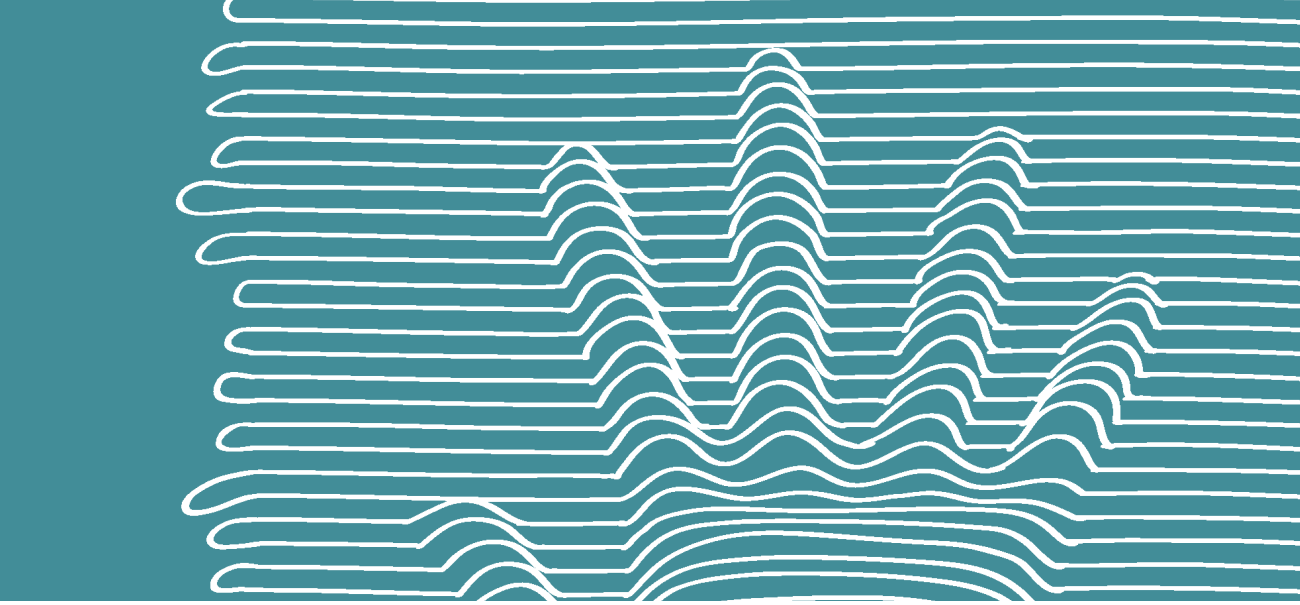

CHAPTER 9

Summary and discussion

Rheumatoid arthritis is a chronic auto-immune disorder, of which persistent synovitis, bone erosions and auto-antibody formation are characteristic features[1]. Although the etiology of the disease remains largely unknown, it is established that genetic risk factors play a pivotal role in disease pathology. Both family and twin studies have shown that the genetic contribution to the disease can be estimated around 50%[2-4].

The most important genetic risk factor for RA was first described in the 1970's and is located in the HLA region. The associated HLA-DRB1 alleles all encode a similar amino-acid sequence at a particular position of the hyper variable region of the HLA class II molecule. This finding resulted in the SE hypothesis, in which was proposed that the SE motif is directly involved in RA pathogenesis. It was postulated that in this binding groove, the HLA class II molecule could present a specific arthritogenic peptide to auto-reactive T-cells[5, 6]. Although it has been shown that epitopes derived from human citrullinated vimentin can be recognized by HLA-DRB1*0401 restricted T-cells, the hypothesis still remains to be elucidated[7].

After the identification of the HLA-DRB1 alleles, it was not until 2004 that the first non-HLA locus that associated with RA susceptibility was identified[8]. Due to advances in genotyping technologies, which could be performed at reasonable costs, the identification of novel genetic risk factors took a flight[9].

In the current thesis the genetic contribution of non-HLA genes to RA susceptibility was further investigated and the functional relevance of these loci in disease pathology was explored.

Establishing genetic risk factors

The recent burst in genetic association studies for disease susceptibility has led to the identification of many risk loci associated with RA susceptibility. Although part of these newly identified loci will truly be associated with RA, part will show to be type I errors, in which the associations appear to be false positive findings. To separate the true findings from the false positive ones, it is of utmost importance to replicate newly identified genes in independent populations of similar origin. In both chapter 2 and chapter 3, such replication studies for recently identified genes were performed.

In chapter 2 three SNPs in STAT4, IL2/IL21 and CTLA4 genes, were investigated in the Leiden RA population. Subsequently, a meta-analysis of all publicly available genotype data was performed to identify the overall genetic association with RA susceptibility. The association with the STAT4 gene was first described in 2007 and replicated consistently in several Caucasian and Asian populations[10-12]. The IL2/IL21 region, on the other hand, had only been described in one study thus far and the CTLA4 gene had shown inconsistent results in several replication studies[13-16]. In the present study, both the SNPs in the STAT4 gene and the CTLA4 gene could be independently replicated in our Dutch RA population, while the SNP in the IL2/IL21 region showed a clear trend towards association. More importantly, all three loci were further confirmed in a well powered meta-analysis, indicating these loci as true genetic risk factors for RA susceptibility.

In chapter 3 a replication study of two SNPs located in the interleukin 2 pathway is described. A recent GWAS in the WTCCC population had indicated the IL2RA and IL2RB regions, located in this pathway, to be novel genetic risk factors for RA[17]. These findings were replicated in our independent Dutch RA population, thereby providing further evidence for true association.

Missing heritability

In addition to the in chapter 2 and chapter 3 described loci, currently over 46 genetic regions have been identified as true genetic risk factors with statistically robust associations. These associations however, confer relatively small increments in risk and thereby contributing only marginally to disease susceptibility[9, 18]. It is thought therefore, that a great part of the genetic risk for RA has not yet been explained. Even including the HLA region, it is likely that more than 50% of the genetic risk for RA remains to be identified. Many explanations for this 'missing heritability' have been suggested, including structural variants that were poorly captured by the existing genotyping arrays, gene-gene interactions and inadequate account of the shared environment among relatives. An additional possibility, on which most speculations have been focused, is that multiple rare variants contribute to disease risk in a significant subset of the RA population. Rare variants in this matter are defined as variants with an allele frequency less than 5%, or even less than 1%, but not rarer than 0.5%. Such variants are not sufficiently frequent to be captured by the current GWAS arrays, nor do they have large enough effect sizes to be identified by linkage analysis in family studies. These variants can be identified by resequencing genomes of patients and healthy individuals[19]. These low frequency variants will have higher effect sizes than the currently identified non-HLA risk loci and could thereby contribute substantially to the current 'missing heritability'.

Next to the investigation of rare variants, structural variants and gene-gene interactions to explore the 'missing heritability', additional common variants can be identified by candidate gene approach. Although the contribution of these variants to disease risk have shown to be marginal and the candidate gene approach has its disadvantages compared to GWAS, a candidate gene study might identify a risk factor that has simply not been typed or identified in the large scale assays.

In chapter 5 a study is described, in which common variants located in the VTCN1 region were investigated for association with RA susceptibility by candidate gene approach. VTCN1 encodes a protein, which has been reported to be a negative regulator of T-cell responses in vitro by inhibiting proliferation and cytokine production[20, 21]. In RA patients the soluble form of the protein is more frequently observed than in healthy individuals and is correlated with disease activity as measured by DAS 28[22]. In knock-out mice, the progression of inflammation in a collagen induced arthritis model was accelerated compared to wild type mice[22]. In addition to these functional indications that VTCN1 plays a role in RA pathology, genetic studies in juvenile idiopathic arthritis identified this region to associate with this disease[23]. Since both functional and genetic data indicated a possible role of VTCN1 in RA, the genetic contribution to RA susceptibility was investigated by a candidate gene approach in our Leiden RA population.

In the discovery phase of this study, a significant association for two SNPs was observed. Replication of these findings in two independent populations from northern-European descent, showed an overall significant association for both variants. These data indicate that this is a novel genetic risk factor for RA susceptibility, although further replication in independent cohorts is necessary to tease apart the role of VTCN1 genetics in RA.

In chapter 7 another novel genetic risk factor is identified by candidate gene approach. In this study the genetic contribution of the C1Q genes was investigated in relation to RA susceptibility. As well as in chapter 5, this region was chosen upon strong indications of both functional and genetic data. It was observed that a complete genetic deficiency of C1Q is strongly associated with the development of Systemic Lupus Erythematosus[24]. Moreover, complement deposits can be found in the synovium of RA patients and a correlation between disease activity and the presence of activated complement fragments bound to C1q in sera of RA patients has been described[25, 26].

In the Leiden RA population five SNPs in the C1Q genes were significantly associated with disease. Subsequent replication of the most significant SNP in three independent populations supported the initial finding. Additionally, in a meta-analysis of six GWAS, containing a substantial number of patients and controls, the association was maintained, even though with borderline significance.

Disease subsets

The identification of additional variants, being rare variants, structural variants, gene-gene interactions or common variants, will probably lead to the identification of a substantial part of the genetic risk conferred to RA. It can be expected however, that part of the genetic background of RA, as described currently, will not be elucidated. RA is a phenotypic disease and is defined by use of classification criteria. In these criteria, points are scored for certain disease features and an individual is classified as having RA when a minimum score is obtained[27]. This results in a heterogeneous disease, with highly variable disease outcomes between patients. It can be expected that the current classification embraces several disease subgroups, each with their own distinct risk factors and pathophysiology. This implies that a more precise definition is needed and further research towards specific immunological or other biologic features that distinct subgroups is desired.

To date, a feature that has been shown to identify specific subgroups of the disease is ACPA. It is known that, to a certain extent, presence or absence of ACPAs can predict disease outcome and these antibodies are one of the best clinical predictors of radiological progression[28]. Moreover, it has been shown that patients with ACPA-negative disease are more likely to achieve drug-free remission than those who test positive for these autoantibodies[29]. Interestingly, the most prominent genetic risk factor for RA, the HLA shared epitope region, predisposes to ACPA positive RA, whereas another haplotype in this region, *HLA-DRB1*03*, predisposes to ACPA-negative disease[30, 31]. The majority of loci that correlate with RA have, however, been identified in ACPA-positive patient populations, and little is known about the genetic contribution to the ACPA-negative subset[32]. Small

studies have suggested a role for *IRF5* and genes encoding C-type lectin-like receptors in ACPA-negative disease, and, in chapter 2 of this thesis, it was shown that a polymorphism of *STAT4* is associated with both disease subsets, whereas a variant of *CTLA4* is associated with ACPA-positive RA only[33, 34]. Together, these studies reveal that ACPA-positive and ACPA-negative RA have different genetic association patterns.

To gain further insight into these association patterns, a genome-wide association study in ACPA-positive and ACPA-negative patient populations was performed recently[35]. More than 1.7 million SNPs were studied for association with disease, in 774 ACPA-negative and 1,147 ACPA-positive patients with RA. The ACPA-positive findings were subsequently replicated in two RA cohorts of Western European descent. When the ACPA-positive subset was compared with the ACPA-negative subset, genome-wide significant differences between the two groups (P <2.9 × 10–8) could be established for several SNPs within the HLA region. No significant associations of loci located outside the HLA region could be identified, which is likely to be due to the insufficient power provided by the relatively small ACPA-negative study population. Nonetheless, these data support the idea that genetic backgrounds contribute differently to the two RA disease subsets characterized by ACPA status. Expansion of the ACPA-negative study population is needed to elucidate the genetic contribution of loci outside the HLA region.

The data of Padyukov *et al.*[35] further suggest, in combination with the results of previous studies, that distinct risk factors operate in the two ACPA subsets. This divergence implies that different pathophysiology underlies ACPA-positive and ACPA-negative RA, and therefore, that these subsets should be considered as two separate diseases, and studied separately in both genetic and functional studies of RA pathophysiology. Previous functional studies have shown that in ACPA-positive RA, immunological responses occur in a citrulline-specific manner, and, in mouse models, it has been shown that citrulline-specific antibodies can induce and promote arthritis[36, 37]. Furthermore, activation of basophils from ACPA-positive patients with RA, in contrast to those from ACPA-negative patients, occurs upon exposure to citrullinated antigens[36]. These findings thus indicate a differential response of immune cells to citrullinated antigens.

Diseases with a distinct pathogenesis might, logically, benefit from different treatment strategies. In RA the mainstays of treatment are DMARDs, which are a heterogeneous collection of therapeutic agents for which mechanisms of action are, largely, not well understood. Methotrexate is the most prominent DMARD and is widely used for the treatment of RA and other inflammatory diseases. ACPA-positive patients with undifferentiated arthritis treated with methotrexate are less likely to progress to RA, and do so at a later time point than a placebo control group[38]. By contrast, no effect of methotrexate therapy on progression to RA could be observed in an ACPA negative patient population, indicating that the two ACPA subgroups respond differently to methotrexate treatment[38].

It is not just methotrexate therapy that produces different outcomes in the two disease subsets. In RA refractory to therapy with tumor necrosis factor blockers, rituximab—a

monoclonal antibody directed towards the B cell marker CD20—has proven to be an effective therapy. Upon binding of rituximab, circulating B cell populations are depleted for periods of at least 3 months. Recently, in a clinical study of rituximab in 208 patients with refractory RA, it was shown that the presence of ACPA predicted a better EULAR (European League Against Rheumatism) response at 24 weeks, indicating that this drug might have a greater role in the ACPA-positive subset than in ACPA-negative patients[39].

Thus, although little is known about the genetic contribution to the development of ACPA-negative RA, data suggests contrasting genetic backgrounds for the disease subsets characterized by ACPA status. This genetic divergence lends further support that distinct genetic risk factors play a part in specific subsets of the disease. To fully identify the genetic risk conferred to RA, genetic risk factors that contribute only to certain disease subgroups should be identified.

In chapter 6, by a candidate gene approach, polymorphisms in the PTGES gene were investigated for association with RA susceptibility and disease characteristics. The PTGES gene is involved in prostaglandin synthesis, a key mediator of inflammation. The prostaglandin production is an important target for non-steroid-anti-inflammatory drugs (NSAIDs), which are key painkillers in the therapy of RA patients[40, 41]. The initial study in a Swedish RA population indicated that no genetic association could be identified in the RA population as a whole. However, upon investigation of disease subgroups it seemed that the association was mainly conferred to female RA patients. Although this finding could not be replicated independently in our Leiden RA population, the minor allele frequencies were biased in the same direction. Further replication of this finding is needed to tease apart the exact role of the PTGES gene in female RA patients, however, it does indicate that genetic risk factors might only contribute to a subset of the disease and further investigation into these subsets is desirable.

Linking genotype to biological pathways

The identification of variants that make up the genetic risk conferred to RA is thought to provide valuable insight into the pathophysiology that underlies the disease. The next step is to translate the genetic associations into biological pathways and use this knowledge for the invention of curative therapies for RA.

As more auto-immune susceptibility loci are being identified, it becomes clear that part of these genes overlap between auto-immune diseases. Family based epidemiology studies have suggested there is a shared genetic basis underlying auto-immune diseases in general, which can explain this overlap[42]. The great number of overlapping genes, however, might partly be due to the way genetic association studies are currently performed. In both GWAS and candidate gene studies, genetic risk factors are identified by comparing patients to healthy individuals. Although this approach is considered best practice, it does not distinguish between RA and inflammation in general. Therefore, in functional studies to identify the biological relevance of a genetic risk factor this should be taken into account.

To date, over 46 genetic risk factors have been established to associate with RA suscep-

tibility[18]. Interestingly, all identified genes are located within immunological pathways, as was expected before hand. However, all the identified genes are possible candidates for further functional studies and their relevance in disease pathology is not yet elucidated. It should be noted however, that the significantly associated variants are generally named after known immune genes that are located in the region. Sometimes only one gene is located nearby, making it a reasonable assumption this gene is related to the associated variant. Other times, however, the variant might be located in a region harboring several genes. An association can also be due to a genetic feature, like a micro-RNA or a regulatory element, which has not been described yet. Although this is for a big part due to the fact that the actual causative variant has not been identified, studies towards exploring the functional consequence of a variant are difficult to perform. Thereby, to translate the genetic associations into biological pathways, great hurdles will have to be taken.

In the candidate gene studies described in this thesis, an attempt to identify the functional relevance of the associated genetic variant is made. Although mRNA expression levels and serum levels of the associated proteins and genes could be significantly related to the genotype, these studies all identified further associations and no functional pathways were observed, further indicating that the identification of the disease pathology by translating the genetic association into a biological pathway involved in disease remains a difficult task to perform.

Concluding remarks

In the current thesis the genetic contribution of non-HLA genes to RA susceptibility was further investigated and the functional relevance of these loci was explored. The studies described were able to establish several previously identified risk factors in a statistical robust manner. Also novel genetic risk factors that are associated with RA susceptibility could be identified, as well as risk factors that are conferred to specific subgroups of the disease. The next step towards understanding the pathophysiology of RA will be to identify the relevance of these risk factors, which might prove to be of great challenge.

REFERENCES

1. Scott DL, Wolfe F, Huizinga TW Rheumatoid arthritis. Lancet 376; 1094-108.
2. Aho K, Koskenvuo M, Tuominen J, Kaprio J Occurrence of rheumatoid arthritis in a nationwide series of twins. J Rheumatol 13; 899-902.
3. MacGregor AJ, Snieder H, Rigby AS *et al*. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 43; 30-7.
4. Silman AJ, MacGregor AJ, Thomson W *et al*. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol 32; 903-7.
5. Gregersen PK, Moriuchi T, Karr RW *et al*. Polymorphism of HLA-DR beta chains in DR4, -7, and -9 haplotypes: implications for the mechanisms of allelic variation. Proc Natl Acad Sci U S A 83; 9149-53.

6. Gregersen PK, Shen M, Song QL *et al.* Molecular diversity of HLA-DR4 haplotypes. Proc Natl Acad Sci U S A 83; 2642-6.

7. Feitsma AL, van der Voort EI, Franken KL *et al.* Identification of citrullinated vimentin peptides as T cell epitopes in HLA-DR4-positive patients with rheumatoid arthritis. Arthritis Rheum 62; 117-25.

8. Suzuki A, Yamada R, Chang X *et al.* Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 34; 395-402.

9. Stahl EA, Raychaudhuri S, Remmers EF *et al.* Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42; 508-14.

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

11. Kobayashi S, Ikari K, Kaneko H *et al.* Association of STAT4 with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in the Japanese population. Arthritis Rheum 58; 1940-6.

12. Remmers EF, Plenge RM, Lee AT *et al.* STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 357; 977-86.

13. Barton A, Jury F, Eyre S *et al.* Haplotype analysis in simplex families and novel analytic approaches in a case-control cohort reveal no evidence of association of the CTLA-4 gene with rheumatoid arthritis. Arthritis Rheum 50; 748-52.

14. Orozco G, Torres B, Nunez-Roldan A, Gonzalez-Escribano MF, Martin J Cytotoxic T-lymphocyte antigen-4-CT60 polymorphism in rheumatoid arthritis. Tissue Antigens 64; 667-70.

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

16. Zhernakova A, Alizadeh BZ, Bevova M *et al.* Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. Am J Hum Genet 81; 1284-8.

17. Barton A, Thomson W, Ke X *et al.* Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. Nat Genet 40; 1156-9.

18. Gregersen PK Susceptibility genes for rheumatoid arthritis - a rapidly expanding harvest. Bull NYU Hosp Jt Dis 68; 179-82.

19. Manolio TA, Collins FS, Cox NJ *et al.* Finding the missing heritability of complex diseases. Nature 461; 747-53.

20. Sica GL, Choi IH, Zhu G *et al*. B7-H4, a molecule of the B7 family, negatively regulates T cell immunity. Immunity 18; 849-61.

21. Yi KH, Chen L Fine tuning the immune response through B7-H3 and B7-H4. Immunol Rev 229; 145-51.

22. Azuma T, Zhu G, Xu H *et al*. Potential role of decoy B7-H4 in the pathogenesis of rheumatoid arthritis: a mouse model informed by clinical data. PLoS Med 6; e1000166.

23. Hinks A, Barton A, Shephard N *et al*. Identification of a novel susceptibility locus for juvenile idiopathic arthritis by genome-wide association analysis. Arthritis Rheum 60; 258-63.

24. Botto M, Walport MJ C1q, autoimmunity and apoptosis. Immunobiology 205; 395-406.

25. Konttinen YT, Ceponis A, Meri S *et al*. Complement in acute and chronic arthritides: assessment of C3c, C9, and protectin (CD59) in synovial membrane. Ann Rheum Dis 55; 888-94.

26. Wouters D, Voskuyl AE, Molenaar ET, Dijkmans BA, Hack CE Evaluation of classical complement pathway activation in rheumatoid arthritis: measurement of C1q-C4 complexes as novel activation products. Arthritis Rheum 54; 1143-50.

27. Aletaha D, Neogi T, Silman AJ *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 69; 1580-8.

28. Syversen SW, Gaarder PI, Goll GL *et al*. High anti-cyclic citrullinated peptide levels and an algorithm of four variables predict radiographic progression in patients with rheumatoid arthritis: results from a 10-year longitudinal study. Ann Rheum Dis 67; 212-7.

29. van der Woude D, Young A, Jayakumar K *et al*. Prevalence of and predictive factors for sustained disease-modifying antirheumatic drug-free remission in rheumatoid arthritis: results from two large early arthritis cohorts. Arthritis Rheum 60; 2262-71.

30. Huizinga TW, Amos CI, van der Helm-van Mil AH *et al*. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 52; 3433-8.

31. Verpoort KN, van Gaalen FA, van der Helm-van Mil AH *et al*. Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis. Arthritis Rheum 52; 3058-62.

32. Stahl EA, Raychaudhuri S, Remmers EF *et al*. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42; 508-14.

33. Lorentzen JC, Flornes L, Eklow C *et al*. Association of arthritis with a gene complex encoding C-type lectin-like receptors. Arthritis Rheum 56; 2620-32.

34. Sigurdsson S, Padyukov L, Kurreeman FA *et al*. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. Arthritis Rheum 56; 2202-10.

35. Padyukov L, Seielstad M, Ong RT *et al*. A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis. Ann Rheum Dis 70; 259-65.

36. Schuerwegh AJ, Ioan-Facsinay A, Dorjee AL *et al*. Evidence for a functional role of IgE anticitrullinated protein antibodies in rheumatoid arthritis. Proc Natl Acad Sci U S A 107; 2586-91.

37. Uysal H, Bockermann R, Nandakumar KS *et al*. Structure and pathogenicity of antibodies specific for citrullinated collagen type II in experimental arthritis. J Exp Med 206; 449-62.

38. van DH, van AJ, Lard LR *et al*. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. Arthritis Rheum 56; 1424-32.

39. Sellam J, Hendel-Chavez H, Rouanet S *et al*. B-cell activation biomarkers as predictive factors of the response to rituximab in rheumatoid arthritis. Arthritis Rheum.

40. Halushka MK, Halushka PV Toward individualized analgesic therapy: functional cyclooxygenase 1 and 2 haplotypes. Clin Pharmacol Ther 79; 404-6.

41. Jakobsson PJ, Thoren S, Morgenstern R, Samuelsson B Identification of human prostaglandin E synthase: a microsomal, glutathione-dependent, inducible enzyme, constituting a potential novel drug target. Proc Natl Acad Sci U S A 96; 7220-5.

42. Hemminki K, Li X, Sundquist K, Sundquist J Shared familial aggregation of susceptibility to autoimmune diseases. Arthritis Rheum 60; 2845-7.

43. Kurreeman FA, Padyukov L, Marques RB *et al*. A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. PLoS Med 4; e278.