Association between *IL2*/*IL21* and *SH2B3* polymorphisms and risk of celiac disease: a meta-analysis

C.C. Guo[1], W.H. Huang[1], N. Zhang[1], F. Dong[1], L.P. Jing[1], Y. Liu[1], X.G. Ye[1], D. Xiao[1], M.L. Ou[1], B.H. Zhang[1], M. Wang[1], W.K. Liang[1], G. Yang[2,3] and C.X. Jing[1,3]

[1]Department of Epidemiology, School of Medicine, Jinan University, Guangzhou, Guangdong, China
[2]Department of Parasitology, School of Medicine, Jinan University, Guangzhou, Guangdong, China
[3]Key Laboratory of Enviromental Exposure and Health in Guangzhou, Jinan University, Guangzhou, Guangdong, China

Corresponding authors: C.X. Jing / G. Yang
E-mail: jcxphd@gmail.com / guangyangphd@gmail.com

Genet. Mol. Res. 14 (4): 13221-13235 (2015)
Received May 4, 2015
Accepted July 29, 2015
Published October 26, 2015
DOI http://dx.doi.org/10.4238/2015.October.26.19

ABSTRACT. Celiac disease (CD) is a common autoimmune disorder characterized by heightened immunological response to ingested gluten. Certain gene polymorphisms of *IL2*/*IL21* (rs6822844 and rs6840978) and *SH2B3* (rs3184504) may influence susceptibility to CD, although the effects remain unclear. We performed a meta-analysis of the associations between rs6822844, rs6840978, and rs3184504 polymorphisms and CD risk. PubMed, EMBASE, and the China National Knowledge Infrastructure were searched. ORs and 95%CIs of each single nucleotide polymorphism (SNP) were estimated using the fixed-effect model if I^2 < 50% in the test of heterogeneity; otherwise, the random-effect model was used. Our meta-analysis included 12,986 CD cases and 28,733 controls from 16 independent samples, and the analysis of each SNP contained a subset of the total. We found that the minor allele T of both rs6822844 (T *vs* G, OR

= 0.72, 95%CI = 0.67-0.78, P < 0.001) and rs6840978 (T *vs* C, OR = 0.76, 95%CI = 0.71-0.83, P < 0.001) in *IL2*/*IL21* significantly decreased the risk of CD. However, the minor allele A of rs3184504 (A *vs* G, OR = 1.18, 95%CI = 1.12-1.24, P < 0.001) in *SH2B3* significantly increased CD susceptibility. The estimated lambda values were 0.49, 0.50, and 0.53 for rs6822844, rs6840978, and rs3184504, respectively, suggesting that a co-dominant model of genotype effect was most appropriate for the three SNPs. Our results support associations between the three SNPs and CD and provide a strong argument for further research.

Key words: *IL2*/*IL21*; *SH2B3*; Single nucleotide polymorphisms; Celiac disease; Meta-analysis

INTRODUCTION

Celiac disease (CD) is a common autoimmune disorder that is characterized by flattened villi on the small bowel mucosa. It is induced in genetically susceptible people who ingest the dietary protein gluten (Di Sabatino and Corazza, 2009). CD primarily occurs in Caucasians, with a prevalence rate of approximately 1% (Catassi et al., 2014). Genetic linkage studies have identified that the human leukocyte antigen (HLA) locus is the strongest genetic factor for CD; however, HLA only accounts for 53% of genetic susceptibility, which suggests that non-HLA genes must be involved in disease susceptibility (Sollid and Lie, 2005).

The first genome-wide association study (GWAS) for CD included 778 cases and 1,422 controls from the UK; the most significant locus outside the HLA region was 4q27, containing the interleukin 2 (*IL2*) and interleukin 21 (*IL21*) genes (van Heel et al., 2007). Both *IL2* and *IL21* are highly interesting candidate genes for CD susceptibility, because the cytokines they encode are T-cell-derived and promote autocrine T-cell activation and proliferation (Adamovic et al., 2008). IL21 cooperates with IL2 to promote interferon gamma (IFN-γ) synthesis (Kasaian et al., 2002). IFN-γ is the dominant pro-inflammatory cytokine inducing T-helper-cell type 1 responses, which leads to the development of celiac lesions. Additionally, IL21 can prolong chronic inflammation and increase celiac damage by facilitating the recruitment of immune cells within the inflamed tissue (Caruso et al., 2007), the expansion of autoreactive T-cells (King et al., 2004), and the synthesis of extracellular matrix metalloproteinases (Monteleone et al., 2006). Two single-nucleotide polymorphisms (SNPs) (rs6822844 and rs6840978), located in the inter-gene region between *IL2* and *IL21*, have been identified as associated risk factors that contribute to the development of CD (van Heel et al., 2007; Hunt et al., 2008; Romanos et al., 2009; Dubois et al., 2010; Maiti et al., 2010; Plaza-Izurieta et al., 2011, Sperandeo et al., 2011). However, several research teams have reported contradictory results that the minor T alleles in rs6822844 and rs6840978 are associated with lower risk of developing CD.

SH2B3 showed significant association with CD in a UK GWAS replication study on the eight most strongly associated non-HLA regions (Hunt et al., 2008). *SH2B3* is located on chromosome 12 (12q24) and is also known as *LNK*. According to evidence from biopsies, it is strongly expressed in the inflamed small intestines of celiac patients, which may reflect leukocyte recruitment and activation (Hunt et al., 2008). SH2B3 regulates T-cell receptor-, growth factor-, and cytokine receptor-mediated signaling, and may be a potential candidate for CD susceptibility (Fitau et al., 2006). The rs3184504 is a non-synonymous SNP in exon 3 of *SH2B3* and several studies have shown that the minor allele A in rs3184504 may increase the risk of CD (Hunt et al., 2008, Romanos et al., 2009). However, the

relationship between the rs3184504 polymorphisms and CD remains unclear.

Although a number of studies have revealed associations between the three SNPs described above and CD risk, relatively small sample sizes and varying population characteristics may have confounded the results. We performed a meta-analysis of all the available studies to accurately estimate the relationship between rs6822844, rs6840978, and rs3184504 polymorphisms and CD risk.

MATERIAL AND METHODS

Search strategy

We conducted electronic searches (up to November 2014) of PubMed, EMBASE, and the China National Knowledge Infrastructure (CNKI) databases to identify relevant studies on the associations between rs6822844, rs6840978, and rs3184504 polymorphisms and CD risk. Data extraction was carried out independently by two investigators (MW and WKL). The search terms were as follows: "[(celiac disease or coeliac disease) and (SH2B3 or SH2B adaptor protein 3) and (IL2/IL21 or Interleukin 2/Interleukin 21) and (rs3184504 or rs6822844 or rs6840978) and (gene or allele or polymorphism)]". Searching was limited to human studies published in English or Chinese. We also examined the reference lists of the retrieved articles for additional relevant publications.

Inclusion and exclusion criteria

Association studies were included if the following criteria were met: 1) the studies were on humans; 2) the studies had a case-control design; 3) the studies were on the associations between *IL2*/*IL21* (rs6822844 or rs6840978) or *SH2B3* (rs3184504) polymorphisms and CD; 4) the data from the studies were sufficient to allow extraction; 5) the ORs and corresponding 95%CIs were reported; and 6) the studies were written in English or Chinese.

Studies were excluded if 1) the case and control subjects were biologically related; 2) the data were insufficient to fulfill further meta-analysis research; or 3) the studies comprised meta-analyses, reviews, or meeting abstracts.

Data extraction

Two investigators (M.W. and W.K.L.) independently extracted the information from all the eligible papers using a standardized data extraction form. Any disagreement was resolved by consensus. We collected the following information from each study: the name of the first author, the year of publication, the gene locus, the nationality, the male/female numbers of cases and controls, the Hardy-Weinberg equilibrium (HWE), the genotyping method, and the frequencies of the minor alleles.

Assessment of bias risk

Study quality was assessed using a risk-of-bias score for genetic association studies developed by Thakkinstian et al. (2011) (**Table S1**). The assessment considered five domains: information bias, confounding bias, selective reporting of outcomes, population stratification, and assessment of HWE in the control group. Each item was classified as "yes", "no", or "unclear", indicating low risk, high risk, and insufficient information, respectively.

Statistical analysis

We used R version 3.1.1 software (University of Auckland, New Zealand) for our statistical analyses. The HWE in the control group was evaluated using a web-based program (http://ihg.gsf.de/cgi-bin/hw/hwa1.pl) and the study was considered to be in disequilibrium if HWE was not evident (at $P < 0.05$). We performed both per-allele and per-genotype analyses to estimate the strength of association between the rs6822844, rs6840978, and rs3184504 polymorphisms and CD risk.

Per-allele analysis

We used the Z-test to determine the statistical significance of the pooled ORs, and P values less than 0.05 were considered statistically significant. Heterogeneity across studies was checked using the Cochran's Q-test, and the significance level was set at 0.10. If a P value was over 0.10, a fixed-effects model was selected, otherwise we chose a random-effects model. The degree of heterogeneity was quantified by calculating I^2 values ($I^2 < 25\%$, no heterogeneity; $25\% < I^2 < 50\%$, moderate heterogeneity; $50\% < I^2 < 75\%$, large heterogeneity; and $I^2 > 75\%$, extreme heterogeneity) (Higgins et al., 2003). The population attributable risk (PAR%) for the minor allele was calculated based on results from discrete-time models (Rossman et al., 2003).

Per-genotype analysis

For each SNP, we used G and g to represent risk and non-risk alleles, and GG, Gg, and gg to represent the minor homozygous, heterozygous, and common homozygous genotypes, respectively. Two distinct ORs: GG versus gg (OR_{GG}) and Gg versus gg (OR_{Gg}) were estimated for each study. The parameter lambda (λ), defined as the ratio of log OR_{GG} to log OR_{Gg}, was used to model the genetic effect (Minelli et al., 2005). The value of λ ranges from 0 to 1 and values equal to 0, 0.5, and 1 suggested the recessive (GG *vs* Gg + gg), co-dominant (GG *vs* gg; Gg *vs* gg), and dominant (GG + Gg *vs* gg) genetic models, respectively. If $\lambda > 1$ or $\lambda < 0$, heterosis was considered. For λ, the WinBUGS 1.4.3 software (University of Cambridge, England, UK) was used. All models were burnt in using 1,000 iterations, followed by 10,000 iterations to estimate the parameters.

A contour-enhanced funnel plot and Egger linear regression were used to assess publication bias. A contour-enhanced funnel plot displays areas of statistical significance (e.g., < 0.01, < 0.05, < 0.1) on a funnel plot (Peters et al., 2008). If the supposed missing studies are in areas of statistical non-significance, the asymmetry may be due to publication bias. Conversely, if missing studies are in areas of statistical significance, it is more likely that the asymmetry has arisen from factors other than publication bias (e.g., variable study quality). A sensitivity analysis was performed by omitting each individual study to reflect the influence of the single study on the pooled OR.

RESULTS

Literature selection

A total of 90 studies were identified in PubMed, EMBASE and CNKI. After duplicates had been removed, 66 titles and abstracts were screened and 54 were determined to be ineligible for the reasons given in Figure 1. After retrieval and review of the 12 remaining studies, we excluded three studies for the following reasons: one concerned other SNPs, one was a review, and one

lacked sufficient data. Ultimately, nine studies involving 16 sub-study collections were finally included in this meta-analysis (van Heel et al., 2007; Hunt et al., 2008; Smyth et al., 2008; Coenen et al., 2009; Romanos et al., 2009; Dubois et al., 2010; Maiti et al., 2010; Plaza-Izurieta et al., 2011; Sperandeo et al., 2011). The characteristics of the included studies are presented in Table 1, with 12,986 CD patients and 28,733 controls. In summary, six studies focused on the *IL2*/*IL21* rs6822844 polymorphism (van Heel et al., 2007; Hunt et al., 2008; Romanos et al., 2009; Maiti et al., 2010; Plaza-Izurieta et al., 2011; Sperandeo et al., 2011); four studies focused on the *IL2*/*IL21* rs6840978 polymorphism (van Heel et al., 2007; Hunt et al., 2008; Dubois et al., 2010; Sperandeo et al., 2011); and six studies focused on the *SH2B3* rs3184504 polymorphism (Hunt et al., 2008; Smyth et al., 2008; Coenen et al., 2009; Romanos et al., 2009; Dubois et al., 2010; Plaza-Izurieta et al., 2011).

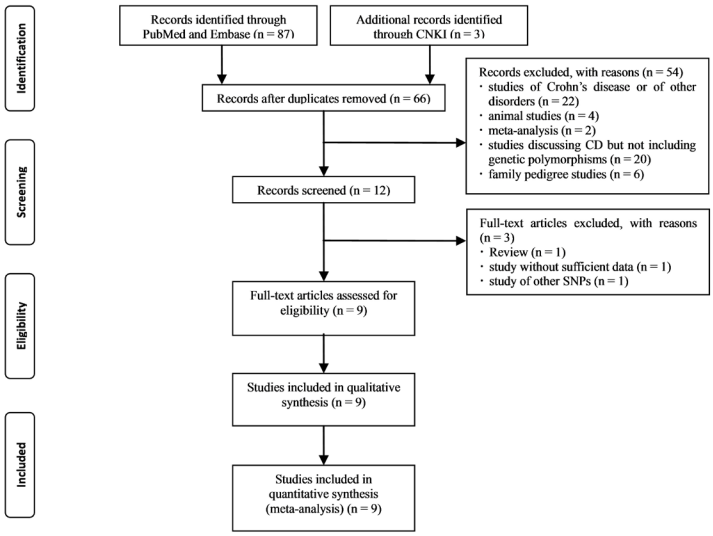

Figure 1. Flow chart showing study selection procedure.

Table 1. Characteristics of the studies included in the meta-analysis.

First author	Year	Gene locus	Nationality	Male/subjects		Genotyping method	HWE test
				Case	Control		
Maiti	2010	rs6822844	Argentinean	NA/189	NA/222	TaqMan assay	Yes
Romanos	2009	rs3184504/rs6822844	Italian	NA/538	NA/593	TaqMan assay	Yes
Smyth	2008	rs3184504	British	637/2560	NA/9339	TaqMan assay	Yes
Plaza-Izurieta	2011	rs3184504/rs6822844	Spanish	NA/1094	NA/540	TaqMan assay	NA
Sperandeo	2011	rs6822844/rs6840978	Italian	NA/643	NA/711	TaqMan assay	NA
Hunt	2008	rs3184504/rs6822844/rs6840978	British	166/719	584/1561	GoldenGate assay	Yes
			Irish	139/416	283/957	GoldenGate assay	Yes
			Dutch	170/508	540/888	GoldenGate assay	Yes
van Heel	2007	rs6822844/rs6840978	Dutch	158/508	569/929	PCR-RFLP	Yes
			Irish	140/483	274/560	PCR-RFLP	Yes
Coenen	2009	rs3184504	Dutch	252/795	1332/1683	TaqMan assay	Yes
Dubois	2010	rs3184504/rs6840978	British1	NA/737	NA/2596	PCR-RFLP	Yes
			British2	NA/1849	NA/4936	TaqMan assay	Yes
			Finnish	NA/647	NA/1829	TaqMan assay	Yes
			Dutch	NA/803	NA/846	TaqMan assay	Yes
			Italian	NA/497	NA/543	TaqMan assay	Yes

NA = not available; HWE = Hardy-Weinberg equilibrium; Yes = in HWE; No = not in HWE; CD = celiac disease.

Bias assessment

The results of the bias assessment are presented in Table 2. The risk of bias was highest in confounding bias (8/9; 88.9%), followed by non-compliance with HWE (3/9; 33.3%), and quality control of genotyping (2/9; 22.2%). However, the remaining four indicators showed no risk of bias (0/9; 0.0%), suggesting the relatively high quality of our meta-analysis.

Table 2. Risks of bias assessment.

Author	Ascertainment of CD	Ascertainment of controls	Quality control for genotyping	Population stratification	Confounding bias	Selective outcome report	HWE
Maiti	Yes	Yes	Yes	Yes	No	Yes	No
Romanos	Yes	Yes	Yes	Yes	No	Yes	Yes
Smyth	Yes	Yes	Yes	Yes	No	Yes	Yes
Plaza-Izurieta	Yes	Yes	Unclear	Yes	No	Yes	No
Sperandeo	Yes	Yes	Unclear	Yes	Yes	Yes	No
Hunt	Yes	Yes	Yes	Yes	No	Yes	Yes
van Heel	Yes	Yes	Yes	Yes	No	Yes	Yes
Coenen	Yes	Yes	Yes	Yes	No	Yes	Yes
Dubois	Yes	Yes	Yes	Yes	No	Yes	Yes

Yes = high risk of bias; No = low risk of bias; Unclear = unclear risk of bias; CD = celiac disease; HWE = Hardy-Weinberg equilibrium.

Association between the *IL2*/*IL21* rs6822844 polymorphism and CD risk

Nine sub-studies, including 4,781 cases and 6,697 controls, explored the association between the *IL2*/*IL21* rs6822844 polymorphism and CD risk (Table 3). The pooled frequency of the minor T allele was 0.11 (95%CI = 0.10-0.13) in the CD group (I^2 = 86%, P < 0.001) and 0.15 (95%CI = 0.12-0.17) in the control group (I^2 = 94%, P < 0.001), estimated by the random-effects model. The pooled OR (T *vs* G) was 0.72 (95%CI = 0.67-0.78, P < 0.001; see Figure 2A) without heterogeneity (I^2 = 0%, P = 0.448), suggesting that individuals carrying the minor T allele had a 28% lower risk of developing CD compared with those carrying the G allele, and the PAR% for the minor T allele was 3.82%, suggesting that Caucasians carrying the T allele in *IL2*/*IL21* rs6822844 had a reduced risk of developing CD.

A contour-enhanced funnel plot is presented in the first part of Figure 3. There were two "missing" studies (Maiti et al., 2010; Plaza-Izurieta et al., 2011), which were both within statistical significance (P < 0.01), suggesting that there was no publication bias. The Egger linear regression test also indicated that there was no publication bias (P = 0.173).

The λ value was 0.49 (95%CI = 0.26-0.83), suggesting that the co-dominant effect was at play (Table 3). Both OR_1 (TT *vs* GG, P = 0.891, I^2 = 0%) and OR_2 (TG *vs* GG, P = 0.467, I^2 = 0%) were homogenous. The pooled OR_1 (0.50; 95%CI = 0.37-0.67) and OR_2 (0.73; 95%CI = 0.67-0.80) values were estimated using a fixed-effects model, which indicated that individuals with TT and TG genotypes had 50 and 27% less risk of developing CD, respectively, compared with those carrying the GG genotype.

Association between the *IL2*/*IL21* rs6840978 polymorphism and CD risk

Eleven sub-studies, including 7,456 cases and 16,090 controls, were examined to determine the association between the *IL2*/*IL21* rs6840978 polymorphism and CD (Table 4). The

pooled frequency of the minor T allele was 15% (95%CI = 14-17%) in the CD group (I^2 = 87%, P < 0.001) and 19% (95%CI = 17-21%) in the non-CD group (I^2 = 96%, P < 0.001), estimated by the random-effects model. The pooled OR (T *vs* C) was 0.76 (95%CI = 0.71-0.82; see Figure 2B) with moderate heterogeneity (P = 0.042, I^2 = 47%), suggesting that individuals carrying the minor T allele had a 24% reduced risk of developing CD than those carrying the C allele, and the PAR% for the minor T allele was 3.44%, suggesting that Caucasians carrying the T allele in *IL2*/*IL21* rs6840978 had a reduced risk of developing CD.

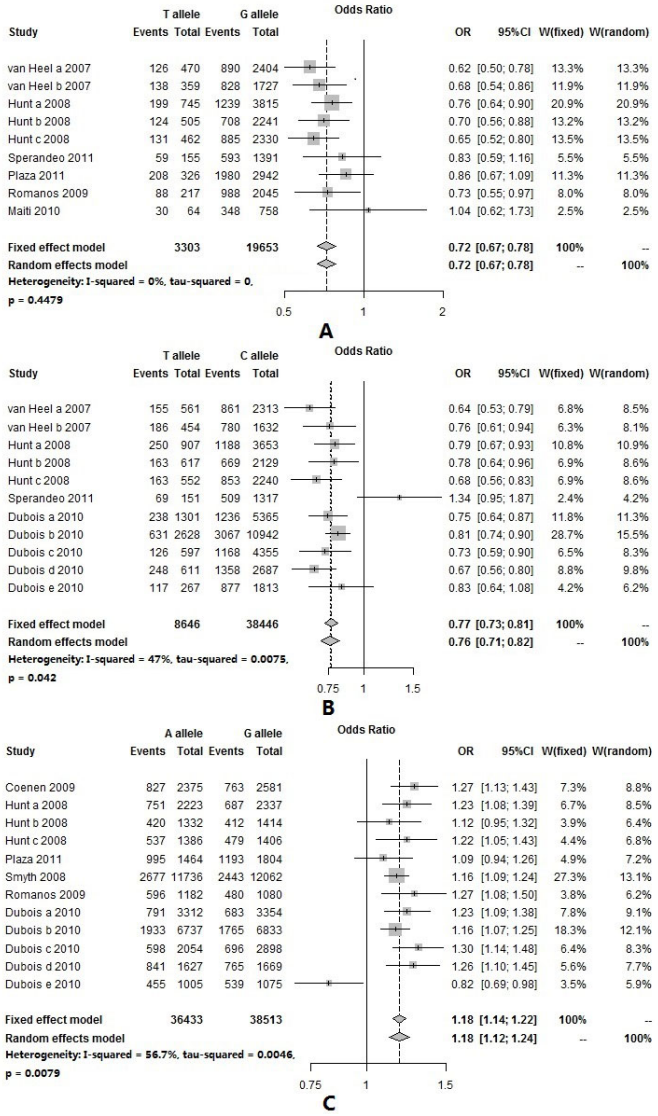

Figure 2. A. Forest plot of the association between the *IL2*/*IL21* rs6822844 polymorphism and celiac disease (CD) risk. **B.** Forest plot of the association between the *IL2*/*IL21* rs6840978 polymorphism and CD risk. **C.** Forest plot of the association between the *SH2B3* rs3184504 polymorphism and CD risk.

Table 3. Genotype frequencies for the *IL2/IL21* rs6822844 polymorphism and genotype effects of studies included in the meta-analysis.

Author	Total number		Case genotype			Control genotype			T allele prevalence	T *vs* G			TT *vs* GG			TG *vs* GG			HWE
	Case	Control	TT	TG	GG	TT	TG	GG		OR	95%CI		OR	95%CI		OR	95%CI		
van Heel[a]	508	929	8	110	390	32	280	617	0.185	0.623	0.500	0.777	0.396	0.180	0.867	0.622	0.482	0.802	0.973
van Heel[b]	483	560	10	118	355	22	177	361	0.197	0.678	0.537	0.856	0.462	0.216	0.990	0.678	0.515	0.893	0.958
Hunt[a]	719	1561	14	171	534	48	450	1063	0.175	0.758	0.636	0.903	0.581	0.317	1.063	0.756	0.617	0.928	0.964
Hunt[b]	416	957	9	106	301	38	305	614	0.199	0.705	0.565	0.879	0.483	0.231	1.012	0.709	0.546	0.920	0.987
Hunt[c]	508	888	8	115	385	31	269	588	0.186	0.646	0.519	0.804	0.394	0.179	0.867	0.653	0.507	0.841	0.973
Sperandeo	326	447	0	59	267	7	82	358	0.107	0.827	0.588	1.163	0.089	0.005	1.571	0.965	0.666	1.397	0.362
Plaza-Izurieta	1094	540	10	188	896	6	106	428	0.109	0.856	0.675	1.087	0.796	0.287	2.205	0.847	0.650	1.104	0.844
Romanos	538	593	4	80	454	7	115	471	0.109	0.730	0.549	0.970	0.593	0.172	2.039	0.722	0.528	0.987	0.995
Maiti	189	222	1	28	160	1	32	189	0.077	1.040	0.623	1.733	1.181	0.073	19.037	1.034	0.597	1.790	1.000
Overall OR										0.723	0.667	0.784	0.496	0.369	0.667	0.733	0.668	0.803	

HWE = Hardy-Weinberg equilibrium; van Heel[a] and van Heel[b] are from the same study but of different cohorts, as is the case for the Hunt group (Hunt[a], Hunt[b], and Hunt[c]).

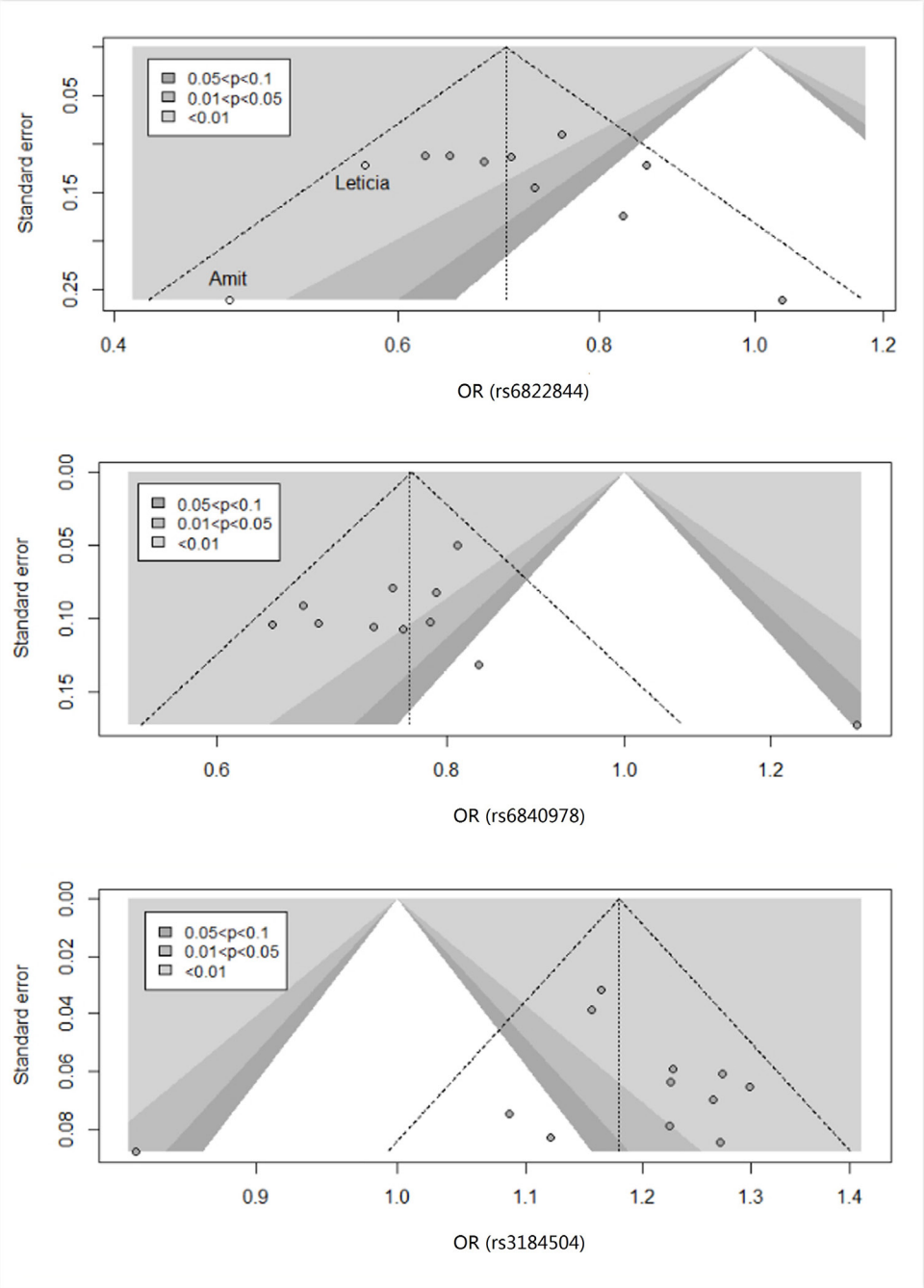

Figure 3. Contour-enhanced funnel plot of *IL2*/*IL21* and *SH2B3* genes with celiac disease (CD) [silvery gray, gray, dark gray, and white areas represent different statistical significances (P < 0.01, 0.01 < P < 0.05, 0.05 < P < 0.1, and P > 0.1, respectively)].

Table 4. Genotype frequencies for the *IL2/IL21* rs6840978 polymorphism and genotype effects of studies included in the meta-analysis.

Author	Total number		Case genotype			Control genotype			T allele prevalence	T *vs* C		TT *vs* CC		CT *vs* CC		HWE
	Case	Control	TT	CT	CC	TT	CT	CC		OR	95%CI	OR	95%CI	OR	95%CI	
van Heel[a]	508	929	12	131	365	44	318	567	0.219	0.644	0.525 0.789	0.424	0.221 0.813	0.640	0.502 0.815	0.945
van Heel[b]	483	560	18	150	315	32	204	324	0.240	0.758	0.614 0.936	0.579	0.318 1.052	0.756	0.582 0.983	0.988
Hunt[a]	719	1561	22	206	491	69	519	973	0.210	0.790	0.672 0.928	0.632	0.386 1.033	0.787	0.648 0.955	0.984
Hunt[b]	416	957	16	131	269	54	346	557	0.237	0.784	0.641 0.958	0.614	0.345 1.092	0.784	0.612 1.005	0.978
Hunt[c]	508	888	13	137	358	43	303	542	0.219	0.681	0.557 0.834	0.458	0.243 0.863	0.685	0.537 0.872	0.938
Sperandeo	289	445	2	65	222	0	82	363	0.092	1.336	0.952 1.875	8.169	0.390 170.926	1.296	0.899 1.869	0.040
Dubois[a]	737	2596	19	200	518	109	845	1642	0.205	0.748	0.641 0.873	0.553	0.336 0.908	0.750	0.625 0.901	0.983
Dubois[b]	1849	4936	54	523	1272	202	1593	3141	0.202	0.811	0.735 0.896	0.660	0.486 0.898	0.811	0.720 0.913	0.999
Dubois[c]	647	1829	6	114	527	30	411	1388	0.129	0.730	0.593 0.898	0.527	0.218 1.273	0.731	0.580 0.920	0.946
Dubois[d]	803	846	19	210	574	39	285	522	0.215	0.669	0.559 0.799	0.443	0.253 0.776	0.670	0.541 0.830	0.990
Dubois[e]	497	543	7	103	387	10	130	403	0.139	0.832	0.643 1.078	0.729	0.275 1.934	0.825	0.615 1.107	0.897
Overall OR										0.764	0.707 0.825	0.579	0.489 0.685	0.767	0.720 0.817	

HWE = Hardy-Weinberg equilibrium; van Heel[a] and van Heel[b] are from the same study but of different cohorts, as is the case for the Hunt group (Hunt[a], Hunt[b], and Hunt[c]) and the Dubois group (Dubois[a], Dubois[b], Dubois[c], Dubois[d], and Dubois[e]).

There were no "missing" studies; however, the study by Sperandeo et al. (2011) was far from the others in the contour-enhanced funnel plot (Figure 3) and we considered this to be a potential cause of heterogeneity. When we excluded this study, I^2 was reduced from 47 to 0% and the pooled OR was 0.76 (95%CI = 0.72-0.80). The Egger linear regression test indicated that no publication bias existed (P = 0.784).

The λ value was 0.50 (95%CI = 0.30-0.82), suggesting that a co-dominant (TT *vs* CC; TC *vs* CC) genetic model was most suitable (Table 4). Both OR_1 (TT *vs* CC, P = 0.786, I^2 = 0%) and OR_2 (TC *vs* CC, P = 0.180, I^2 = 28%) were homogeneous and the pooled OR_1 (0.58; 95%CI = 0.49-0.68) and OR_2 (0.77; 95%CI = 0.72-0.82) values were estimated using the fixed-effects model. The results indicated that individuals with TT and TC genotypes had 42 and 23% reduced risks of developing CD, respectively, compared with those carrying the CC genotype.

Association between the *SH2B3* rs3184504 polymorphism and CD risk

Twelve sub-studies, with 11,163 cases and 26,311 controls, explored the *SH2B3* rs3184504 polymorphism and CD risk (Table 5). The pooled frequency of the minor A allele was 51% (95%CI = 49-53%) in the CD group (I^2 =85%, P < 0.001) and 47% (95%CI = 45-49%) in the control group (I^2 = 91%, P < 0.001), according to the random-effects model. The pooled OR (A *vs* G) was 1.18 (P < 0.001, 95%CI = 1.12-1.24, Figure 2C) with large heterogeneity (P = 0.008, I^2 = 57%), and the PAR% for the minor A allele was 4.84%, suggesting that Caucasians carrying the A allele in *SH2B3* rs3184504 had an increased risk of developing CD.

There were no "missing" studies, but we noticed that the Italian sample collection of the study from Dubois et al. (2010) was far from the others in the contour-enhanced funnel plot (Figure 3), and we considered that this population might be a potential cause of heterogeneity. When we excluded this sample collection, I^2 was reduced from 57 to 0% and the pooled OR was 1.19 (95%CI = 1.16-1.23). The Egger linear regression test indicated that no publication bias existed (P = 0.857).

The λ value was 0.53 (95%CI = 0.27-0.86), suggesting that a co-dominant (AA *vs* GG; AG *vs* GG) genetic model was at play (Table 5). The OR_1 value (AA *vs* GG, P = 0.009, I^2 = 56%) was large and heterogeneous using the random-effects model and the OR_2 (AG *vs* GG, P = 0.617, I^2 = 0%) was homogeneous using the fixed-effects model. The pooled OR_1 and OR_2 values were 1.39 (95%CI = 1.25-1.54) and 1.18 (95%CI = 1.11-1.26), respectively, which suggests that individuals with AA and AG genotypes have 39 and 18% higher risk of CD, respectively, compared with those carrying the GG genotype.

DISCUSSION

Several studies have considered the rs6822844 and rs6840978 polymorphisms of *IL2*/*IL21* and the rs3184504 polymorphism of *SH2B3* as candidates for susceptibility to CD (van Heel et al., 2007; Hunt et al., 2008; Smyth et al., 2008; Coenen et al., 2009; Romanos et al., 2009; Dubois et al., 2010; Maiti et al., 2010; Plaza-Izurieta et al., 2011; Sperandeo et al., 2011). However, there is little consensus amongst the results owing to the small sample sizes and the varying population characteristics. We performed a meta-analysis of nine case-control studies, including a total of 12,986 CD patients and 28,733 controls, to assess the effects of three polymorphisms (rs6822844 and rs6840978 in *IL2*/*IL21*, and rs3184504 in *SH2B3*) on CD risk. For the *IL2*/*IL21* gene, the minor T allele in rs6822844 and rs6840978 might be considered a potentially protective factor for CD. For the *SH2B3* gene, the minor A allele in rs3184504 increases the risk of developing CD by 18%.

Table 5. Genotype frequencies for the *SH2B3* rs3184504 polymorphism and genotype effects of studies included in the meta-analysis.

Author	Total number		Case genotype			Control genotype			A allele prevalence	A *vs* G			AA *vs* GG			AG *vs* GG			HWE
	Case	Control	AA	AG	GG	AA	AG	GG		OR	95%CI		OR	95%CI		OR	95%CI		
Coenen	795	1683	215	397	183	356	836	491	0.460	1.273	1.130	1.434	1.620	1.275	2.059	1.274	1.035	1.568	0.997
Hunt[a]	719	1561	196	359	164	347	778	436	0.471	1.225	1.081	1.389	1.502	1.169	1.930	1.227	0.986	1.527	0.998
Hunt[b]	416	957	106	208	102	217	478	262	0.476	1.120	0.952	1.318	1.255	0.906	1.738	1.118	0.844	1.480	0.971
Hunt[c]	508	888	142	253	113	203	443	242	0.478	1.224	1.049	1.428	1.498	1.099	2.042	1.223	0.932	1.605	0.992
Plaza-Izurieta	1094	540	226	543	325	102	265	173	0.434	1.087	0.938	1.258	1.179	0.876	1.589	1.091	0.862	1.381	0.977
Smyth	2560	9339	700	1277	583	2197	4665	2477	0.485	1.164	1.094	1.238	1.354	1.196	1.532	1.163	1.043	1.298	0.995
Romanos	538	593	165	266	107	145	296	152	0.494	1.271	1.077	1.500	1.617	1.159	2.255	1.277	0.948	1.719	0.970
Dubois[a]	737	2596	212	367	158	612	1297	687	0.486	1.227	1.093	1.378	1.506	1.193	1.901	1.230	0.999	1.515	0.997
Dubois[b]	1849	4936	505	923	421	1169	2466	1301	0.487	1.155	1.071	1.246	1.335	1.147	1.553	1.157	1.012	1.322	0.995
Dubois[c]	647	1829	138	322	187	290	876	663	0.398	1.299	1.144	1.477	1.687	1.301	2.187	1.303	1.060	1.602	0.982
Dubois[d]	803	846	220	401	182	183	420	242	0.465	1.264	1.103	1.450	1.599	1.215	2.104	1.270	1.003	1.607	0.976
Dubois[e]	497	543	104	247	146	139	272	132	0.506	0.823	0.692	0.977	0.676	0.478	0.956	0.821	0.613	1.099	0.963
Overall OR										1.179	1.118	1.243	1.389	1.250	1.544	1.180	1.115	1.248	

HWE = Hardy-Weinberg equilibrium; Hunt[a], Hunt[b], and Hunt[c] are from the same study but of different cohorts, as is the case for the Dubois group (Dubois[a], Dubois[b], Dubois[c], Dubois[d], and Dubois[e]).

The exact role played by rs6822844 G>T and rs6840978 C>T in CD is unclear, but both SNPs, located in the inter-gene region of *IL2* and *IL21*, are convincing candidates for CD pathogenesis (Romanos et al., 2009). Both IL2 and IL21 cytokines are T-cell-derived and stimulate T-cell maturation and proliferation (Adamovic et al., 2008). More importantly, IL21 cooperates with IL2 to promote IFN-γ synthesis, thereby amplifying T helper cell type 1 (Th1) responses, which leads to enterocyte apoptosis by the Fas/Fas ligand (FasL) system, or interleukin 15 (IL-15)-induced perforin-granzyme and NFG2D-MIC signaling pathways (Kasaian et al., 2002). IL2 is critical in cellular activation, and primary and secondary T-cell responses (Bachmann and Oxenius, 2007). In addition, it promotes the proliferation of T-cells, B-cells, and natural killer cells (Adamovic et al., 2008), and may, therefore, promote both the Th-1 and Th-2 cell responses, which induce the apoptotic death of enterocytes and/or enterocyte cytoskeleton changes. Moreover, IL21 can prolong chronic inflammation and favor tissue damage by promoting the recruitment of immune cells, the increase of autoreactive T cells, and the synthesis of extracellular matrix metalloproteinases (Fina et al., 2008). The SNPs rs6822844 and rs6840978 map to a non-coding region upstream of *IL2* and downstream of *IL21*, and previous studies have shown that they may modulate *IL2* and *IL21* expression (Maiti et al., 2010; Warren et al., 2011). In our meta-analysis, the minor rs6822844[T] and rs6840978[T] alleles had significantly lower frequencies in CD patients compared with the controls, which adds credence to the pathogenetic role of *IL2*/*IL21* in CD.

rs3184504, located in exon 3 of the *SH2B3* gene, is a non-synonymous R262W SNP. However, its specific effects on CD have yet to be established. The *SH2B3* gene is mainly expressed in monocytes and dendritic cells, and is a good candidate for CD (Su et al., 2004). SH2B3 is thought to negatively regulate lymphopoiesis and early hematopoiesis (Takaki, 2008), and may, therefore, reduce the Th-1 cell response in the context of CD etiology. SH2B3-deficiency leads to enhanced production of B-cells, revealing a negative regulatory function of SH2B3 in cytokine signaling and Th-2 cell response (Takaki, 2008). Given the reported association of the rs3184504[T] allele with higher expression of *SH2B3* in CD patients (Plaza-Izurieta et al., 2011), our meta-analysis showed that the minor rs3184504[A] allele was associated with an increased risk of CD, adding credence to the protective effect of *SH2B3* in CD.

The genetic associations between *IL2*/*IL21* and *SH2B3* and CD reveal the role of T-cell-mediated signaling in the pathogenesis of CD. The innate molecules within CD may in some sense explicate the link between mucosal immunity state and infections in predisposition to CD (Stene et al., 2006). A better understanding of the pathogenesis of CD could propel the development of pharmacologic agents that would effectually regulate the immune mechanism (Levy et al., 2014).

There were several limitations to our meta-analysis. First, the selection criteria for the controls were unclear, which might have led to inaccurate estimations. Second, heterogeneity existed in two SNPs (rs6840978 T *vs* C and rs3184504 A *vs* G), which may have influenced the pooled results. We evaluated the sources of heterogeneity by omitting the most isolated study in the contour-enhanced funnel plots. For example, in the analysis of the *SH2B3* polymorphism and CD risk, when we excluded the Italian population from the study by Dubois et al. (2010), the heterogeneity disappeared. Therefore, we think that the Italian population may have caused the heterogeneity. However, we were unable to further identify the exact sources of heterogeneity because limited relevant data were provided by the included studies. Third, the number of included studies for each gene (six studies on rs6822844, four studies on rs6840978, and six studies on rs3184504) was low, but the relatively large sample size (12,986 CD patients and 28,733 controls) improved the reliability of the results. Finally, we only included articles written in English or Chinese, so we may have missed some relevant studies published in other languages.

In conclusion, our meta-analysis results provide a strong argument for conducting further studies to better understand the roles of *IL2*/*IL21* and *SH2B3* in CD pathogenesis. The exact roles played by re6822844, rs6840978, and rs3184504 in CD development require further investigation.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research financially supported in part by the National Natural Science Foundation of China (Grant #30901249; #81101267; and #81541070), the Guangdong Natural Science Foundation (Grant #10151063201000036; #S2011010002526), the Guangdong Province Medical Research Foundation (Grant #A2014374 and #A2015310), and by a Project from Jinan university (Grant #21612426, #21615427). We would like to thank Ammarin Thakkinstian from the Section for Clinical Epidemiology and Biostatistics, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Thailand; John R. Thompson from the Department of Health Sciences, University of Leicester, UK; and Sasivimol Rattanasiri from the Section for Clinical Epidemiology and Biostatistics, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Thailand for providing statistical computing algorithm support.

Supplementary material

REFERENCES

Adamovic S, Amundsen SS, Lie BA, Gudjónsdóttir AH, et al. (2008). Association study of IL2/IL21 and FcgRIIa: significant association with the IL2/IL21 region in Scandinavian coeliac disease families. *Genes Immun.* 9: 364-367.

Bachmann MF and Oxenius A (2007). Interleukin 2: from immunostimulation to immunoregulation and back again. *EMBO Rep.* 8: 1142-1148.

Caruso R, Fina D, Peluso I, Stolfi C, et al. (2007). A functional role for interleukin-21 in promoting the synthesis of the T-cell chemoattractant, MIP-3alpha, by gut epithelial cells. *Gastroenterology* 132: 166-175.

Catassi C, Gatti S and Fasano A (2014). The new epidemiology of celiac disease. *J. Pediatr. Gastroenterol. Nutr.* 59: S7-S9.

Coenen MJ, Trynka G, Heskamp S, Franke B, et al. (2009). Common and different genetic background for rheumatoid arthritis and coeliac disease. *Hum. Mol. Genet.* 18: 4195-4203.

Di Sabatino A and Corazza GR (2009). Coeliac disease. *Lancet* 373: 1480-1493.

Dubois PC, Trynka G, Franke L, Hunt KA, et al. (2010). Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet.* 42: 295-302.

Fina D, Sarra M, Caruso R, Del Vecchio Blanco G, et al. (2008). Interleukin 21 contributes to the mucosal T helper cell type 1 response in coeliac disease. *Gut* 57: 887-892.

Fitau J, Boulday G, Coulon F, Quillard T, et al. (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.

Higgins JP, Thompson SG, Deeks JJ and Altman DG (2003). Measuring inconsistency in meta-analyses. *BMJ* 327: 557-560.

Hunt KA, Zhernakova A, Turner G, Heap GA, et al. (2008). Newly identified genetic risk variants for celiac disease related to the immune response. *Nat. Genet.* 40: 395-402.

Kasaian MT, Whitters MJ, Carter LL, Lowe LD, et al. (2002). IL-21 limits NK cell responses and promotes antigen-specific T cell activation: a mediator of the transition from innate to adaptive immunity. *Immunity* 16: 559-569.

King C, Ilic A, Koelsch K and Sarvetnick N (2004). Homeostatic expansion of T cells during immune insufficiency generates autoimmunity. *Cell* 117: 265-277.

Levy J, Bernstein L and Silber N (2014). Celiac disease: an immune dysregulation syndrome. *Curr. Probl. Pediatr. Adolesc.*

Health Care 44: 324-327.

Maiti AK, Kim-Howard X, Viswanathan P, Guillén L, et al. (2010). Confirmation of an association between rs6822844 at the Il2-Il21 region and multiple autoimmune diseases: evidence of a general susceptibility locus. *Arthritis Rheum.* 62: 323-329.

Minelli C, Thompson JR, Abrams KR, Thakkinstian A, et al. (2005). The choice of a genetic model in the meta-analysis of molecular association studies. *Int. J. Epidemiol.* 34: 1319-1328.

Monteleone G, Caruso R, Fina D, Peluso I, et al. (2006). Control of matrix metalloproteinase production in human intestinal fibroblasts by interleukin 21. *Gut* 55: 1774-1780.

Peters JL, Sutton AJ, Jones DR, Abrams KR, et al. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J. Clin. Epidemiol.* 61: 991-996.

Plaza-Izurieta L, Castellanos-Rubio A, Irastorza I, Fernández-Jimenez N, et al. (2011). Revisiting genome wide association studies (GWAS) in coeliac disease: replication study in Spanish population and expression analysis of candidate genes. *J. Med. Genet.* 48: 493-496.

Romanos J, Barisani D, Trynka G, Zhernakova A, et al. (2009). Six new coeliac disease loci replicated in an Italian population confirm association with coeliac disease. *J. Med. Genet.* 46: 60-63.

Rossman MD, Thompson B, Frederick M, Maliarik M, et al. (2003). HLA-DRB1*1101: a significant risk factor for sarcoidosis in blacks and whites. *Am. J. Hum. Genet.* 73: 720-735.

Smyth DJ, Plagnol V, Walker NM, Cooper JD, et al. (2008). Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N. Engl. J. Med.* 359: 2767-2777.

Sollid LM and Lie BA (2005). Celiac disease genetics: current concepts and practical applications. *Clin. Gastroenterol. Hepatol.* 3: 843-851.

Sperandeo MP, Tosco A, Izzo V, Tucci F, et al. (2011). Potential celiac patients: a model of celiac disease pathogenesis. *PLoS One* 6: e21281.

Stene LC, Honeyman MC, Hoffenberg EJ, Haas JE, et al. (2006). Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. *Am. J. Gastroenterol.* 101: 2333-2340.

Su AI, Wiltshire T, Batalov S, Lapp H, et al. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. U. S. A.* 101: 6062-6067.

Takaki S (2008). Sh2b3/Lnk family adaptor proteins in the regulation of lymphohematopoiesis. *Nihon Rinsho Meneki Gakkai Kaishi.* 31: 440-447.

Thakkinstian A, McKay GJ, McEvoy M, Chakravarthy U, et al. (2011). Systematic review and meta-analysis of the association between complement component 3 and age-related macular degeneration: a HuGE review and meta-analysis. *Am. J. Epidemiol.* 173: 1365-1379.

van Heel DA, Franke L, Hunt KA, Gwilliam R, 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.

Warren RB, Smith RL, Flynn E, Bowes J, et al. (2011). A systematic investigation of confirmed autoimmune loci in early-onset psoriasis reveals an association with IL2/IL21. *Br. J. Dermatol.* 164: 660-664.