Human Molecular Genetics, 2009, Vol. 18, No. 21 ***4195–4203***
doi:10.1093/hmg/ddp365
Advance Access published on July 31, 2009

Common and different genetic background for rheumatoid arthritis and coeliac disease

Marieke J.H. Coenen[1,†], Gosia Trynka[3,†], Sandra Heskamp[1], Barbara Franke[1], Cleo C. van Diemen[3], Joanna Smolonska[3], Miek van Leeuwen[4], Elisabeth Brouwer[4], Marike H. Boezen[5], Dirkje S. Postma[6], Mathieu Platteel[3], Pieter Zanen[7], Jan-Willem W.J. Lammers[7], Harry J.M. Groen[6], Willem P.T.M. Mali[8], Chris J. Mulder[11], Greetje J. Tack[11], Wieke H.M. Verbeek[11], Victorien M. Wolters[9], Roderick H.J. Houwen[9], M. Luisa Mearin[12], David A. van Heel[13], Timothy R.D.J. Radstake[2], Piet L.C.M. van Riel[2], Cisca Wijmenga[3], Pilar Barrera[2] and Alexandra Zhernakova[10,*]

[1]Department of Human Genetics, Institute for Genetic and Metabolic Diseases and [2]Department of Rheumatology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, [3]Genetics Department, [4]Department of Rheumatology, [5]Department of Epidemiology and [6]Department of Pulmonology, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands, [7]Department of Respiratory Diseases, Division of Heart and Lungs, [8]Department of Radiology, Radiotherapy and Nuclear Medicine, [9]Department of Paediatric Gastroenterology and [10]Department of Medical Genetics, University Medical Centre Utrecht, Utrecht, The Netherlands, [11]Department of Gastroenterology, VU Medical Centre, Amsterdam, The Netherlands, [12]Department of Paediatric Gastroenterology, Leiden University Medical Centre, Leiden, The Netherlands and [13]Institute of Cell and Molecular Science, Barts and The London School of Medicine and Dentistry, London, UK

Received May 22, 2009; Revised and Accepted July 29, 2009

Recent genome-wide association studies (GWAS) have revealed genetic risk factors in autoimmune and inflammatory disorders. Several of the associated genes and underlying pathways are shared by various autoimmune diseases. Rheumatoid arthritis (RA) and coeliac disease (CD) are two autoimmune disorders which have commonalities in their pathogenesis. We aimed to replicate known RA loci in a Dutch RA population, and to investigate whether the effect of known RA and CD risk factors generalize across the two diseases. We selected all loci associated to either RA or CD in a GWAS and confirmed in an independent cohort, with a combined *P*-value cut-off $P < 5 \times 10^{-6}$. We genotyped 11 RA and 11 CD loci in 1368 RA patients, 795 CD patients and 1683 Dutch controls. We combined our results in a meta-analysis with UK GWAS on RA (1860 cases; 2938 controls) and CD (767 cases; 1422 controls). In the Dutch RA cohort, the *PTPN22* and *IL2/IL21* variants showed convincing association ($P = 3.4 \times 10^{-12}$ and $P = 2.8 \times 10^{-4}$, respectively). Association of RA with the known CD risk variant in the *SH2B3* was also observed, predominantly in the subgroup of rheumatoid factor-positive RA patients ($P = 0.0055$). In a meta-analysis of Dutch and UK data sets, shared association with six loci (*TNFAIP3*, *IL2/IL21*, *SH2B3*, *LPP*, *MMEL1/TNFRSF14* and *PFKFB3/PRKCQ*) was observed in both RA and CD cohorts. We confirmed two known loci and identified four novel ones for shared CD–RA genetic risk. Most of the shared loci further emphasize a role for adaptive and innate immunity in these diseases.

*To whom correspondence should be addressed at: Department of Biomedical Genetics, University Medical Centre Utrecht, Stratenum 2.112, PO Box 85060, 3508 AB Utrecht, The Netherlands. Tel: +31 887568790; Fax: +31 887568479; Email: a.zhernakova@umcutrecht.nl
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic inflammatory disorder affecting ~1% of the adult population. The disease is characterized by inflammation of the synovial tissue of multiple joints leading to pain, deformities and a reduced quality of life. The aetiology of RA is complex and largely unknown; however, epidemiologic data support a polygenic susceptibility model (1). Besides this, environmental factors may also play a role in the pathogenesis of RA (2).

RA shares common pathogenic mechanisms with other autoimmune disorders. This is reflected by the co-occurrence of several autoimmune disorders in some patients and families, and the shared genetic background of autoimmune diseases (3,4). One of the autoimmune disorders showing similar pathogenic mechanisms to those seen in RA is coeliac disease (CD). This is a chronic intestinal inflammatory disorder which develops through an intolerance to gluten—the major dietary protein in wheat—, and related proteins from barley and rye (5). Although the diseases are clearly distinct in their phenotype, common features in RA and CD include the association of the HLA molecules with the diseases, T-cell infiltration in the target organs, the development of disease-specific auto-antibodies and the role of enzymes involved in post-translational modifications in the pathogenesis of the diseases (6). In addition, several studies show co-occurrence of CD and RA (7–10). However, there are no well-designed studies to assess the true prevalence of RA and CD co-morbidity.

Recently performed genome-wide association studies (GWAS) have identified 11 loci associated with RA and 10 loci associated with CD, in addition to the already known HLA loci (11–21) (Supplementary Material, Table S1). These GWAS showed that, besides HLA, two chromosomal regions are shared between RA and CD: a region on chromosome 4q27 (including the genes *IL2* and *IL21)* and the 6q23.3 locus containing the *OLIG3* and *TNFAIP3* genes (22). Interestingly, the *IL2/IL21* locus has also been associated with psoriatic arthritis, Grave's disease and type 1 diabetes, whereas the *OLIG3/TNFAIP3* locus has also been associated to systemic lupus erythematosus and type 1 diabetes (13,23–27). These results strongly suggest that overlapping genetic mechanisms underlie the development of multiple autoimmune disorders (4). We therefore hypothesize that other susceptibility genes identified for RA might also contribute to the development of CD and vice versa. In this study, we investigated a total of 22 SNPs associated with either RA (11 SNPs) or CD (11 SNPs) by association testing in Dutch RA and CD cohorts and by meta-analyses including well-comparable data sets from two earlier GWAS on RA and CD.

RESULTS

In this study, we performed an association analysis with all SNPs showing replicated association to either RA or CD with a *P*-value cut-off of 5×10^{-6}. In total, 22 SNPs (11 primary CD-related and 11 primary RA-related SNPs) were genotyped or imputed in 1368 Dutch RA cases, 795 Dutch CD cases and 1683 controls. Since the distribution of males and females was substantially different in our cases and control cohort (frequency female in both case cohorts 60%, whereas only 20% of controls were female, see Supplementary Material, Table S2), we first compared the allele frequency of all the SNPs in males and females in our control cohort. None of the SNPs showed a significant difference (Supplementary Material, Table S4), which we took as evidence for the comparability of the samples.

Replication of RA loci in Dutch RA cohort

A replication of known RA variants in a Dutch RA cohort has not been performed previously. From a total of 11 RA loci, only the well-established risk variant in *PTPN22* (rs2476601) showed a clear association with RA in our cohort ($P = 3.43 \times 10^{-12}$; OR = 1.70 (95% confidence interval (CI): 1.46–1.98)). Three other RA SNPs (rs3890745, rs4810485 and rs3218253 located in the *MMEL1/TNFRSF14*, *CD40* and the *IL2RB* gene regions, respectively) were nominally replicated in Dutch RA samples ($P < 0.05$). The remaining RA loci were not associated in our RA data set (Table 1).

Association of CD loci in Dutch CD cohort

All previously confirmed coeliac variants, except the *LPP* SNP rs1464510, showed association to CD in the Dutch cohort. This is not surprising since we used 508 of the 795 cases, and 833 of the 1683 controls as part of the multinational cohorts in our previous studies to establish the association of the CD loci (14,20,21).

Cross-disease association study in the Dutch cohorts

Three out of 11 primary CD-related SNPs were associated with RA in the Dutch cohort. The *IL2/IL21* variant rs13151961 showed the strongest association to RA [$P = 0.0003$; OR = 0.78 (95% CI: 0.68–0.89)]. This association was reported previously in a subset of our cohort (22). We now confirmed the association of *IL2/21* variants in an extended group of RA and CD cases. The other two variants, *SH2B3* rs3184504 and *LPP* rs1464510, showed a moderate association with RA [$P = 0.024$; OR = 1.12 (95% CI: 1.02–1.25) and $P = 0.012$; OR = 1.14 (95% CI: 1.03–1.26), respectively]. Opposing allelic association for RA and CD was observed for the LPP variant: the frequency of LPP rs1464510*C allele was increased in RA compared with controls, whereas the rs1464510*A allele was more frequent in CD patients than controls.

From the RA-specific variants, only rs3890745 from the *MMEL1/TNFRSF14* locus showed a trend for association to CD [$P = 0.04$, OR = 0.87 (95% CI: 0.77–0.99)]. The rs10499194 variant in the *TNFAIP3* locus, previously reported only in RA samples, also showed moderate association to CD [$P = 0.018$; OR = 0.84 (95% CI: 0.73–0.97)] (Table 1).

Stratification of RA samples for rheumatoid factor

As the association of several known RA genetic risk variants has been shown to be different in auto-antibody positive and negative cases (28), we performed a separate association

Table 1. Results of association analysis in Dutch RA and CD cohorts

Gene(s) present in risk locus	Reported association	rsID	CHR	Minor allele	MAF controls	RA MAF	*P*-value	OR	95% CI	CD MAF	*P*-value	OR	95% CI
RGS1	CD	rs2816316	1	C	0.19	0.18	0.797	0.98	0.86–1.12	0.14	7.37E−05	0.72	0.61–0.84
REL	CD	rs842647	2	G	0.36	0.34	0.106	0.92	0.82–1.02	0.29	1.38E−05	0.75	0.66–0.86
IL18RAP	CD	rs917997	2	A	0.23	0.24	0.628	1.03	0.91–1.17	0.29	1.73E−05	1.34	1.17–1.54
CCR3	CD	rs6441961	3	A	0.32	0.31	0.855	0.99	0.89–1.10	0.38	1.52E−05	1.32	1.16–1.49
IL12A/SCHIP	CD	rs17810546	3	G	0.12	0.11	0.613	0.96	0.82–1.13	0.17	2.73E−07	1.55	1.31–1.84
IL12A	CD	rs9811792	3	G	0.45	0.44	0.566	0.97	0.88–1.08	0.50	0.002	1.21	1.08–1.37
LPP	CD	rs1464510	3	C	0.49	0.53	0.013	1.14	1.03–1.26	0.47	0.102	0.91	0.80–1.02
IL2-21	CD	rs13151961	4	G	0.18	0.15	0.0003	0.78	0.68–0.89	0.13	5.08E−07	0.65	0.54–0.77
TAGAP	CD	rs1738074	6	A	0.40	0.39	0.579	0.97	0.88–1.08	0.44	0.004	1.19	1.06–1.34
SH2B3	CD	rs3184504	12	A	0.46	0.49	0.024	1.12	1.02–1.25	0.52	1.15E−05	1.31	1.16–1.47
OLIG3-TNFAIP3	CD/RA	rs2327832	6	G	0.21	0.22	0.230	1.08	0.95–1.22	0.23	0.0611	1.15	0.99–1.32
OLIG3-TNFAIP3	RA	rs10499194	6	T	0.24	0.24	0.987	1.00	0.89–1.13	0.21	0.018	0.84	0.73–0.97
MMEL1/TNFRSF14	RA	rs3890745	1	G	0.32	0.29	0.014	0.87	0.78–0.97	0.29	0.041	0.87	0.77–0.99
PTPN22	RA	rs2476601	1	A	0.10	0.16	3.43E−12	1.70	1.46–1.98	0.10	0.604	0.95	0.78–1.16
STAT4	RA	rs7574865	2	A	0.24	0.24	0.930	1.01	0.89–1.13	0.23	0.759	0.98	0.85–1.13
CDK6	RA	rs42041[a]	7	G	0.25	0.26	0.462	1.06	0.91–1.22	0.26	0.704	1.03	0.88–1.22
CCL21	RA	rs2812378	9	G	0.35	0.37	0.167	1.08	0.97–1.20	0.36	0.453	1.05	0.93–1.19
TRAF1/C5	RA	rs3761847	9	G	0.44	0.45	0.829	1.01	0.91–1.12	0.44	0.860	0.99	0.88–1.12
PFKFB3/PRKCQ	RA	rs4750316	10	C	0.19	0.21	0.094	1.12	0.98–1.27	0.17	0.080	0.87	0.74–1.02
KIF5A	RA	rs1678542	12	G	0.33	0.34	0.331	1.06	0.95–1.18	0.32	0.670	0.97	0.86–1.11
CD40	RA	rs4810485	20	T	0.24	0.21	0.015	0.86	0.76–0.97	0.25	0.347	1.07	0.93–1.23
IL2RB	RA	rs3218253	22	A	0.27	0.30	0.030	1.13	1.01–1.27	0.29	0.270	1.08	0.94–1.23

MAF, minor allele frequency; CHR, chromosome; OR odds ratio.
[a]rs42041 was genotyped in a subgroup of controls ($n = 833$).

Table 2. Association of RA and CD SNPs in rheumatoid factor-positive subgroup of RA patients ($n = 775$), compared to controls ($n = 1683$)

Gene(s) present in risk locus	CD/RA SNP	rsID	CHR	Minor allele	MAF_cont	MAF RF pos	*P*-value	OR	95% CI
RGS1	CD	rs2816316	1	C	0.19	0.18	0.640	0.96	0.82–1.13
REL	CD	rs842647	2	G	0.36	0.33	0.070	0.89	0.78–1.01
IL18RAP	CD	rs917997	2	A	0.23	0.25	0.234	1.09	0.95–1.26
CCR3	CD	rs6441961	3	A	0.32	0.32	0.824	1.02	0.89–1.16
IL12A/SCHIP	CD	rs17810546	3	G	0.12	0.11	0.693	0.96	0.80–1.17
IL12A	CD	rs9811792	3	G	0.45	0.42	0.066	0.89	0.79–1.01
LPP	CD	rs1464510	3	C	0.49	0.52	0.161	1.09	0.97–1.23
IL2-21	CD	rs13151961	4	G	0.18	0.14	0.0004	0.74	0.62–0.87
TAGAP	CD	rs1738074	6	A	0.40	0.38	0.104	0.90	0.80–1.02
SH2B3	CD	rs3184504	12	A	0.46	0.50	0.0055	1.19	1.05–1.34
OLIG3-TNFAIP3	CD/RA	rs2327832	6	G	0.21	0.21	0.990	1.00	0.86–1.16
OLIG3-TNFAIP3	CD/RA	rs10499194	6	T	0.24	0.25	0.602	1.04	0.90–1.20
MMEL1-TNFRSF14	RA	rs3890745	1	G	0.32	0.29	0.016	0.85	0.75–0.97
PTPN22	RA	rs2476601	1	A	0.10	0.17	1.45E−11	1.81	1.52–2.15
STAT4	RA	rs7574865	2	A	0.24	0.24	0.982	1.00	0.87–1.15
CDK6	RA	rs42041[a]	7	G	0.25	0.25	0.818	0.98	0.83–1.16
CCL21	RA	rs2812378	9	G	0.35	0.37	0.150	1.10	0.97–1.24
TRAF1/C5	RA	rs3761847	9	G	0.44	0.45	0.712	1.02	0.91–1.16
PFKFB3/PRKCQ	RA	rs4750316	10	C	0.19	0.22	0.024	1.19	1.02–1.38
KIF5A[a]	RA	rs1678542	12	G	0.33	0.34	0.322	1.07	0.94–1.21
CD40	RA	rs4810485	20	T	0.24	0.23	0.413	0.94	0.82–1.09
IL2RB	RA	rs3218253	22	A	0.27	0.29	0.245	1.08	0.95–1.24

CHR, chromosome; RF, rheumatoid factor; MAF, minor allele frequency; OR, odds ratio.
[a]rs42041 was genotyped in a subgroup of controls ($n = 833$).

analysis in rheumatoid factor (RF)-positive and RF-negative cases. We had information on RF status available for a subgroup of Dutch RA cases: 776 cases were RF positive and 204 RF negative. In the RF-positive group the *SH2B3* variant rs3184504 showed stronger association compared to the total RA cohort ($P = 0.0055$, OR = 1.19 (95% CI: 1.05–1.34) (Table 2). We did not perform stratification on anti-CCP auto-antibodies, as the anti-CCP status was only available for a minority of cases.

Meta-analysis

To increase the power of the study, we combined our data with those of a GWAS in RA and CD patients from the UK into a meta-analysis. The UK cohorts included 1860 cases and 2938 controls from the WTCCC study on RA, and 767 CD cases and 1422 controls from a UK GWAS in CD (13,21). In the meta-analysis, four risk loci showed association to both RA and CD with $P < 0.01$, including the genes *IL2/IL21*, *LPP*, *TNFAIP3* and *SH2B3* (Table 3). Association with the *LPP* locus in CD and RA was observed for opposing alleles. A trend for association with SNPs in *MMEL1-TNFRSF14* and *PRKCQ* was observed for both diseases ($P < 0.05$) (Table 3). An overview of the common and separate associations of the tested genes with RA and CD is shown in Figure 1.

DISCUSSION

In this study, we (i) performed the replication of known RA associated loci in our Dutch cohort of RA cases and controls; (ii) performed the cross-study of RA and CD associated variants in Dutch RA and CD cohorts and (iii) combined our results in a meta-analysis with the UK GWAS in RA and CD (13,21). We were able to replicate 4 out of 11 RA loci in our Dutch RA cohort, and identified 6 loci which showed shared association to CD and RA in the meta-analysis. Strikingly, most of the previously established RA SNPs were not replicated. There are several reasons for this observation. First, the endophenotyping difference might be the major reason for non-replication: several of the RA loci were established mainly in subgroups of ACPA-positive RA patients. Differential association in auto-antibody positive and negative subgroups of RA has been reported previously (28). We could not stratify for ACPA in our cohort due to lack of ACPA status for most patients. However, when stratified for RF, we were able to observe a stronger effect for the variant in the SH2B3 gene. Secondly, population heterogeneity might be another explanation for lack of association with the initial reported variants. To control for the heterogeneity of the association test, we included the Breslow–Day test in our analysis. Two of the SNPs (rs4750316 and rs1678542, the PFKFB3/PRKCQ and KIF5A locus, respectively) showed significant results for this test, indicating the presence of heterogeneity between the two populations. Further replication in other populations is essential for establishing the true risk effect of these genes. Since we only tested a single marker (the most associated one for each locus), we may have missed the association due to low/other linkage disequilibrium (LD) between the established 'tagging' SNP and the true causal variant in the Dutch population. Finally, the relative risk of associated regions in the initial studies was rather low (OR between 0.75 and 1.32 for most SNPs). Moreover, similar to the single gene association studies, the effect of some of the established loci might be overestimated and our

Table 3. Meta-analysis of SNPs in Dutch and UK RA and CD populations

Locus	CD/RA SNP	rsID	Ref allele	RA pCMH	OR	95% CI	CD pCMH	OR	95% CI
RGS1	CD	rs2816316	C	0.6349	0.98	0.90–1.06	1.76E−07	0.73	0.65–0.82
REL	CD	rs842647	G	0.1470	0.95	0.89–1.02	2.87E−06	0.80	0.73–0.88
IL18RAP	CD	rs917997	A	0.1071	1.06	0.99–1.15	8.62E−09	1.34	1.21–1.48
CCR3	CD	rs6441961	A	0.2672	1.04	0.97–1.11	2.46E−07	1.27	1.16–1.39
IL12A/SCHIP	CD	rs17810546	G	0.4926	0.97	0.88–1.07	1.45E−09	1.47	1.30–1.67
IL12A	CD	rs9811792	G	0.9945	1.00	0.94–1.07	1.16E−05	1.21	1.11–1.32
LPP	CD	rs1464510	C	0.0054	1.10	1.03–1.17	0.00010	0.84	0.77–0.92
IL2-21	CD	rs13151961	G	1E−05	0.83	0.76–0.90	6.61E−12	0.65	0.58–0.74
TAGAP	CD	rs1738074	A	0.0839	0.94	0.88–1.01	1.81E−05	1.21	1.11–1.32
SH2B3	CD	rs3184504	A	0.0011	1.11	1.04–1.19	4.42E−08	1.27	1.17–1.39
OLIG3-TNFAIP3	CD/RA	rs2327832	G	4E−05	1.17	1.09–1.27	0.00069	1.19	1.08–1.32
OLIG3-TNFAIP3	CD/RA	rs10499194	T	0.0419	0.93	0.86–1.00	0.0419	0.93	0.86–1.00
MMEL1, TNFRSF14	RA	rs3890745	G	5E−07	0.84	0.78–0.90	0.0275	0.90	0.82–0.99
PTPN22	RA	rs2476601	A	2E−27	1.67	1.52–1.84	0.2623	1.08	0.94–1.25
STAT4	RA	rs7574865	A	0.1041	1.06	0.99–1.15	0.3319	1.05	0.95–1.16
CDK6	RA	rs42041[a]	G	0.0085	1.11	1.03–1.20	n/a	n/a	n/a
CCL21	RA	rs2812378	G	0.0007	1.12	1.05–1.20	0.3189	1.05	0.96–1.15
TRAF1, C5	RA	rs3761847	G	0.9508	1.00	0.94–1.06	0.6610	0.98	0.90–1.07
PFKFB3, PRKCQ	RA	rs4750316	C	0.0390	0.92	0.85–1.00	0.0135	0.87	0.77–0.97
KIF5A	RA	rs1678542	G	0.0076	0.91	0.85–0.98	0.8002	0.99	0.90–1.08
CD40	RA	rs4810485	T	0.0033	0.89	0.83–0.96	0.2007	1.07	0.97–1.18
IL2RB	RA	rs3218253	A	2E−05	1.17	1.09–1.26	0.0568	1.10	1.00–1.21

pCMH, *P*-value Mantel–Haenszel chi-square.
[a]The meta-analysis for rs42041 was not done due to poor imputation of this SNP in the UK and Dutch GWAS data sets.

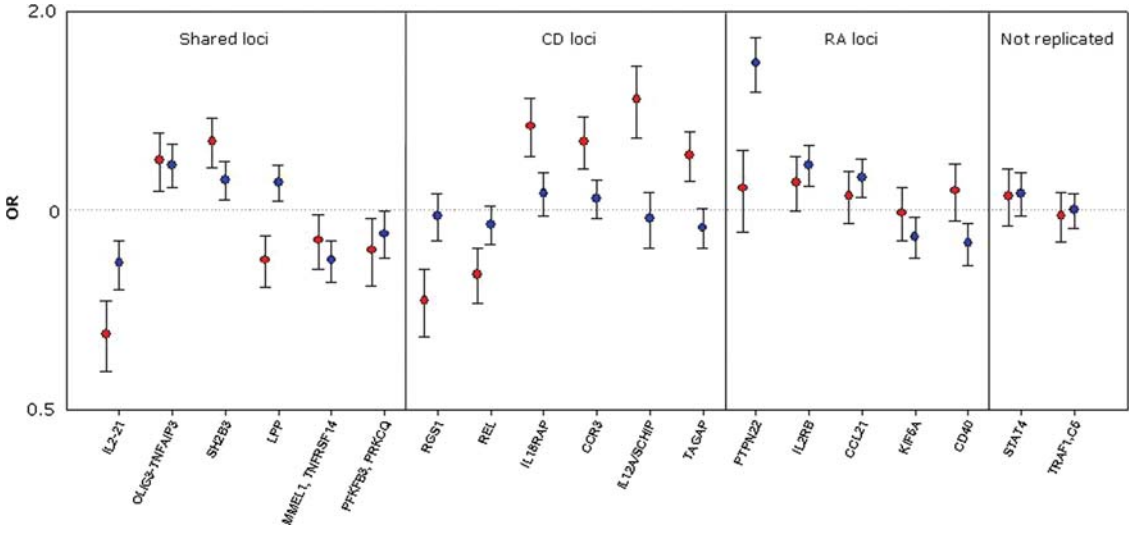

Figure 1. Association of 20 investigated loci to CD and RA. From the IL12A-SHIP locus and OLIG3-TNFAIP3 locus, one of two SNPs (the most associated one, rs17810546 and rs2327832, respectively) are included in the figure. The SNP rs42041 was excluded from the meta-analysis due to poor imputation quality. Blue dots and corresponding line—OR and 95% CI for RA. Red dots and lines—OR and 95% CI for CD. OR was calculated for the minor allele of associated SNP.

study might not have had enough power to replicate all loci. Depending on the SNP, the power of our study ranged from 0.2–1 in the Dutch cohorts and increased in the meta-analysis to a range of 0.34–1 (Supplementary Material, Table S5). The power was at 80% or above for 16 out of 22 SNPs for the RA NL-UK combined analysis and above 80% for 8 out of 22 SNPs for the CD NL-UK analysis. All these factors, or combinations of them, might explain the negative results for most of the genes we tested. Further replication studies in various populations and endophenotypes would shed light on the effect and heterogeneity of associated loci in RA.

In the combined Dutch and UK data sets, we observed a convincing shared association of four genes to RA and CD ($P < 0.01$ in both diseases), and suggestive association of two more loci ($P < 0.05$ in both diseases). Two of the shared loci, *IL2/IL21* and *TNFAIP3*, are already known to be involved in both diseases and were confirmed in our study.

Both genes show association with several other immune-related diseases: *IL2/IL21* with inflammatory bowel disease (29–31), type 1 diabetes (22,24) and psoriasis (23), and *TNFAIP3* with systemic lupus erythematosus (26), type 1 diabetes (25) and psoriasis (32). This points to a general role for these genes in the development of autoimmunity.

The association of IL2-IL21 locus to several immune-related diseases is especially interesting, although strong LD in this locus makes it difficult to locate the true associated gene. Both IL2 and IL21 are important in T-regulatory and Th17 cells, respectively. Both CD and RA show high levels of Th17 in affected tissues (33,34). Interestingly, IL21 is produced by Th17 cells and is important for maintaining Th17 cells (35). High levels of IL21 have been found in the intestine of patients with CD (21), whereas IL21 receptor is overexpressed in synovial tissues of RA patients (36). Blockading the IL21/IL21R pathway ameliorates disease in a murine model of RA (37).

The association of SH2B3 with both RA and CD has not been reported previously. The SH2B3 rs3184504 variant is a non-synonymous SNP R262W located in exon 3 of the gene, and has previously been associated with CD, type 1 diabetes and myocardial infarction (14,24,38). It encodes the T-cell adapter protein LNK, which regulates T-cell receptor-, growth factor- and cytokine receptor-mediated signalling, and is therefore an attractive candidate gene for shared autoimmune susceptibility (39).

Another shared gene identified in this study, *LPP* (LIM domain containing preferred translocation partner in lipoma), is involved in cell adhesion, cytoskeletal remodelling and maintaining cell shape and motility (40,41). Chromosomal aberrations including the *LPP* region have been observed in leukaemia, indicating a potential role for this chromosomal region in regulating the immune system (42). The exact function of *LPP* in autoimmunity has not yet been defined. Interestingly, in our study, *LPP* shows a differential association for RA and CD. The effect of *LPP* rs1464510*A allele confers protection in RA, and susceptibility in CD. The reason for the opposing allelic associations could be the presence of distinct RA- and CD-causing variants, located on different haplotypes and tagged by the opposite alleles of the same SNP. Another possibility is that the same variant confers truly susceptibility to one disease and protection from another, similar to the *PTPN22* functional variant Arg620Trp, which confers susceptibility to several autoimmune diseases, but protection from Crohn's disease (43,44). Sequencing of the whole associated block in both diseases is required to define the RA and CD causal variants and explain the exact nature of this observation. Association of opposite alleles to different autoimmunities has been also observed for *IL18RAP* and *TAGAP* variants in CD and type 1 diabetes (45).

Two loci showed moderate association to both diseases—the *PFKFB3/PRKCQ* and *MMEL1-TNFRSF14* ($P < 0.05$). The *PFKFB3/PRKCQ* variant was originally reported in a meta-analysis including RA patients (17). In addition, the same locus has recently been associated with type 1 diabetes in a meta-analysis (46). Thus, our findings, although only moderately significant, provide additional support for this locus being a shared autoimmune gene. PRKCQ is involved in regulating and controlling T-cell-mediated signalling and is therefore a plausible candidate for autoimmune traits (47) The *MMEL1/TNFRSF14* locus includes the *TNFRSF14* (*HVEM*, herpes virus entry mediator) gene, which functions as a co-stimulatory molecule during T-cell activation (48), and enhances the bactericidal activities of human monocytes and neutrophils (49). In antiviral responses, TNFRSF14 is involved in NF-κB activation.

Overall, from six shared CD-RA loci, five contain genes directly involved in immune function. Association to the *SH2B3*, *PRKCQ*, *IL2/21* and *TNFRSF14* genes points to the role of T-cell-mediated signalling, whereas both the *TNFAIP3* and *TNFRSF14* genes are linked to NF-kb signalling, innate immunity and the response to pathogens. The association of innate molecules to RA and CD may now explain the link of mucosal immunity state and infections in predisposition to both diseases (50–52).

Our study extends the knowledge on genes that are shared between RA and CD. Most of the shared genes fit with the current hypotheses on the pathogenesis of these diseases, involving both innate and adaptive immunity. Several of the shared genes also contribute to the susceptibility to additional autoimmune diseases. Creating genetic profiles for autoimmunity may help understand the basic mechanisms of pathogenesis and predict common immune-related risk factors/phenotypes. Furthermore, it will open up possibilities for the development of new drug targets for a better treatment of autoimmune diseases.

MATERIALS AND METHODS

Study population

Rheumatoid arthritis cohorts. We combined two independent Dutch RA inception cohorts: from Nijmegen ($n = 960$) and Groningen ($n = 408$). Both cohorts have been described elsewhere (22,53,54). All patients were diagnosed according to the American College of Rheumatology (ACR) criteria for RA (55). Due to the lack of whole genome data, we ascertained subjects were of Dutch descent based on their surname.

Coeliac disease cohorts. Our study analysed 795 unrelated Dutch individuals with CD. All affected individuals were diagnosed according to the revised ESPGAN criteria (56). The cohort encompassed individuals that showed a Marsh II or Marsh III lesion in the initial diagnostic small-bowel biopsy specimens upon re-evaluation by one of two experienced pathologists, or presented with dermatitis herpetiformis and were HLA-DQ2 positive.

Control cohort. The control cohort comprised unrelated blood bank donors ($n = 833$) and NELSON controls ($n = 850$). The blood bank control cohort was described earlier (14,20). Other controls were included from the NELSON project—an ongoing population-based, randomized multi-centre lung cancer screening trial, studying male smokers (57). These controls were collected from the north and centre of the Netherlands (Groningen, Utrecht and Drenthe, The Netherlands). All the control subjects were heavy smokers or ex-smokers (a minimum of 16 cigarettes/day for 25 years or 11 cigarettes/day

for 30 years), but did not develop airway obstruction or emphysema suggesting chronic obstructive pulmonary disease (COPD) until the end of a 4 year observation period.

The current study was approved by the local ethics committees and all the patients and controls gave their written informed consent. Information on male/female ratio of the cohorts is shown in Supplementary Material, Table S2.

SNP selection and genotyping

We selected SNPs with confirmed association to either RA or CD with $P < 5 \times 10^{-06}$ from existing literature. Information on the original studies and the associated SNPs is given in Supplementary Material, Table S1. In total, we tested 22 SNPs from 20 loci, 11 known for association to CD and 11 primarily associated to RA, with two of the genes known to be associated to diseases.

Genotyping of Dutch RA samples. All 22 SNPs were genotyped in the Dutch RA cohort using TaqMan probes and primers developed by Applied Biosystems, on an ABI 7900HT system (Applied Biosystems, Nieuwerkerk a/d IJssel, The Netherlands). Genotyping was performed following the manufacturer's specifications. DNA samples were processed in 384-well plates, each plate contained 8 negative controls and 5 duplicated samples. All duplicates showed consistent results for all SNPs.

Genotyping of Dutch CD samples and controls. Of the 22 SNPs, 15 were present on the Illumina HAP550 platform. The genotyping data of CD cases and controls for these 15 SNPs were extracted from an ongoing GWAS in the Dutch population, performed on the Human670-QuadCustom Illumina BeadChips, which contains the HumanHap550 SNP set (manuscript in preparation). The remaining seven SNPs were not present on the genotyping platforms. In these cases, a 4 MB window around the SNP of interest was imputed using Plink v1.05 (http://pngu.mgh.harvard.edu/purcell/plink/) and the phase 2, HapMap CEU samples, release 23a, as the imputation reference panel (58). We included the imputed genotypes only if the imputation quality was above 0.8 (Supplementary Material, Table S3). The SNP rs42041 showed deviation from Hardy–Weinberg equilibrium (HWE) in the Dutch GWAS imputed data set, so this SNP was genotyped by TaqMan in Dutch CD cases and the blood bank control group ($n = 833$). NELSON controls ($n = 850$) were not genotyped for the rs42041. TaqMan genotyping was performed following the manufacturer's specifications. DNA samples were processed in 384-well plates; each plate contained 8 negative controls and 16 genotyping controls [four duplicates of four different samples obtained from the Centre d'Etude du Polymorphisme Humain (CEPH)].

UK data sets. For RA, we used freely available genotype data for 9 SNPs from RA patients from the GWAS preformed by the Wellcome Trust Case Control Consortium (WTCCC) in 2007 (13), comprising 1860 UK RA cases and 2938 UK controls. For the remaining 13 SNPs, we used imputed genotypes, all with imputation quality >0.8 (https://www.wtccc.org.uk/ccc1/summary_stats.shtml data access 21 August 2008).

For CD, genotyping results for 15 SNPs were extracted from the UK GWAS study (21). The non-genotyped SNPs were imputed in the UK cases and controls (imputation quality (info) >0.8, except for the SNP rs42041, which could not be imputed with sufficient quality and was therefore excluded from the meta-analysis (Supplementary Material, Table S3).

The genotyping method and imputation quality in each cohort is presented in Supplementary Material, Table S3A–C. The frequency of missing genotypes was below 5% for all the SNPs in all the genotyped cohorts.

Population substructure analysis

Multidimensional scaling analysis (implemented in Plink) was applied to all the samples for which genome-wide genotype data were available. The values of the first five components were further plotted against each other. We detected outlying samples only for the first three components. Samples exceeding two standard deviations from the mean for the first three components were excluded, ensuring that only the most homogeneous samples were included in our final analysis and that there was no population stratification.

As a part of the ongoing GWAS in celiac disease, we applied a multidimensional scaling (MDS) analysis to assess whether there was a large population substructure between the UK and Dutch cohorts. Both cohorts created tight clusters and, as expected, the two first components could distinguish between the two populations. However, over 40% of the samples shared a common part in the plot. Both cohorts were also merged with the HapMap2 data and the MDS analysis was repeated. Both cohorts mapped perfectly to the CEU population, any outlying samples that mapped outside the cluster were excluded from further analysis.

Statistical analysis

We calculated HWE for all the genotyped SNPs by comparing the expected and observed genotypes in a 2×3 chi-square table. None of the markers deviated significantly from HWE in cases or controls ($P > 0.01$) in any of the populations. Association analysis was performed using chi-square statistics with two-tailed P-values (implemented in Plink v1.05) (58). In the meta-analysis, we combined the information on allele counts for the Dutch and UK cohorts separately for CD and RA, in the Mantel–Haenszel chi-square association test with two clusters. P-values, odds ratios (ORs) and 95% confidence intervals (95% CIs) were calculated using Plink v1.05.

To estimate the heterogeneity of the association tests between the populations, we performed the Breslow–Day test. All except two SNPs described in this study were negative for this test. Rs4750316 and rs1678542 showed significant results for the Breslow–Day test in the RA cohort and we therefore applied the random-effect model for non-heterogeneous studies (the DerSimonian–Laird method implemented in the rmeta package for R; www.r-project.org).

Power calculation

We calculated the power of our sample size to detect significant associations using the Genetic Power Calculator (http://pngu.

mgh.harvard.edu/~purcell/gpc/) (59). The calculations were performed separately for the Dutch cohort and for the meta-analysis with UK collections and are summarized in Supplementary Material, Table S5.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

G.T. was awarded a Ter Meulen Fund travel grant by the Royal Netherlands Academy of Arts and Sciences (KNAW). We thank all the individuals who participated in the study, Jackie Senior for critically reading the manuscript, and Flip Mulder for help with graphic design.

Conflict of Interest statement. None declared.

FUNDING

The study was supported by the Celiac Disease Consortium, an Innovative Cluster approved by the Netherlands Genomics Initiative and partially funded by the Dutch Government (BSIK03009), by the Netherlands Organization for Scientific Research (NWO, VICI grant 918.66.620 to CW), the Wellcome Trust (WT084743MA). We acknowledge use of DNA from the British 1958 Birth Cohort collection, funded by the UK Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. M.C. was supported by a personal grant from the Netherlands Organization for Scientific Research (NOW, VENI grant 916.76.020). GT was awarded a Ter Meulen Fund travel grant by the Royal Netherlands Academy of Arts and Sciences (KNAW). We thank all the individuals who participated in the study, Jackie Senior for critically reading the manuscript, and Flip Mulder for help with graphic design.

REFERENCES

1. Oliver, J.E., Worthington, J. and Silman, A.J. (2006) Genetic epidemiology of rheumatoid arthritis. *Curr. Opin. Rheumatol.*, **18**, 141–146.
2. Oliver, J.E. and Silman, A.J. (2006) Risk factors for the development of rheumatoid arthritis. *Scand. J. Rheumatol.*, **35**, 169–174.
3. Somers, E.C., Thomas, S.L., Smeeth, L. and Hall, A.J. (2006) Autoimmune diseases co-occurring within individuals and within families: a systematic review. *Epidemiology*, **17**, 202–217.
4. Zhernakova, A., van Diemen, C.C. and Wijmenga, C. (2009) Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat. Rev. Genet.*, **10**, 43–55.
5. Sollid, L.M. (2002) Coeliac disease: dissecting a complex inflammatory disorder. *Nat. Rev. Immunol.*, **2**, 647–655.
6. Molberg, O. and Sollid, L.M. (2006) A gut feeling for joint inflammation—using coeliac disease to understand rheumatoid arthritis. *Trends Immunol.*, **27**, 188–194.
7. Bourne, J.T., Kumar, P., Huskisson, E.C., Mageed, R., Unsworth, D.J. and Wojtulewski, J.A. (1985) Arthritis and coeliac disease. *Ann. Rheum. Dis.*, **44**, 592–598.
8. Collin, P., Korpela, M., Hallstrom, O., Viander, M., Keyrilainen, O. and Maki, M. (1992) Rheumatic complaints as a presenting symptom in patients with coeliac disease. *Scand. J. Rheumatol.*, **21**, 20–23.
9. Neuhausen, S.L., Steele, L., Ryan, S., Mousavi, M., Pinto, M., Osann, K.E., Flodman, P. and Zone, J.J. (2008) Co-occurrence of celiac disease and other autoimmune diseases in celiacs and their first-degree relatives. *J. Autoimmun.*, **31**, 160–165.
10. Parke, A.L., Fagan, E.A., Chadwick, V.S. and Hughes, G.R. (1984) Coeliac disease and rheumatoid arthritis. *Ann. Rheum. Dis.*, **43**, 378–380.
11. Barton, A., Thomson, W., Ke, X., Eyre, S., Hinks, A., Bowes, J., Gibbons, L., Plant, D., Wilson, A.G., Marinou, I. *et al.* (2008) Re-evaluation of putative rheumatoid arthritis susceptibility genes in the post-genome wide association study era and hypothesis of a key pathway underlying susceptibility. *Hum. Mol. Genet.*, **17**, 2274–2279.
12. Barton, A., Thomson, W., Ke, X., Eyre, S., Hinks, A., Bowes, J., Plant, D., Gibbons, L.J., Wilson, A.G., Bax, D.E. *et al.* (2008) Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. *Nat. Genet.*, **40**, 1156–1159.
13. Consortium, W.T.C.C. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, **447**, 661–678.
14. Hunt, K.A., Zhernakova, A., Turner, G., Heap, G.A., Franke, L., Bruinenberg, M., Romanos, J., Dinesen, L.C., Ryan, A.W., Panesar, D. *et al.* (2008) Newly identified genetic risk variants for celiac disease related to the immune response. *Nat. Genet.*, **40**, 395–402.
15. Plenge, R.M., Cotsapas, C., Davies, L., Price, A.L., de Bakker, P.I., Maller, J., Pe'er, I., Burtt, N.P., Blumenstiel, B., DeFelice, M. *et al.* (2007) Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nat. Genet.*, **39**, 1477–1482.
16. Plenge, R.M., Seielstad, M., Padyukov, L., Lee, A.T., Remmers, E.F., Ding, B., Liew, A., Khalili, H., Chandrasekaran, A., Davies, L.R. *et al.* (2007) TRAF1-C5 as a risk locus for rheumatoid arthritis—a genomewide study. *N. Engl. J. Med.*, **357**, 1199–1209.
17. Raychaudhuri, S., Remmers, E.F., Lee, A.T., Hackett, R., Guiducci, C., Burtt, N.P., Gianniny, L., Korman, B.D., Padyukov, L., Kurreeman, F.A. *et al.* (2008) Common variants at CD40 and other loci confer risk of rheumatoid arthritis. *Nat. Genet.*, **40**, 1216–1223.
18. Remmers, E.F., Plenge, R.M., Lee, A.T., Graham, R.R., Hom, G., Behrens, T.W., de Bakker, P.I., Le, J.M., Lee, H.S., Batliwalla, F. *et al.* (2007) STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N. Engl. J. Med.*, **357**, 977–986.
19. Thomson, W., Barton, A., Ke, X., Eyre, S., Hinks, A., Bowes, J., Donn, R., Symmons, D., Hider, S., Bruce, I.N. *et al.* (2007) Rheumatoid arthritis association at 6q23. *Nat. Genet.*, **39**, 1431–1433.
20. Trynka, G., Zhernakova, A., Romanos, J., Franke, L., Hunt, K., Turner, G., Platteel, M., Ryan, A.W., de Kovel, C., Barisani, D. *et al.* (2009) Coeliac disease associated risk variants in TNFAIP3 and REL implicate altered NF-{kappa}B signalling. *Gut.*, **58**, 1078–1083.
21. van Heel, D.A., Franke, L., Hunt, K.A., Gwilliam, R., Zhernakova, A., Inouye, M., Wapenaar, M.C., Barnardo, M.C., Bethel, G., Holmes, G.K. *et al.* (2007) A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nat. Genet.*, **39**, 827–829.
22. Zhernakova, A., Alizadeh, B.Z., Bevova, M., van Leeuwen, M.A., Coenen, M.J., Franke, B., Franke, L., Posthumus, M.D., van Heel, D.A., van der Steege, G. *et al.* (2007) 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**, 1284–1288.
23. Liu, Y., Helms, C., Liao, W., Zaba, L.C., Duan, S., Gardner, J., Wise, C., Miner, A., Malloy, M.J., Pullinger, C.R. *et al.* (2008) A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet.*, **4**, e1000041.
24. Todd, J.A., Walker, N.M., Cooper, J.D., Smyth, D.J., Downes, K., Plagnol, V., Bailey, R., Nejentsev, S., Field, S.F., Payne, F. *et al.* (2007) Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat. Genet.*, **39**, 857–864.
25. Fung, E.Y., Smyth, D.J., Howson, J.M., Cooper, J.D., Walker, N.M., Stevens, H., Wicker, L.S. and Todd, J.A. (2009) Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus. *Genes Immun.*, **10**, 188–191.
26. Graham, R.R., Cotsapas, C., Davies, L., Hackett, R., Lessard, C.J., Leon, J.M., Burtt, N.P., Guiducci, C., Parkin, M., Gates, C. *et al.* (2008) Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nat. Genet.*, **40**, 1059–1061.

27. Musone, S.L., Taylor, K.E., Lu, T.T., Nititham, J., Ferreira, R.C., Ortmann, W., Shifrin, N., Petri, M.A., Kamboh, M.I., Manzi, S. *et al.* (2008) Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. *Nat. Genet.*, **40**, 1062–1064.
28. van der Helm-van Mil, A.H. and Huizinga, T.W. (2008) Advances in the genetics of rheumatoid arthritis point to subclassification into distinct disease subsets. *Arthritis Res. Ther.*, **10**, 205.
29. Festen, E.A., Goyette, P., Scott, R., Annese, V., Zhernakova, A., Lian, J., Lefebvre, C., Brant, S.R., Cho, J.H., Silverberg, M.S. *et al.* (2009) Genetic variants in the region harbouring IL2/IL21 associated with ulcerative colitis. *Gut*, **58**, 799–804.
30. Glas, J., Stallhofer, J., Ripke, S., Wetzke, M., Pfennig, S., Klein, W., Epplen, J.T., Griga, T., Schiemann, U., Lacher, M. *et al.* (2009) Novel genetic risk markers for ulcerative colitis in the IL2/IL21 region are in epistasis with IL23R and suggest a common genetic background for ulcerative colitis and celiac disease. *Am. J. Gastroenterol.*, **104**, 1737–1744.
31. Marquez, A., Orozco, G., Martinez, A., Palomino-Morales, R., Fernandez-Arquero, M., Mendoza, J.L., Taxonera, C., Diaz-Rubio, M., Gomez-Garcia, M., Nieto, A. *et al.* (2009) Novel association of the Interleukin 2-Interleukin 21 Region with inflammatory bowel disease. *Am. J. Gastroenterol.* [Epub ahead of print].
32. Nair, R.P., Duffin, K.C., Helms, C., Ding, J., Stuart, P.E., Goldgar, D., Gudjonsson, J.E., Li, Y., Tejasvi, T., Feng, B.J. *et al.* (2009) Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat. Genet.*, **41**, 199–204.
33. Castellanos-Rubio, A., Santin, I., Irastorza, I., Castano, L., Carlos Vitoria, J. and Ramon Bilbao, J. (2009) TH17 (and TH1) signatures of intestinal biopsies of CD patients in response to gliadin. *Autoimmunity*, **42**, 69–73.
34. Pernis, A.B. (2009) Th17 cells in rheumatoid arthritis and systemic lupus erythematosus. *J. Intern. Med.*, **265**, 644–652.
35. Deenick, E.K. and Tangye, S.G. (2007) Autoimmunity: IL-21: a new player in Th17-cell differentiation. *Immunol. Cell. Biol.*, **85**, 503–505.
36. Jungel, A., Distler, J.H., Kurowska-Stolarska, M., Seemayer, C.A., Seibl, R., Forster, A., Michel, B.A., Gay, R.E., Emmrich, F., Gay, S. *et al.* (2004) Expression of interleukin-21 receptor, but not interleukin-21, in synovial fibroblasts and synovial macrophages of patients with rheumatoid arthritis. *Arthritis Rheum.*, **50**, 1468–1476.
37. Young, D.A., Hegen, M., Ma, H.L., Whitters, M.J., Albert, L.M., Lowe, L., Senices, M., Wu, P.W., Sibley, B., Leathurby, Y. *et al.* (2007) Blockade of the interleukin-21/interleukin-21 receptor pathway ameliorates disease in animal models of rheumatoid arthritis. *Arthritis Rheum.*, **56**, 1152–1163.
38. Gudbjartsson, D.F., Bjornsdottir, U.S., Halapi, E., Helgadottir, A., Sulem, P., Jonsdottir, G.M., Thorleifsson, G., Helgadottir, H., Steinthorsdottir, V., Stefansson, H. *et al.* (2009) Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.*, **41**, 342–347.
39. Fitau, J., Boulday, G., Coulon, F., Quillard, T. and Charreau, B. (2006) The adaptor molecule Lnk negatively regulates tumor necrosis factor-alpha-dependent VCAM-1 expression in endothelial cells through inhibition of the ERK1 and -2 pathways. *J. Biol. Chem.*, **281**, 20148–20159.
40. Jin, L., Kern, M.J., Otey, C.A., Wamhoff, B.R. and Somlyo, A.V. (2007) Angiotensin II, focal adhesion kinase, and PRX1 enhance smooth muscle expression of lipoma preferred partner and its newly identified binding partner palladin to promote cell migration. *Circ. Res.*, **100**, 817–825.
41. Petit, M.M., Meulemans, S.M. and Van de Ven, W.J. (2003) The focal adhesion and nuclear targeting capacity of the LIM-containing lipoma-preferred partner (LPP) protein. *J. Biol. Chem.*, **278**, 2157–2168.
42. Daheron, L., Veinstein, A., Brizard, F., Drabkin, H., Lacotte, L., Guilhot, F., Larsen, C.J., Brizard, A. and Roche, J. (2001) Human LPP gene is fused to MLL in a secondary acute leukemia with a t(3;11) (q28;q23). *Genes Chromosomes Cancer*, **31**, 382–389.
43. Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., Silverberg, M.S., Taylor, K.D., Barmada, M.M. *et al.* (2008) Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.*, **40**, 955–962.
44. Vang, T., Miletic, A.V., Bottini, N. and Mustelin, T. (2007) Protein tyrosine phosphatase PTPN22 in human autoimmunity. *Autoimmunity*, **40**, 453–461.
45. Smyth, D.J., Plagnol, V., Walker, N.M., Cooper, J.D., Downes, K., Yang, J.H., Howson, J.M., Stevens, H., McManus, R., Wijmenga, C. *et al.* (2008) Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N. Engl. J. Med.*, **359**, 2767–2777.
46. Cooper, J.D., Smyth, D.J., Smiles, A.M., Plagnol, V., Walker, N.M., Allen, J.E., Downes, K., Barrett, J.C., Healy, B.C., Mychaleckyj, J.C. *et al.* (2008) Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat. Genet.*, **40**, 1399–1401.
47. Hayashi, K. and Altman, A. (2007) Protein kinase C theta (PKCtheta): a key player in T cell life and death. *Pharmacol. Res.*, **55**, 537–544.
48. Gonzalez, L.C., Loyet, K.M., Calemine-Fenaux, J., Chauhan, V., Wranik, B., Ouyang, W. and Eaton, D.L. (2005) A coreceptor interaction between the CD28 and TNF receptor family members B and T lymphocyte attenuator and herpesvirus entry mediator. *Proc. Natl Acad. Sci. USA*, **102**, 1116–1121.
49. Heo, S.K., Ju, S.A., Lee, S.C., Park, S.M., Choe, S.Y., Kwon, B., Kwon, B.S. and Kim, B.S. (2006) LIGHT enhances the bactericidal activity of human monocytes and neutrophils via HVEM. *J. Leukoc. Biol.*, **79**, 330–338.
50. Rashid, T. and Ebringer, A. (2008) Rheumatoid arthritis in smokers could be linked to Proteus urinary tract infections. *Med. Hypotheses*, **70**, 975–980.
51. Stene, L.C., Honeyman, M.C., Hoffenberg, E.J., Haas, J.E., Sokol, R.J., Emery, L., Taki, I., Norris, J.M., Erlich, H.A., Eisenbarth, G.S. *et al.* (2006) Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. *Am. J. Gastroenterol.*, **101**, 2333–2340.
52. Vaahtovuo, J., Munukka, E., Korkeamaki, M., Luukkainen, R. and Toivanen, P. (2008) Fecal microbiota in early rheumatoid arthritis. *J. Rheumatol.*, **35**, 1500–1505.
53. Toonen, E.J., Coenen, M.J., Kievit, W., Fransen, J., Eijsbouts, A.M., Scheffer, H., Radstake, T.R., Creemers, M.C., de Rooij, D.J., van Riel, P.L. *et al.* (2008) The tumour necrosis factor receptor superfamily member 1b 676T>G polymorphism in relation to response to infliximab and adalimumab treatment and disease severity in rheumatoid arthritis. *Ann. Rheum. Dis.*, **67**, 1174–1177.
54. Welsing, P.M. and van Riel, P.L. (2004) The Nijmegen inception cohort of early rheumatoid arthritis. *J. Rheumatol. Suppl.*, **69**, 14–21.
55. Arnett, F.C., Edworthy, S.M., Bloch, D.A., McShane, D.J., Fries, J.F., Cooper, N.S., Healey, L.A., Kaplan, S.R., Liang, M.H., Luthra, H.S. *et al.* (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.*, **31**, 315–324.
56. United European Gastroenterology. (2001) When is a coeliac a coeliac? Report of a working group of the United European Gastroenterology Week in Amsterdam, 2001. *Eur. J. Gastroenterol. Hepatol.*, **13**, 1123–1128.
57. van Iersel, C.A., de Koning, H.J., Draisma, G., Mali, W.P., Scholten, E.T., Nackaerts, K., Prokop, M., Habbema, J.D., Oudkerk, M. and van Klaveren, R.J. (2007) Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). *Int. J. Cancer*, **120**, 868–874.
58. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
59. Purcell, S., Cherny, S.S. and Sham, P.C. (2003) Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*, **19**, 149–150.