm termed the random forest method (randomForest 4.5-25, R-project) (14,15). Applying this method, the 29 analyzed factors and anti-CCP antibodies were used to classify disease states, taking advantage of combinations of the predictors. The basic principle of the random forest method is similar to that of decision trees; however, to classify objects to a specific disease state, many bootstrap samples are obtained from the original data. Each of the bootstrap samples is used to create decision trees independent of each other. On average, one will not use more than approximately two-thirds of the observations in each tree construction, due to replacement in the bootstrap procedure. The resulting one-third of the observations are used to validate each tree structure. At each branch in any decision tree, only a sample of all the predictors will be tested as the best splitter. Sampling ~5 predictors at each branch seems to perform close to optimally in this case. When predicting the class of an observation from the model, each tree votes for a disease state, and then the disease state with the highest number of votes across all trees will give the overall disease state. The random forest estimates of classification accuracy are unbiased, and the models are robust to overfitting due to properties of the sampling. Generally, this approach has been shown to perform well compared with other classifiers such as discriminate analysis, support vector machines, and neural networks (16). The importance of each predictor variable in the classification process is estimated by permuting the value of the predictor in all trees at all branches where it is applied. Thereafter, the mean decrease in prediction performance associated with this predictor can be calculated. The discrimination of the random forest models can be visualized graphically by using multidimensional scaling of the proximity matrix (17), where the proximity is a relative measure of pairwise relationships derived from the set of trees. The prediction performance of the random forest model can be translated to sensitivity and specificity.

RESULTS

The frequency distributions of anti-CCP antibodies, RF, the HLA SE, and the *PTPN22* 1858T variant as well as clinical and demographic data for the pre-patients, patients at disease onset, and control subjects are presented in Table 1.

Pre-patients versus control subjects. Analysis of the plasma samples using a multiplex bead–based system

Table 1. Characteristics of the pre-patients, patients at disease onset, and control subjects*

	Pre-patients (n = 85)	Patients at disease onset (n = 69)	Controls (n = 254)
Age, mean (range) years	52.3 (30–69)	56.4 (34–72)	52.3 (30–69)
No. of women/no. of men	64/21	52/17	192/62
Anti-CCP antibody positive	32 (37.6)†	47 (68.1)†	3 (1.4)
IgM-RF positive	23 (27.1)†	57 (82.6)†	10 (6.0)
PTPN22 1858T carrier	33 (42.9)†	29 (42.6)†	40 (19.1)
HLA SE carrier	47 (55.3)‡	41 (59.4)‡	47 (35.9)
Smoking ever	46 (59.0)‡	40 (59.7)‡	79 (38.5)
Prednisolone treatment	–	27 (39.1)	–

* Except where indicated otherwise, values are the number (%). Pre-patients were defined as individuals in whom the onset of joint symptoms had not yet occurred. Anti-CCP = anti–cyclic citrullinated peptide; RF = rheumatoid factor; SE = shared epitope.
† $P < 0.001$ versus controls.
‡ $P < 0.01$ versus controls.

Table 2. Cytokines, cytokine receptors, and chemokines in controls, pre-patients, and patients*

Cytokine/chemokine	Controls (n = 254)	Pre-patients (n = 85)	*P*	Patients (n = 69)	*P*
General activation					
IL-1β	2.6 (0.8–3.9)	3.4 (1.2–5.6)	4.0×10^{-4}	5.5 (2.2–22.6)	0.002
IL-1Ra	132.6 (89.2–198.4)	155.0 (107.2–243.8)	0.003	234.6 (134.9–537.5)	1.8×10^{-4}
IL-2Rα	36.5 (25.0–51.7)	39.1 (25.5–62.0)	0.153	73.5 (36.1–123.9)	5.5×10^{-6}
TNFα	35.3 (14.2–61.7)	42.7 (12.8–109.6)	0.013	35.6 (15.3–284.9)	0.223
IL-6	4.6 (1.1–10.8)	6.5 (2.2–17.6)	0.001	37.1 (12.7–189.6)	4.1×10^{-7}
IL-2	1.1 (1.1–9.2)	4.7 (1.1–22.8)	0.001	18.0 (1.1–152.4)	0.001
IL-15	0.5 (0.3–4.2)	0.5 (0.3–4.2)	0.182	1.8 (0.3–4.2)	0.007
Th1-related					
IL-12	15.8 (8.6–27.2)	21.5 (10.9–41.2)	0.001	30.5 (15.7–256.9)	3.7×10^{-4}
IFNγ	77.1 (48.4–127.6)	92.8 (56.1–190.4)	2.7×10^{-4}	164.8 (71.2–793.2)	0.002
Th2-related					
IL-4	2.7 (1.9–3.8)	3.2 (2.3–4.8)	2.7×10^{-4}	5.1 (3.0–27.0)	2.2×10^{-4}
IL-5	4.4 (2.1–6.3)	4.9 (2.6–8.1)	0.042	4.2 (2.1–9.3)	0.561
IL-9	12.4 (0.9–35.5)	20.0 (5.3–95.9)	0.005	51.0 (7.9–321.5)	0.014
IL-13	4.4 (2.7–7.9)	5.1 (3.4–13.2)	0.019	6.5 (3.6–19.5)	0.138
Eotaxin	30.8 (18.2–42.8)	41.6 (21.9–79.2)	0.001	109.7 (33.3–312.7)	5.0×10^{-4}
Th17-related					
IL-17	21.1 (6.5–38.5)	28.7 (9.3–39.6)	0.149	6.0 (0.2–26.3)	6.1×10^{-5}
Treg cell–related					
IL-10	4.3 (2.4–6.6)	5.0 (2.8–9.4)	0.024	7.6 (2.8–21.8)	0.002
Bone marrow–derived					
IL-7	21.3 (13.9–30.3)	24.1 (15.0–35.6)	0.030	23.9 (16.1–46.7)	0.058
GM-CSF	5.6 (2.3–14.7)	11.7 (3.2–30.8)	0.003	16.4 (2.1–90.2)	0.065
G-CSF	52.4 (42.3–68.5)	59.1 (43.6–75.5)	0.010	59.4 (34.0–89.6)	0.669
Stromal cells and angiogenic factors					
bFGF	6.8 (2.2–6.8)	6.8 (2.2–6.8)	0.268	6.6 (2.2–6.8)	0.478
PDGF-BB	1,571.2 (987.0–2,318.2)	1,736.8 (969.2–2,462.4)	0.304	6,265.4 (1,371.9–13,960.2)	1.7×10^{-6}
VEGF	11.0 (4.6–20.3)	14.4 (6.5–30.2)	0.051	56.5 (15.0–156.3)	2.0×10^{-6}
Chemokines					
MIF	250.7 (115.7–444.7)	207.8 (112.3–421.2)	0.196	375.3 (167.6–728.6)	0.053
Mig	294.2 (188.5–394.0)	342.0 (240.0–484.4)	0.059	707.8 (431.6–1,433.5)	1.7×10^{-8}
IL-8	3.9 (1.0–8.2)	3.7 (0.6–8.3)	0.882	7.3 (1.4–19.0)	1.8×10^{-4}
IP-10	591.7 (372.3–854.1)	702.8 (391.5–1,077.7)	0.109	1,039.2 (442.1–1,961.6)	0.001
MCP-1	16.0 (10.6–24.0)	21.6 (14.8–38.5)	0.003	42.9 (20.5–96.8)	1.5×10^{-4}
MIP-1α	6.9 (4.9–9.6)	8.0 (5.9–10.6)	0.011	6.1 (4.1–9.1)	0.169
MIP-1β	35.4 (25.8–45.3)	36.3 (27.4–42.8)	0.823	44.2 (19.2–78.6)	0.001

* Values are the median (interquartile range) pg/ml. Statistical analyses were performed using conditional logistic regression for pre-patients (defined as individuals in whom the onset of joint symptoms had not yet occurred) versus controls and Wilcoxon's rank sum test for matched pairs (pre-patients versus patients). IL-1β = interleukin-1β; IL-1Ra = IL-1 receptor antagonist; IL-1Rα = IL-1 receptor α; TNFα = tumor necrosis factor α; IFNγ = interferon-γ; GM-CSF = granulocyte–macrophage colony-stimulating factor; G-CSF = granulocyte colony-stimulating factor; bFGF = basic fibroblast growth factor; PDGF-BB = platelet-derived growth factor BB; VEGF = vascular endothelial growth factor; MIF = macrophage migration inhibitory factor; Mig = monokine induced by IFNγ; IP-10 = IFN-inducible protein 10; MCP-1 = monocyte chemoattractant protein 1; MIP-1α = macrophage inflammatory protein 1α.

showed that concentrations of two-thirds of the proinflammatory cytokines, IL-1β, TNFα, IL-6, IL-2, IL-12, IFNγ, IL-4, eotaxin, IL-10, and IL-7, were significantly increased in pre-patients compared with matched control subjects (Table 2). In 50 of the 85 pre-patients, levels of at least two-thirds of the cytokines or cytokine-related factors were above the median values for control subjects. These cytokines represented both antiinflammatory and proinflammatory responses and were related to both the Th1 and Th2 lineage and to Treg cells. IL-17 expression was increased in pre-patients, although the difference did not reach statistical significance. Of the chemokines, the levels of MCP-1 and MIP-1α were significantly increased in pre-patients compared with control subjects, as were the levels of GM-CSF and G-CSF.

Sensitivity and specificity for RA in pre-patients, and associations with anti-CCP antibodies. The sensitivity and specificity for the development of RA were calculated for the cytokines, cytokine-related factors, and chemokines defined as positive above the 95th percentile of the value for control subjects. The highest

Table 3. Sensitivity and specificity for the development of RA in pre-patients versus controls*

Variable	Sensitivity	Specificity	OR	95% CI
Anti-CCP	38.7	98.6	41.4	12.2–140.6
IgM-RF	27.4	94.0	5.9	2.66–13.16
Eotaxin	22.4	95.3	5.8	2.70–12.62
IL-1Ra	18.8	95.3	4.7	2.12–10.40
IL-2	18.8	95.3	4.7	2.12–10.40
GM-CSF	18.8	95.3	4.7	2.12–10.40
IFNγ	18.8	95.3	4.7	2.12–10.40
IL-4	17.6	95.3	4.3	1.94–9.70
IL-9	17.6	95.3	4.7	2.08–10.80
TNFα	16.5	95.3	4.0	1.77–9.02
IL-12	16.5	95.3	4.0	1.77–9.02
IL-10	16.5	95.3	4.0	1.77–9.02
IL-1β	16.5	95.3	4.0	1.77–9.02
IL-6	15.3	95.3	3.6	1.60–8.36
MCP-1	14.3	95.3	3.8	1.47–9.56
IL-15	10.6	95.3	2.6	1.05–6.55
IP-10	10.6	95.7	2.6	1.05–6.55

* Values are the percentage and were calculated for the cytokines, cytokine-related factors, and chemokines defined as positive above the 95th percentile of the value for control subjects. Variables are presented with the lower 95% confidence interval (95% CI) limit >1.0. Pre-patients were defined as individuals in whom the onset of joint symptoms had not yet occurred. RA = rheumatoid arthritis; OR = odds ratio; anti-CCP = anti–cyclic citrullinated peptide; IgM-RF = IgM rheumatoid factor; IL-1Ra = interleukin-1 receptor antagonist; GM-CSF = granulocyte–macrophage colony-stimulating factor; IFNγ = interferon-γ; TNFα = tumor necrosis factor α; MCP-1 = monocyte chemoattractant protein 1; IP-10 = IFN-inducible protein 10.

sensitivity for predicting RA, following that for anti-CCP antibodies and of the same magnitude as RF, was for eotaxin, with a sensitivity of 22.4% and an OR of 5.8 (95% confidence interval [95% CI] 2.70–12.62) at a preset specificity of 95.3%. Among the cytokines and cytokine-related factors with significant predictive values for the development of RA (95% CI >1), one-half were related to Th1, Th2, and Treg cells (Table 3). The combination of anti-CCP antibodies and the cytokines stratified for Th1 (IFNγ and/or IL-12), Th2 (IL-4, IL-5, IL-9, IL-13, and/or eotaxin), or general activation (IL-1β, IL-2, IL-6, IL-7, IL-15, IL-1Ra, IL-2Rα, TNF, and/or GM-CSF) yielded sensitivities of 14.1%, 12.9%, and 20.0%, respectively, at a specificity of 100% for all, because none of the control subjects had this combination.

In pre-patients, significant associations were observed between anti-CCP antibodies and cytokine positivity, as defined above, and were more prominent for Th2-related cytokines as one group ($\chi^2 = 14.6$, $P < 0.0001$) than for Th1-related cytokines ($\chi^2 = 5.6$, $P < 0.05$) or cytokines involved in more general immune activation ($\chi^2 = 7.1$, $P < 0.01$) or for Th17 ($\chi^2 = 4.06$, $P < 0.05$).

Time relationships in pre-patients. Time-dependent analyses revealed that the concentrations of most of the analyzed variables increased in relation to the closer to the onset of symptoms that the plasma sample was collected (additional information available from the corresponding author). However, stratifying for more or fewer than 3 years before the onset of symptoms of RA, statistical significance was reached for IL-2Rα ($P = 0.035$), IL-1β ($P = 0.039$), IL-9 ($P = 0.002$), IL-10 ($P = 0.036$), eotaxin ($P = 0.021$), GM-CSF ($P = 0.018$), bFGF ($P = 0.038$), IP-10 ($P = 0.029$), and Mig ($P = 0.003$).

Patients versus pre-patients and control subjects. In samples obtained from patients at the time of the diagnosis of RA, the concentrations of all analytes except IL-17, MIF, MIP-1α, and bFGF were significantly increased compared with the concentrations in samples from control subjects. In the comparison between matched pairs, i.e., the same individual sampled before and after the onset of disease (n = 69), expression of most of the analyzed variables increased further, with the exception of TNF, IL-5, IL-7, IL-13, GM-CSF, G-CSF, bFGF, MIP-1α, MIF, and IL-17 (Table 2). There was actually a significant decrease in the IL-17 concentration after disease onset.

Stratifications for anti-CCP antibodies and RF in pre-patients and patients. Stratification for anti-CCP antibodies in pre-patients resulted in the identification of additional significantly increased levels of IL-2Rα ($P = 0.001$), which after stratification reached significant levels in anti-CCP antibody–positive individuals (n = 32) compared with control subjects (additional information available from the corresponding author). Stratification for IgM-RF showed that the levels of IL-2Rα, Th2-related cytokines (IL-5, IL-9, IL-13), IL-17, and Mig were now significantly higher in RF-positive pre-patients (n = 23) compared with control subjects, whereas the levels of TNF and IL-7 were not significantly increased. In RF-negative pre-patients (n = 61), the significance was lost for most of the cytokines compared with control subjects (data not shown) but was gained for IL-13 and IL-8. No significant correlation between RF and each separate factor was observed.

Among the patients, the levels of most cytokines, cytokine-related factors, and chemokines were significantly increased compared with those in control subjects and remained so after stratification for anti-CCP antibody–positive patients (n = 47) (additional information available from the corresponding author). The levels of some of the cytokines and most of the chemokines remained increased in anti-CCP antibody–negative patients (n = 22). In RF-positive patients (n =

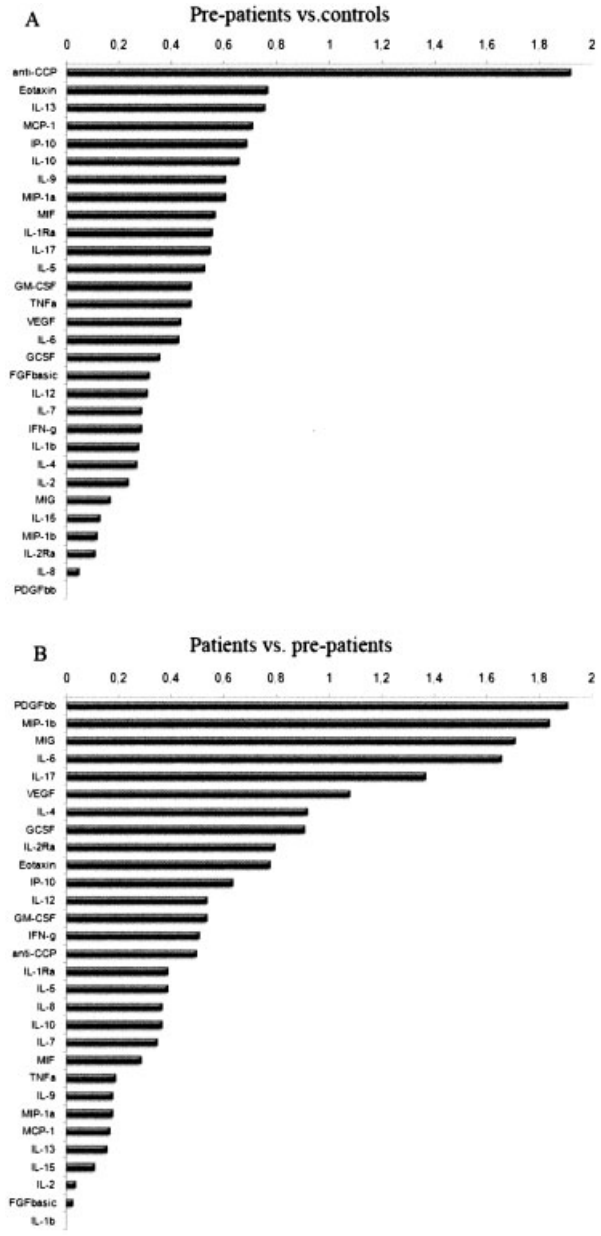

Figure 1. The order of factors in terms of their accuracy for discriminating pre-patients from control subjects and patients from pre-patients. Pre-patients were defined as individuals in whom symptoms of rheumatoid arthritis (RA) had not yet occurred; patients were defined as the same individuals after the onset of RA. The horizontal axis represents the average decrease in classification accuracy when the values for each factor were permutated for pre-patients compared with matched control subjects (**A**) and for patients compared with pre-patients (**B**). Anti-CCP = anti–cyclic citrullinated peptide; IL-13 = interleukin-13; MCP-1 = monocyte chemoattractant protein 1; IP-10 = interferon (IFN)–inducible protein 10; MIP-1α = macrophage inflammatory protein 1α; MIF = macrophage migration inhibitory factor; IL-1Ra = IL-1 receptor antagonist; GM-CSF = granulocyte–macrophage colony-stimulating factor; TNFα = tumor necrosis factor α; VEGF = vascular endothelial growth factor; G-CSF = granulocyte colony-stimulating factor; FGF = fibroblast growth factor; Mig = monokine induced by IFNγ; MIP-1β = macrophage inflammatory protein 1β; PDGF-BB = platelet-derived growth factor BB.

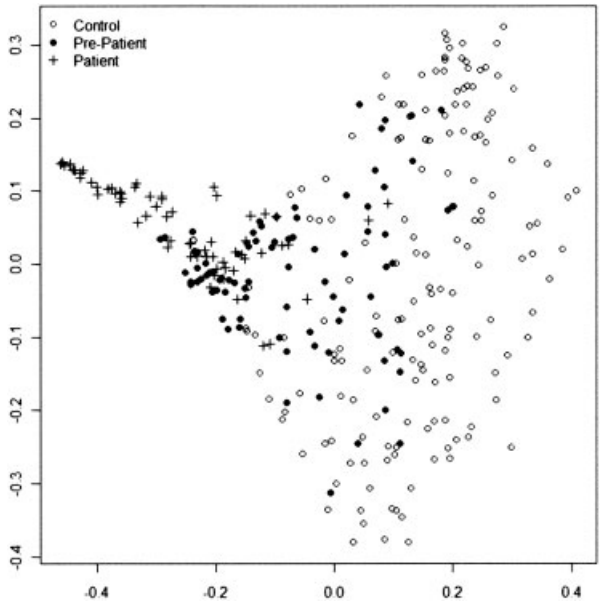

Figure 2. Multidimensional scaling using random forest modeling (summarizing all factors), demonstrating the clustering of control subjects, pre-patients, and patients. Pre-patients were defined as individuals in whom symptoms of rheumatoid arthritis (RA) had not yet occurred; patients were defined as the same individuals after the onset of RA. The 2 axes represent the dominant clustering directions between the groups.

57), the levels of all of the markers except IL-7, IL-8, G-CSF, IP-10, bFGF, MIP-1α, MIP-1β, and MIF were significantly increased compared with the levels in control subjects. In RF-negative patients (n = 12), significantly elevated concentrations in patients compared with control subjects were observed for IL-6, eotaxin, Mig, IP-10, GM-CSF, PDGF-BB, and TNF ($P < 0.05$–0.001). The levels of all cytokines, VEGF, G-CSF, and MCP-1 correlated with the levels of RF (median correlation coefficient 0.478, range 0.298–0.696).

Random forest analyses. The relative importance of individual cytokines, cytokine-related factors, and chemokines in classifying between the different disease states (pre-patient and patient) and control subjects was determined by random forest modeling. The order of the factors in terms of their accuracy to discriminate pre-patients from control subjects and patients from pre-patients is shown in Figure 1. Important cytokines and chemokines are associated with their contribution to reducing accurate classification. Following anti-CCP antibodies, the most important factors for discriminating pre-patients from control subjects were eotaxin, IL-13, MCP-1, IP-10, IL-10, IL-9, MIP-1α, MIF, and IL-1Ra (Figure 1A), whereas PDGF-BB, MIP-1β, Mig, IL-6, IL-17, VEGF, IL-4, and G-CSF, in that order, were the most important cytokines and chemokines for discrimi-

nating between pre-patients and patients at the onset of disease (Figure 1B). Anti-CCP antibodies, PDGF-BB, IL-6, MIP-1β, Mig, eotaxin, IL-10, and G-CSF were the factors that best distinguished between patients after disease onset and control subjects (data not shown).

Based on a summary of the random forest modeling, including all analyzed cytokines and chemokines (multidimensional scaling), the relationship and gradual progression between control subjects, pre-patients, and patients could be observed (Figure 2). The random forest analyses gave a sensitivity of 51.2% for predicting the development of RA among pre-patients and control subjects, with all analyzed factors and anti-CCP antibodies included, with a specificity of 91.9%. When only anti-CCP antibodies were included in the model, the sensitivity was 36.9%, and the specificity increased to 98.8%. When all variables were included, the sensitivity for discriminating patients from pre-patients was 72.7%, and the specificity was 84.5%; when only anti-CCP antibodies were considered, the sensitivity was 68.2%, and the specificity was 63.1%. When all analyzed factors were included, the sensitivity for discriminating patients from control subjects was 86.4%, and the specificity was 95.0%; when only anti-CCP antibodies were considered, the sensitivity decreased to 68.2%, and the specificity increased to 98.8%.

Stratification or adjustment for gene carriage of *PTPN22* 1858T, HLA SE alleles, smoking habits, or sex did not add any further information or reveal any statistical differences.

DISCUSSION

In this explorative study comparing individuals before the onset of symptoms (pre-patients) and after the onset of RA with matched control subjects, we observed several cytokines, cytokine-related factors, and chemokines to be up-regulated. In pre-patients, the most prominent of these were the Th2-related cytokines, eotaxin, and IL-4, followed by Th1 cytokines, IL-12, IFNγ, and IL-10. Furthermore, by using random forest modeling, it was evident that those cytokines best distinguishing pre-patients from control subjects were related to Th1, Th2, and Treg cells, representing the adaptive immune system. The factors separating individuals before and after the onset of disease represented not only a more general immunologic response but also stromal cells and angiogenic factors. It was also evident from the analysis models (i.e., random forest) that these factors, in addition to anti-CCP antibodies, further enhanced identification of the individuals in whom RA would subsequently develop.

The finding of an early (even before the onset of symptoms) elevation in the levels of IL-4 and IFNγ is consistent with a previous report by Jørgensen et al (5). In the patients with very early RA in the study by Raza et al (4), IL-4 was among the 5 most important variables in synovial fluid, whereas IFNγ was not. In the present study, another Th2-related cytokine, eotaxin, was among the most significant cytokines appearing early. An interesting finding was that IL-17, which is suggested to be a proinflammatory cytokine, was present at its highest concentrations in pre-patients, and the level had already decreased within 7.7 months following the onset of disease. This observation endorses the role of IL-17 in the initiating phase, and, as the pathogenesis progresses, other factors are subsequently brought into action. The diagnostic sensitivity of the individual cytokines and chemokines was relatively low in pre-patients, although the OR, with the 95% CI, was significantly related to the development of RA.

There was a clear trend for the concentrations of many of the cytokines and chemokines to be increased the closer to the onset of symptoms that plasma samples were collected, although statistical significance was reached only for some of them, and the increased concentrations were related to the presence of anti-CCP antibodies. We observed a clear relationship between cytokines related not only to Th1, Th2, and Treg cells but also to Th17 (above the 95th percentile of control subjects) and the presence of anti-CCP antibodies, thereby supporting the concept that the immune system was already up-regulated and disease was developing toward RA. The highest significance level of a relationship was for Th2-related cytokines (as a group) and anti-CCP antibodies. However, sensitivity was not increased by the combination of anti-CCP antibodies and cytokines stratified for T cell subtype. Our data do not allow us to conclude that the presence of anti-CCP antibodies predated increased cytokine and/or chemokine levels, as was suggested by Jørgensen et al (5).

Of the chemokines, MCP-1 was increased most prominently. In a previous study using an enzyme-linked immunosorbent assay to analyze plasma from individuals before disease onset, the level of MCP-1 was increased only in anti-CCP antibody–positive individuals (9). The most significant finding by Jørgensen et al (5) was the increased level of TNF, which was also significantly increased among pre-patients compared with control subjects in this study, although the increase that we observed was not of the same magnitude as that reported by Jørgensen et al. In addition, it was recently shown that expression of sTNFRII, which could be a pseudomarker for TNF, was significantly increased before disease onset (6).

When considering the patients in whom RA

developed, the immunologic response has expanded and involved, in addition to macrophages and T cells, stromal cells, fibroblasts, and other cells producing angiogenic factors, although the patients had early disease (the mean duration of symptoms was only 7.7 months).

When the pre-patients and patients were stratified according to the presence of autoantibodies, i.e., anti-CCP antibodies and IgM-RF, some of the significant differences were lost in both autoantibody-positive and antibody-negative individuals and patients. In pre-patients, the presence of anti-CCP antibodies and RF could indicate initiation of the disease process and explain the higher levels of some of the cytokines and chemokines. In these analyses, however, the number of individuals in each subgroup was relatively small. Additionally, one must bear in mind the possibility of false-positive binding of the antibody, particularly that of IgM-RF, which could contribute to the increased levels detected. According to their published reports, some investigators have added blocking agents (4), while others have not discussed this matter (5). Before the present study was undertaken, we evaluated various blocking protocols without producing reproducible results; consequently, we preferred not to use any blocking agent.

Statistical analyses showed a very low variable correlation between the concentration of IgM-RF and some, but not all, of the cytokines. This observation provides evidence against a general, nonspecific binding of IgM-RF to the cytokine-specific monoclonal antibodies attached to the Luminex beads. Furthermore, this interference would require that the detection antibody also bound the IgM-RF. We tested the ability of IgM-RF to act as a nonspecific bridge between the capture and detection antibodies displaying different specificities. In these assays, we failed to detect any signal at all, even though the sera used contained high levels of IgM-RF. Moreover, in our random forest analysis, a clear difference in the classification accuracy between the different cytokines was observed. This result would be unexpected if the cytokine levels were merely a reflection of IgM-RF levels, because under such circumstances all cytokines would have contributed equally.

We are aware of the statistical limitations of the present study, mainly in terms of power, due to the low number of samples from the same patients before and after the onset of disease. However, we believe that the use of samples from the same individuals adds strength to this type of study. The potential effects of storage time must be considered, because some of the samples have been stored in the Medical Biobank of Northern Sweden for several years. However, we compensated for this effect by selecting control subjects who were matched for the date of sampling and storage conditions, i.e., the storage conditions were equivalent for both patients and control subjects.

This explorative study can help to generate an hypothesis regarding the development of RA, particularly in terms of the role of the immune system in the initiation of the disease. Based on our results, we are unable to conclude which agent initiates pathogenesis, but we can summarize the processes that are activated. We envision that in pre-patients, compared with control subjects, there is continuous enhanced bone marrow production of granulocytes and lymphocytes, which supports initiation of the inflammatory process, as detected by increased levels of G-CSF, GM-CSF, and IL-7. Both Th1 and Th2 cells are engaged with notable involvement of Th17 cells, primarily in the initiation phase of the disease. Moreover, levels of chemokines (e.g., MCP-1, MIP-1α) promoting both Th1 and Th2 cells and monocyte migration are enhanced. The macrophages are up-regulated by IFNγ and IL-10 to produce IL-1β, IL-1Ra, TNFα, and IL-6. At this stage, the process is not necessarily located in the joints but could occur in the peripheral lymphoid organs. In the later stages of disease development process (i.e., in patients), some cytokines such as soluble TNFα are no longer as abundant, whereas the expression of others, such as IL-6, IL-1β, and IL-1Ra, is clearly enhanced. Furthermore, enhancement of factors involved in tissue remodeling, such as VEGF and PDGF-BB, is also up-regulated, facilitating the growth of synovial tissue leading to pannus formation. In addition, the levels of chemokines (e.g., MIF, Mig, and IP-10) promoting Th1 and Th2 cells and monocyte migration are further elevated.

Based on the results of this study, we conclude that blood samples obtained from individuals before the onset of symptoms of RA have elevated concentrations of proinflammatory cytokines, cytokine-related factors, and chemokines, indicating activation of the immune system. In the present study, such activation occurred before any symptoms of joint involvement. These findings present an opportunity for better predicting the risk of developing RA and, therefore, possibly preventing disease progression.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Rantapää Dahlqvist had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Kokkonen, Söderström, Rantapää Dahlqvist.
Acquisition of data. Kokkonen, Söderström, Rantapää Dahlqvist.
Analysis and interpretation of data. Kokkonen, Söderström, Rocklöv, Hallmans, Lejon, Rantapää Dahlqvist.

REFERENCES

1. Feldmann M, Brennan FM, Maini RN. Role of cytokines in rheumatoid arthritis. Annu Rev Immunol 1996;14:397–440.
2. Klareskog L, Alfredsson L, Rantapaa-Dahlqvist S, Berglin E, Stolt P, Padyukov L. What precedes development of rheumatoid arthritis? Ann Rheum Dis 2004;63(Suppl 2):ii28–31.
3. McInnes IB, Schett G. Cytokines in the pathogenesis of rheumatoid arthritis. Nat Rev Immunol 2007;7:429–42.
4. Raza K, Falciani F, Curnow SJ, Ross EJ, Lee CY, Akbar AN, et al. Early rheumatoid arthritis is characterized by a distinct and transient synovial fluid cytokine profile of T cell and stromal cell origin. Arthritis Res Ther 2005;7:R784–95.
5. Jorgensen KT, Wiik A, Pedersen M, Hedegaard CJ, Vestergaard BF, Gislefoss RE, et al. Cytokines, autoantibodies and viral antibodies in premorbid and postdiagnostic sera from patients with rheumatoid arthritis: case-control study nested in a cohort of Norwegian blood donors. Ann Rheum Dis 2008;67:860–6.
6. Karlson EW, Chibnik LB, Tworonger SS, Lee IM, Buring JE, Shadick NA, et al. Biomarkers of inflammation and development of rheumatoid arthritis in women from two prospective cohort studies. Arthritis Rheum 2009;60:641–52.
7. Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003;48:2741–9.
8. Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004;50:380–6.
9. Rantapaa-Dahlqvist S, Boman K, Tarkowski A, Hallmans G. Up regulation of monocyte chemoattractant protein-1 expression in anti-citrulline antibody and immunoglobulin M rheumatoid factor positive subjects precedes onset of inflammatory response and development of overt rheumatoid arthritis. Ann Rheum Dis 2007;66:121–3.
10. Berglin E, Padyukov L, Sundin U, Hallmans G, Stenlund H, van Venrooij W, et al. A combination of autoantibodies to cyclic citrullinated peptide (CCP) and HLA-DRB1 locus antigen is strongly associated with future onset of rheumatoid arthritis. Arthritis Res Ther 2004;6:R303–8.
11. Johansson M, Arlestig L, Hallmans G, Rantapaa-Dahlqvist S. PTPN22 polymorphism and anti-cyclic citrullinated peptide antibodies in combination strongly predicts future onset of rheumatoid arthritis and has a specificity of 100% for the disease. Arthritis Res Ther 2005;8:R19.
12. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
13. De Jager W, Rijkers GT. Solid-phase and bead based cytokine immunoassay: a comparison. Methods 2006;38:294–303.
14. R Development Core Team (2008). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL: http://www.R-project.org.
15. Breiman L. Manual on setting up, using, and understanding random forests V3.1. URL: http://oz.berkeley.edu/users/breiman/Using_random_forests_V3.1.pdf.
16. Breiman L. Random forests. Machine Learning 2001;45:5–32.
17. Cox TF, Cox MA. Multidimensional scaling. London: Chapman and Hall; 2000.