Identification of Genetic Loci Associated With *Helicobacter pylori* Serologic Status

Julia Mayerle, MD
Caroline M. den Hoed, MD
Claudia Schurmann, MSc
Lisette Stolk, PhD
Georg Homuth, PhD
Marjolein J. Peters, MD, PhD
Lisette G. Capelle, MD
Kathrin Zimmermann, MD
Fernando Rivadeneira, MD, PhD
Sybille Gruska, PhD
Henry Völzke, MD
Annemarie C. de Vries, PhD
Uwe Völker, PhD
Alexander Teumer, PhD
Joyce B. J. van Meurs, PhD
Ivo Steinmetz, MD
Matthias Nauck, MD
Florian Ernst, PhD
Frank-Ulrich Weiss, PhD
Albert Hofman, MD, PhD
Martin Zenker, MD
Heyo K. Kroemer, PhD
Holger Prokisch, PhD
Andre G. Uitterlinden, MD, PhD
Markus M. Lerch, MD, FRCP
Ernst J. Kuipers, MD, PhD

Importance *Helicobacter pylori* is a major cause of gastritis and gastroduodenal ulcer disease and can cause cancer. *H pylori* prevalence is as high as 90% in some developing countries but 10% of a given population is never colonized, regardless of exposure. Genetic factors are hypothesized to confer *H pylori* susceptibility.

Objective To identify genetic loci associated with *H pylori* seroprevalence in 2 independent population-based cohorts and to determine their putative pathophysiological role by whole-blood RNA gene expression profiling.

Design, Setting, and Participants Two independent genome-wide association studies (GWASs) and a subsequent meta-analysis were conducted for anti-*H pylori* IgG serology in the Study of Health in Pomerania (SHIP) (recruitment, 1997-2001 [n = 3830]) as well as the Rotterdam Study (RS-I) (recruitment, 1990-1993) and RS-II (recruitment, 2000-2001 [n = 7108]) populations. Whole-blood RNA gene expression profiles were analyzed in RS-III (recruitment, 2006-2008 [n = 762]) and SHIP-TREND (recruitment, 2008-2012 [n = 991]), and fecal *H pylori* antigen in SHIP-TREND (n = 961).

Main Outcomes and Measures *H pylori* seroprevalence.

Results Of 10 938 participants, 6160 (56.3%) were seropositive for *H pylori*. GWASs identified the toll-like receptor (*TLR*) locus (*4p14*; top-ranked single-nucleotide polymorphism (SNP), rs10004195; $P = 1.4 \times 10^{-18}$; odds ratio, 0.70 [95% CI, 0.65 to 0.76]) and the *FCGR2A* locus (*1q23.3*; top-ranked SNP, rs368433; $P = 2.1 \times 10^{-8}$; odds ratio, 0.73 [95% CI, 0.65 to 0.81]) as associated with *H pylori* seroprevalence. Among the 3 *TLR* genes at *4p14*, only *TLR1* was differentially expressed per copy number of the minor rs10004195-A allele ($\beta = -0.23$ [95% CI, −0.34 to −0.11]; $P = 2.1 \times 10^{-4}$). Individuals with high fecal *H pylori* antigen titers (optical density >1) also exhibited the highest 25% of *TLR1* expression levels ($P = .01$ by χ^2 test). Furthermore, TLR1 exhibited an Asn248Ser substitution in the extracellular domain strongly linked to the rs10004195 SNP.

Conclusions and Relevance GWAS meta-analysis identified an association between *TLR1* and *H pylori* seroprevalence, a finding that requires replication in nonwhite populations. If confirmed, genetic variations in *TLR1* may help explain some of the observed variation in individual risk for *H pylori* infection.

JAMA. 2013;309(18):1912-1920 www.jamo.com

Author Affiliations: Department of Medicine A (Drs Mayerle, Gruska, Weiss, and Lerch), Interfaculty Institute for Genetics and Functional Genomics (Drs Homuth, Völker, Teumer, and Ernst and Ms Schurmann), Friedrich Loeffler Institute of Medical Microbiology (Drs Zimmermann and Steinmetz), Institutes for Community Medicine (Dr Völzke) and Clinical Chemistry and Laboratory Medicine (Dr Nauck), and Department of Pharmacology, Center of Drug Absorption and Transport (CDAT) (Dr Kroemer), University Medicine Greifswald, Greifswald, Germany; Departments of Gastroenterology and Hepatology (Drs den Hoed, Capelle, de Vries, and Kuipers), Internal Medicine (Drs Stolk, Peters, Rivadeneira, van Meurs, Uitterlinden, and Kuipers), and Epidemiology (Drs Hofman and Uitterlinden), Erasmus University Medical Center, Rotterdam, the Netherlands; Netherlands Consortium for Healthy Aging, Rotterdam, Leiden, the Netherlands (Drs Stolk, Peters, Rivadeneira, van Meurs, Hofman, and Uitterlinden); Institute of Human Genetics, Otto-von-Guericke University, Magdeburg, Germany (Dr Zenker); Helmholtz Zentrum München, German Research Center for Environmental Health, Institut für Humangenetik, Neuherberg, Germany (Dr Prokisch); and Technische Universität München, Institut für Humangenetik, Munich, Germany (Dr Prokisch).
Corresponding Author: Markus M. Lerch, MD, FRCP, Department of Medicine A, University Medicine Greifswald, Ferdinand-Sauerbruchstrasse, 17475 Greifswald, Germany (lerch@uni-greifswald.de).

THE GRAM-NEGATIVE PATHOGEN *Helicobacter pylori* is specifically adapted to colonize the mucus layer covering the gastric mucosa, with little invasion of the gastric glands.[1,2] It is the major cause

For editorial comment see p 1939.

of gastritis (80%) and gastroduodenal ulcer disease (15%-20%) and the only bacterial pathogen believed to cause cancer (IARC Working Group 1994).[3,4] Prevalence of infection with *H pylori* varies from less than 10% in asymptomatic children in Western countries to approximately 90% in some developing countries. Most infections occur during childhood, whereas the *H pylori* status of adults remains stable.[5,6] Furthermore, although there are wide interindividual variations in the level of gastritis as well as in the inflammatory response to *H pylori*, the intraindividual gastritis pattern is constant over time.[7]

Approximately 5% to 10% of a population is never infected with *H pylori,* even in the presence of high exposure rates.[8] A contribution of genetic factors to *H pylori* susceptibility is supported by differences in *H pylori* susceptibility between African Americans and US residents of European ancestry after adjusting for socioeconomic status, age, and living conditions.[9] Significantly higher concordance for *H pylori* infection in monozygotic compared with dizygotic twins, or for household members who are siblings rather than unrelated persons,[10] also argues for a genetic influence, with a heritability estimate in twins of 57%.[10]

The key pathophysiological event in *H pylori* infection is the initiation of an inflammatory response. This response is triggered by bacterial membrane components, namely, lipopolysaccharides and lipid A, as well as cytotoxins and *H pylori* urease activity.[11] So far, few candidate gene studies have analyzed human host genetic factors for susceptibility to *H pylori* infection and persistence.[12-14] Several studies have demonstrated that genetic variations, eg, in *IL1B*, modulate the susceptibility for gastric cancer among *H pylori*–infected individuals.[15,16]

This genome-wide association study (GWAS) for determinants of *H pylori* seroprevalence was conducted in 2 large population-based cohorts, the Study of Health in Pomerania (SHIP) and the Rotterdam Study (RS-I and RS-II). Subsequent whole-blood transcriptome analyses were conducted in the independent SHIP-TREND and RS-III populations.

METHODS

Study Cohorts

The SHIP study consists of 2 independent prospectively collected population-based cohorts in Northeastern Germany, SHIP and SHIP-TREND. The study design of SHIP has been previously described in detail.[17] The first patient for the SHIP study was recruited in October 1997 and the last in May 2001. SHIP-TREND is an additional independent cohort from the same region, with individuals newly recruited between September 2008 and summer 2012; for details of SHIP-TREND, see the eAppendix available at http://www.jama.com.

The SHIP study has 2 main objectives: to assess prevalence and incidence of common risk factors, subclinical disorders, and clinical diseases; and to investigate the complex associations among risk factors, subclinical disorders, and clinical diseases. A particular characteristic of SHIP is that it does not specifically address a single selected disease; rather, it attempts to describe health-related conditions with the widest focus possible.

The Rotterdam Study is a large, population-based prospective study of elderly individuals of European ancestry consisting of 3 cohorts (RS-I, RS-II, RS-III) of individuals residing in a suburb of Rotterdam, the Netherlands, and has been described in detail.[18-20] The study targets cardiovascular, endocrine, hepatic, neurologic, ophthalmic, psychiatric, and respiratory diseases. Baseline recruitment and measurements for the RS-I study were obtained between 1990 and 1993. A second cohort, RS-II, was established in 2000-2001. The third cohort, RS-III, started in 2006, with recruitment ending in December 2008.

Data from SHIP, RS-I, and RS-II were used for the GWAS; data from SHIP-TREND and RS-III were used for the transcriptome analysis.

Written informed consent was obtained from all participants, and the medical ethics committee of the Erasmus Medical Center Rotterdam and University Medicine Greifswald approved the study.

Phenotype Determination: Seroprevalence and Bacterial Load

Anti–*H pylori* serum IgG antibody titers were measured using commercial enzyme immunoassays (Pyloriset EIA-G III ELISA; Orion). Seroprevalence, an indicator for current or previous infection, was defined as an anti–*H pylori* IgG titer equal to or greater than 20 U/mL, according to the manufacturer's recommendation.[21] In comparison with culture or CLO (*Campylobacter*-like organism) testing (rapid urease activity testing), using this cutoff value should detect *H pylori* infection with a sensitivity of 97.8%, a specificity of 58.0%, and an accuracy of 78.7%. The positive predictive value for the Pyloriset EIA-G III ELISA immunoassay is reported as 71.5% and the negative predictive value as 96.2%.[22] Individuals with the lowest 75% of the IgG titer distribution comprised the control group. Infection was defined in accordance with international convention when *H pylori* was detected by fecal *H pylori* antigen testing.[23] A significant correlation between titer levels and actual infection has been reported.[22] To investigate the association of gene expression levels and the fecal *H pylori* antigen titer, individuals with high bacterial load (based on fecal *H pylori* antigen titer, optical density [OD]>1; see below) were studied to determine if they also exhibited the highest 25% of gene expression levels of the respective *4p14*-region genes.

The *H pylori* antigen ELISA kit (Immunodiagnostics) was used to detect *H pylori* antigen in stool. One hundred mg of feces was stored at −20°C before analysis. According to the manufacturer's instructions, all participants with an OD greater than or equal to 0.025 at 450 nm are positive for *H pylori* infection. Sensitivity for this test is reported as 97.7% and specificity as

96.3%. For the quantitative correlation between fecal *H pylori* antigen and bacterial load, a Pearson correlation coefficient of 0.222 is reported in the literature.[22]

Genotyping and Imputation

Genotyping of the SHIP probands using the Affymetrix SNP 6.0 array was performed as described previously.[24] Genotyping of the SHIP-TREND participants was performed using the Illumina HumanOmni2.5-Quad array, of RS-I participants using the Illumina 550K (V.3) single and duo arrays, and of RS-II participants using the Illumina 550K (V.3) duo and Illumina 610K Quad arrays, following manufacturer's protocols. All data sets were imputed to the HapMap v22 CEU reference panel (≈2.5 million single-nucleotide polymorphisms [SNPs]) for the meta-analysis. The discovery stage of the GWAS was independently performed in SHIP and RS-I and RS-II. SHIP and RS-I and RS-II served as replication cohorts for each other. For the meta-analysis the 2 data sets were combined. For details, see Study Population and Genotyping in the eAppendix.

Whole-Blood Transcriptome Analysis

For SHIP-TREND and RS-III, whole blood was collected in PAXgene tubes (BD). Total RNA in SHIP-TREND was prepared using a QIA cube device in combination with the Blood miRNA Kit (Qiagen), according to manufacturer's protocols. Subsequent RNA sample processing and hybridization with Illumina HumanHT-12 v3 Expression BeadChips was performed as described by the manufacturer (Illumina) at the Helmholtz Zentrum München. The SHIP-TREND expression data set is available at the GEO (Gene Expression Omnibus) public repository under accession number GSE36382. For details, see the eAppendix.

In RS-III, RNA was amplified and labeled (Ambion TotalPrep RNA) and hybridized with the Illumina HumanHT-12 v4 Expression BeadChips as described by the manufacturer's protocol. The RS-III expression data set is available at the GEO public repository under accession number GSE33828. For details see the eAppendix.

Statistical Analyses

The case and control groups were defined according to their anti–*H pylori* IgG-titer as described above. Detailed information on the study cohort characteristics are provided in **Table 1**. Genome-wide association analysis for SHIP was performed with Quicktest (http://toby.freeshell.org/software/quicktest.shtml) using a logistic regression model with adjustment for sex and age. For RS-I and RS-II, sex- and age-adjusted GWASs were performed using MACH2DAT (http://www.sph.umich.edu/csg/abecasis/MaCH/) implemented in GRIMP.[25] Only SNPs with a minor allele frequency (MAF) greater than 1% as well as those available in all 3 cohorts were considered for further analyses.

Meta-analyses were conducted by an inverse-variance weighted fixed-effects model using METAL (www.sph.umich.edu/csg/abecasis/metal).[26] The random-effects model was

Table 1. Characteristics of Study of Health in Pomerania (SHIP) and Rotterdam Study (RS) Population-Based Cohorts

	No. (%)				
	SHIP		RS		
Characteristic	SHIP	SHIP-TREND	RS-I	RS-II	RS-III
Total sample size, No.	3830	1001	4542	2566	762
Age (range), y	50 (20-81)	50 (20-81)	69 (55-99)	65 (55-95)	60 (46-89)
Women	1957 (51.1)	561 (56.0)	2454 (54.0)	1401 (54.6)	410 (53.8)
Total No. of samples with measured anti–*Helicobacter pylori* IgG titer	3830	988	4542	2566	NA
No. of samples with anti–*H pylori* IgG titer >20 U/mL	2269 (59.2)	551 (55.8)	2695 (59.3)	1196 (46.6)	NA
No. of samples with the highest 25% IgG titer values (cutoff in U/mL)	958 (124.5)	247 (82.8)	1136 (136.8)	642 (88.9)	NA
IgG titer values, mean (SD), [median], U/mL					
Group with the lowest 75% values	32.0 (29.3) [18.1]	19.0 (15.2) [12.4]	32.9 (31.6) [18.0]	20.6 (17.1) [12.9]	NA
Group with the highest 25% values	268.6 (120.0) [230.9]	262.2 (140.1) [227.3]	459.7 (589.0) [320.7]	309.5 (255.0) [233.4]	NA
Total No. of samples with measured *H pylori* antigen	NA	961	NA	NA	NA
No. of samples with *H pylori* antigen OD>0.025	NA	325 (33.8)	NA	NA	NA
No. of samples with *H pylori* antigen OD>1	NA	139 (14.5)	NA	NA	NA

Abbreviations: NA, not available; OD, optical density; RS, Rotterdam Study; SHIP, Study of Health in Pomerania.

calculated in R using the package metafor.[27] Genomic control was applied to the individual studies' GWAS results ($\lambda_{RS\text{-}I}=1.008$, $\lambda_{RS\text{-}II}=1.012$, $\lambda_{SHIP}=1.000$). Because the genomic inflation factor of the combined meta-analysis was below 1 ($\lambda=0.999$), no genomic control was applied for those results.

To assess whether there were multiple independently associated SNPs within the loci found in the combined meta-analysis, a clumping analysis was performed using PLINK[28] (1-megabase distance, $r^2>0.01$, HapMap R28 CEU genotype data set), but no additional hits were found. $P=5\times10^{-8}$ was used as the threshold for genome-wide significance, and 2-sided significance testing was performed.

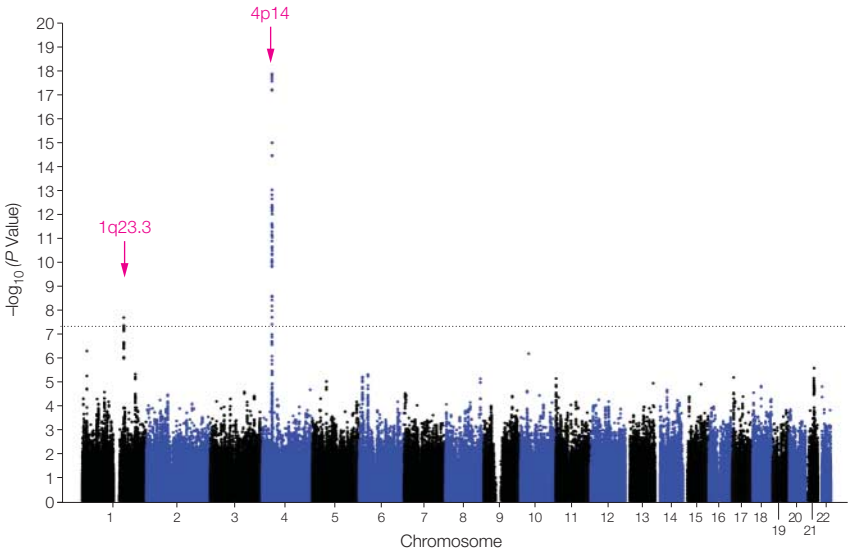
Expression analyses in SHIP-TREND were performed using quantile-normalized and $\log_2$-transformed gene expression data. Expression quantitative trait loci (eQTL) analyses in SHIP-TREND and RS-III used quantile-normalized, $\log_2$-transformed, probe-centered, and standardized gene expression data. To reduce the number of false-positive and false-negative eQTL association results, in a first step a principal component analysis of the matrix of gene expression data was performed.[29] In a second step, the gene expression profiles of the individuals were adjusted for the first 50 principal components,[29] which represent the largest factors related to variation of gene expression levels. For details see eAppendix, eTable 1A and B, and eFigure 1A-F.

RESULTS

Prevalence and Frequency of *H pylori* Seroprevalence

Some level of seroprevalence throughout the study was detected in 6160 of 10 938 participants (56.3%) (SHIP, RS-I, and RS-II) (Table 1). Based on the predefined phenotypic seroprevalence in the top 25% of the study population, a total of 2623 cases (25%) and 7862 controls (75%) were used for GWAS meta-analysis (SHIP, RS-I, and RS-II) (Table 1). To increase specificity and reduce the number of false-positive *H pylori* infections in the case group of the GWAS, the cutoff was set to the upper 25% of the IgG titer distribution of the corresponding cohort (124.5 U/mL for SHIP, 136.8 U/mL for RS-I, and 88.9 U/mL for RS-II).

Figure 1. Genome-wide Association Studies Meta-analysis: Statistical Significance of Association for All Single-Nucleotide Polymorphisms (SNPs) With Minor Allele Frequency Greater Than 1%



Manhattan plot showing significance of association for all single-nucleotide polymorphisms (SNPs) with a minor allele frequency greater than 1% in the meta-analysis with the anti–*Helicobacter pylori* IgG phenotype defined as a dichotomous variable comparing samples with the highest 25% IgG titers vs those with the lowest 75% IgG titers. SNPs available in all 3 cohorts are plotted on the x-axis according to their chromosomal position against the association with the phenotype (shown as $-\log_{10} P$ value) on the y-axis. The dotted line indicates the threshold for genome-wide statistical significance ($P=5\times10^{-8}$).

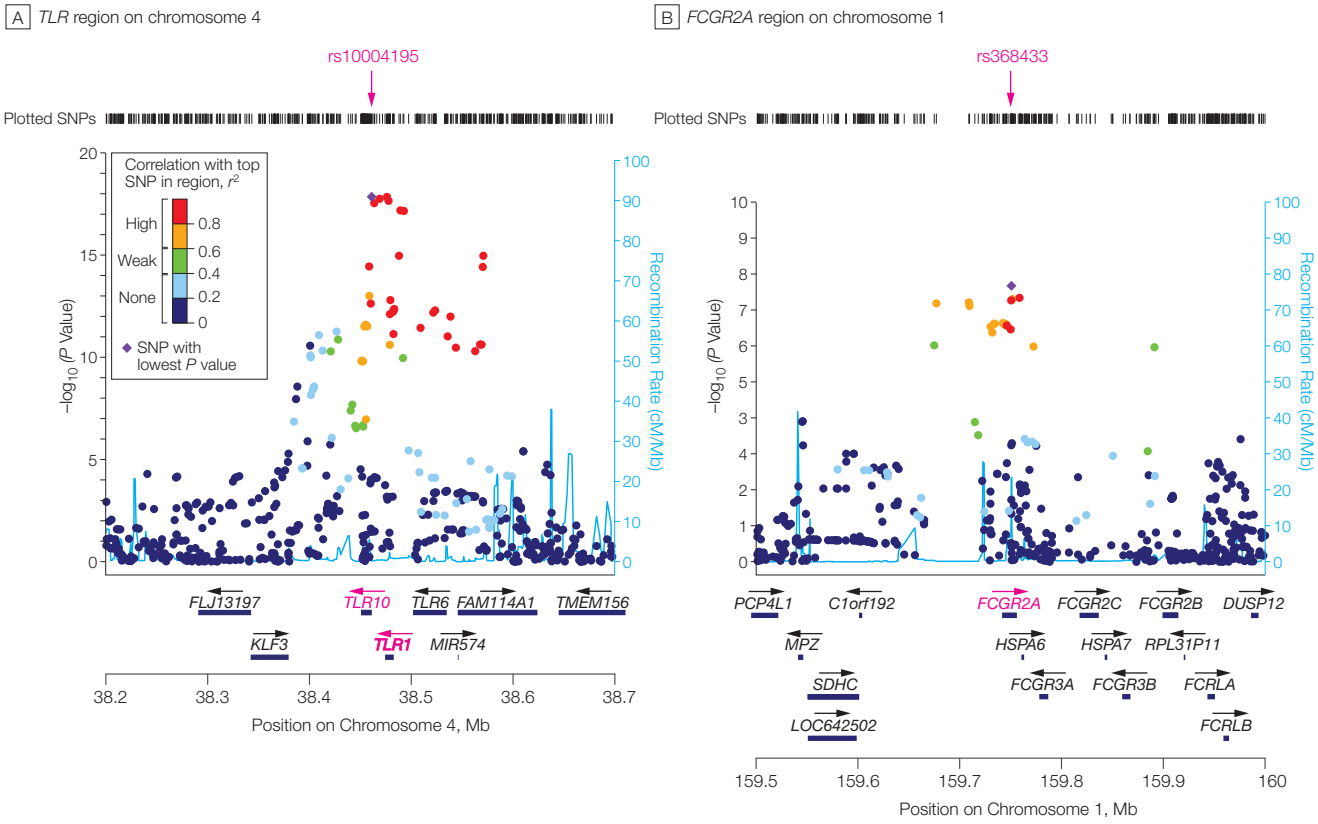
GWAS Meta-analysis

Two genome-wide significant loci were identified (FIGURE 1) by combining the data from RS-I, RS-II, and SHIP (n=10 485, for which *H pylori* serology and genotyping data were available) using a fixed-effects meta-analysis model. The *TLR* locus on *4p14* exhibited the lowest *P* value (FIGURE 2A), with rs10004195 as the lead SNP (odds ratio [OR] for the minor allele, 0.70 [95% CI, 0.65-0.76]; $P=1.42\times10^{-18}$; MAF=24.7%), closely followed by rs4833095 (OR for the minor allele, 0.70 [95% CI, 0.65-0.76]; $P=1.43\times10^{-18}$; MAF=24.9%) (TABLE 2). The second genome-wide significant locus was located on chromosome *1q23.3*, with $P=2.1\times10^{-8}$ for the lead SNP, rs368433 (OR for the minor allele, 0.73 [95% CI, 0.65-0.81]; MAF=16%) (Figure 2B). This SNP is located in an intron of *FCGR2A* encoding the Fcγ receptor 2a. The I^2 heterogeneity measures for the top-ranked SNPs[30] (57.3% for rs10004195; 22.6% for rs368433) indicated high and low between-study heterogeneity, respectively. To take into account that the effect of the SNP might differ between the studies as implicated by the high heterogeneity, a random-effects meta-analysis was applied to the genome-wide significant findings to ensure that the combined effect of the SNP was not the result of a large effect in a single cohort. For rs10004195 the association *P* value increased to 6.5×10^{-9} using this model, and the combined effect estimate was nearly the same (OR, 0.69 [95% CI, 0.61-0.79]), indicating that the observed association was not completely driven by a single study. For SNP rs368433, the association *P* value and

combined effect size were essentially unchanged compared with the fixed-effects meta-analysis (OR, 0.73 [95% CI, 0.65-0.82]; $P = 9.2 \times 10^{-8}$), but the *P* value no longer met the genome-wide significance threshold (eTable 2).

In the complete meta-analysis comprising SHIP, RS-I, and RS-II (n = 10 938), the numbers of cases (samples with the highest 25% IgG titer values) were 985 for SHIP, 1136 for RS-I, and 642 for RS-II (n = 2763), whereas the numbers of con-

Figure 2. Genome-wide Association Studies Meta-analysis: Loci Associated With Anti–*Helicobacter* pylori IgG Titers on a Genome-Wide Level of Significance



Regional plots of the 2 loci associated with anti–*H pylori* IgG titers on a genome-wide level of significance. The y-axis on the left indicates the $-\log_{10}$ *P* value for the association with the anti–*H pylori* IgG phenotype. Single-nucleotide polymorphisms (SNPs) available in all 3 cohorts are plotted on the x-axis according to their chromosomal position against the association with the phenotype on the y-axis. Shown are the *TLR* (left) and *FCGR2A* (right) regions: purple diamonds indicate the top-ranked SNPs of the respective regions exhibiting the lowest *P* value for association with the phenotype. The blue y-axes on the right of each plot indicate the estimated recombination rates (based on HapMap phase II); the bottom of each panel shows the respective annotated genes at the locus and their transcriptional direction. cM indicates centimorgans; Mb, megabases.

Table 2. Top-Ranked SNPs From the Genome-Wide Association Studies Meta-analysis of the Defined Anti–*Helicobacter pylori* IgG Titer Phenotype in RS-I, RS-II, and SHIP[a]

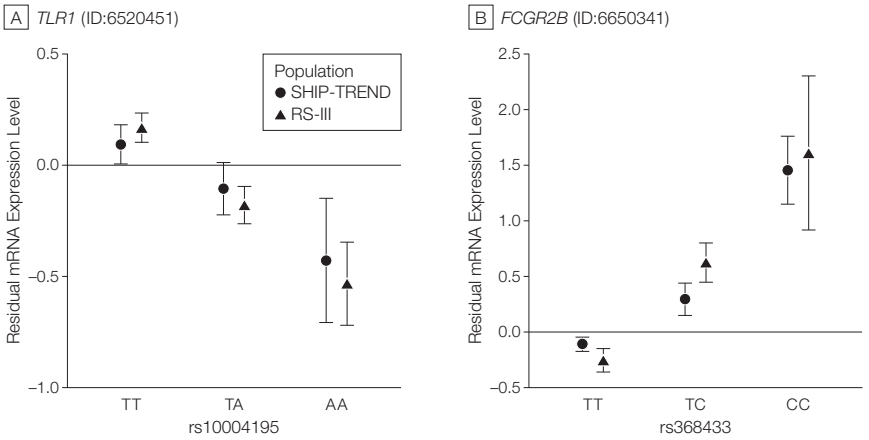
Chromosome	Analysis	MAF	OR (95% CI)	*P* Value
4p14 (top-ranked SNP: rs10004195)[b]	Meta-analysis	0.25	0.70 (0.65-0.76)	1.4×10^{-18}
	SHIP	0.22	0.62 (0.54-0.72)	5.0×10^{-11}
	RS-I and RS-II	0.26	0.74 (0.67-0.81)	7.4×10^{-10}
1q23.3 (top-ranked SNP: rs368433)[c]	Meta-analysis	0.16	0.73 (0.65-0.81)	2.1×10^{-8}
	SHIP	0.15	0.67 (0.55-0.81)	2.4×10^{-5}
	RS-I and RS-II	0.16	0.76 (0.66-0.87)	1.3×10^{-4}

Abbreviations: MAF, minor allele frequency; OR, odds ratio; RS, Rotterdam Study; SHIP, Study of Health in Pomerania; SNP, single-nucleotide polymorphism.
[a] Top-ranked SNP refers to the locus' SNP with the smallest *P* value.
[b] Genes: *TLR10, TLR1, TLR6, FAM114A1*. Major/minor allele: A/T.
[c] Genes: *FCGR2A, HSPA6*. Major/minor allele: C/T.

trols (samples with the lowest 75% IgG titer values) in each population were 2845 for SHIP, 3406 for RS-I, and 1924 for RS-II (n=8175). The MAF of the protective *TLR* locus lead SNP rs10004195 was 0.22 in SHIP, 0.26 in RS-I/II, and 0.25 in the meta-analysis. The MAF of the protective *FCGR2A* locus lead SNP rs368433 was 0.15 in SHIP, 0.16 in RS-I/II, and 0.16 in the meta-analysis.

From a pathophysiological point of view, associations of other *TLR* genes with *H pylori* seroprevalence might be of relevance. Therefore, the meta-analysis data were used to search for associated SNPs (genotyped or imputed in all 3 cohorts, MAF >0.01) within or in ±100-kilobase windows upstream and downstream of *TLR9*, *TLR2*, or *TLR4*. The smallest *P* value observed was for *TLR2* (rs11736691; OR, 0.79 [95% CI, 0.65-0.96]; *P*=.02); this did not meet genome-wide significance. A complete list of these results is provided in eTable 3; a list of all genome-wide significant SNPs is provided in eTable 4.

Figure 3. *Cis*-eQTL Analysis of the Significant Genome-wide Association Study Top-Ranked SNPs in Whole Blood



A, Residual (ie, after adjustment for technical effects and potential confounders; overall mean centered) mean $\log_2$-transformed gene expression levels corresponding to gene-specific mRNA levels in whole blood and 95% CIs (error bars) per genotype group of *TLR1* (ID:6520451), adjusted for the first 50 eigenvectors with respect to rs10004195 for Study of Health in Pomerania–TREND (SHIP-TREND) (genotype distribution: TT=614, TA=315, AA=47) and Rotterdam Study III (RS-III) (genotype distribution: TT=439, TA=269, AA=54). B, Residual (ie, after adjustment for technical effects and potential confounders, and overall mean-centered) mean $\log_2$-transformed gene expression levels corresponding to gene-specific mRNA levels in whole blood and 95% CIs (error bars) per genotype group of *FCGR2B* (ID:6650341), adjusted for the first 50 eigenvectors with respect to rs368433 for both SHIP-TREND (genotype distribution: TT=765, TC=191, CC=20) and RS-III (genotype distribution: TT=551, TC=198, CC=13). For *cis*-eQTL (expression quantitative trait loci) analysis, all genes in a ±250-kilobase region of the 2 top-ranked single-nucleotide polymorphisms (SNPs) from the genome-wide association study were chosen.

Putative Functional Context of the Genome-wide Significant SNPs

The protein-coding sequences of the genes closest to the top-ranked SNPs were analyzed for nonsynonymous SNPs in linkage disequilibrium with $r^2 \geq 0.6$. Using the 1000 Genomes database in the SNAP SNP Proxy search tool,[31] 2 nonsynonymous SNPs were identified in *TLR1* and 1 SNP in *TLR10* in linkage disequilibrium with the top-ranked SNP (rs10004195). In *TLR1*, the previously mentioned rs4833095 (r^2=1.0) and rs5743618 (r^2=0.95) are nonsynonymous SNPs. Whereas rs4833095 causes the amino acid substitution Asn248Ser, rs5743618 results in the Ser602Ile substitution. In *TLR10*, rs4129009 (r^2=0.77) is located close to the 3' end of the single protein–coding exon of the gene and corresponds to the Ile775Val substitution. This TLR10 amino acid position is localized within the intracellular TIR (toll/interleukin-1 receptor) domain that participates in the transduction of extracellular signaling.[32] For the top-ranked SNP at the 1q23.3 locus (rs368433; Figure 2B), no nonsynonymous SNPs in linkage disequilibrium were found.

Association of Anti–*H pylori* IgG Titers and *H pylori* Stool Antigen in SHIP-TREND

To explore the relationship between the *TLR* locus and *H pylori* infection in more detail, additional analyses in SHIP-TREND were carried out. *H pylori* stool antigen levels and measured anti–*H pylori* IgG titers were found to have a significant positive correlation (Spearman ρ=0.59, $P=2\times10^{-90}$).

Cis-eQTL Analysis of the Significant GWAS Hits in Whole Blood

Integrative analysis of blood expression profiles with genome-wide SNP data in 1763 participants from SHIP-TREND and RS-III were used to investigate the potential causal relationship between gene(s) located at the loci shown to be associated with *H pylori* susceptibility. Variation at rs10004195, the top-ranked SNP at the *TLR* locus (*4p14*), was significantly correlated with the mRNA levels of *TLR1* (ID: 6520451; $P=2.1\times10^{-4}$ for SHIP-TREND and $P=3.2\times10^{-17}$ for RS-III) (FIGURE 3A, eTable 5, eFigure 2). Variation at the linked nonsynonymous SNP rs4833095 was also significantly associated with *TLR1* (ID:6520451) mRNA expression levels. Among the three *TLR* genes only *TLR1* was differentially expressed. *TLR1* was differentially expressed per copy number of the minor rs10004195-A allele (β=−0.23 [95% CI, −0.34 to −0.11]; $P=2.1\times10^{-4}$). The eQTL-results for rs10004195 and all available *TLR* probes are shown in eTable 6.

Variation at rs368433, the top-ranked SNP of the second genome-wide significant locus (*1q23.3*), was significantly associated with expression levels of *FCGR2A*, *FCGR2B,* and *HSPA6* represented by 7 probes (Figure 3B, eTable 6, eFigure 3A and B). Significant associations were found for

FCGR2B (ID:6650341; $P=9.5\times10^{-19}$ for SHIP-TREND and $P=2.5\times10^{-20}$ for RS-III). *FCGR2B* expression levels increased in participants carrying 1 or more minor alleles in both populations, whereas expression levels of *HSPA6* and *FCGR2A* decreased with each minor allele.

Relationship Between Bacterial Load, *TLR* Gene Expression, and Genotype

TLR6, *TLR1*, or *TLR10* mRNA amounts were analyzed to determine if expression levels increased in parallel with bacterial load as determined by *H pylori* stool antigen testing. To minimize false-positive results, only samples with an OD greater than 1 (high bacterial load as defined in previously published studies[22]) were used to study correlations between *H pylori* fecal load and gene expression levels. Individuals with high fecal *H pylori* antigen titers (OD>1) also exhibited the highest 25% of *TLR1* expression levels ($P=.01$ by χ^2 test; *TLR1* gene probe ID: 6520451). This result was independent of the rs10004195 SNP genotype. This significantly increased expression implied a specific and genotype-independent transcriptional upregulation of *TLR1*, but not of *TLR6* or *TLR10*, in the presence of *H pylori* (*TLR6*: $P=.26$ [TLR6 gene probe ID: 5340427] and $P=.80$ [TLR6 gene probe ID:2600735]; *TLR10*: $P=.05$ [TLR10 gene probe ID:4480543], $P=.72$ [TLR10 gene probe ID:620441], and $P=.95$ [TLR10 gene probe ID:380639] [all results by χ^2 test]). For *TLR6* and *TLR10*, more than 1 gene-specific probe was present on the HumanHT-12 Expression BeadChips. These findings identify TLR1 to be the receptor most likely involved in recognition of *H pylori*.

No significant association between fecal *H pylori* antigen and the expression of *FCGR2A* and *2B* was detected in whole blood.

DISCUSSION

This GWAS on *H pylori* seroprevalence, conducted in 2 population-based cohorts including a total of 10 938 participants, identified 2 genome-wide significant loci located at *4p14* and *1q23.3* associated with *H pylori* seroprevalence. These findings were further explored in 1763 additional study participants in whom gene expression levels were studied in whole blood and 961 participants in whom bacterial load was measured in stool samples.

The *4p14* region encodes the *TLR1*, *TLR6*, and *TLR10* genes; TLR1 was identified as the receptor most likely causatively associated with *H pylori* seroprevalence. This conclusion is biologically plausible because TLRs are known to be essential for protective immunity against infection. Murine models have previously suggested that the TLRs 2, 9, and 8 could be involved in the recognition of *H pylori* by dendritic cells, which traverse epithelial tight junctions in the intestine to sample luminal bacteria.[33,34] Among these, TLR2 represents the only cell surface receptor/ligand system that can cause a pronounced anti-inflammatory signature.[33]

TLR1, on the other hand, is known to represent one of the coreceptors of TLR2. Both proteins can form a heterodimer that recognizes triacylated lipopeptides from the cell envelope of gram-negative bacteria.[35,36] The fine structure of *H pylori* lipid A can consist of triacylated lipopeptides, which would make it an ideal ligand for TLR2-TLR1 binding.[37] However, this conclusion remains speculative, because information was not available regarding the identities of the different *H pylori* strains in the study cohorts. Recent studies showed the association of rs5743618, a nonsynonymous SNP in linkage with rs10004195 identified in this study, with *Chlamydia trachomatis* infection[38] and leprosy,[39] indicating the involvement of TLR1 in bacterial infection.

Additional studies have shown that the specific *H pylori*–induced dendritic cell cytokine profile polarizes the balance between mucosal T helper 1 (T_H1) and T helper 17 (T_H17) cells on one side and regulatory T helper (T_{reg}) cells on the other side, toward a T_{reg}-biased response. This would cause suppression of the *H pylori*–specific T_H1/T_H17–dependent responses and correlated with a higher degree of *H pylori* infection.[40,41] These results demonstrated that *H pylori*, by inducing a T_{reg}-skewed response via active tolerogenic programming of dendritic cells, limits the host's ability to eradicate the pathogen and can result in persistence of *H pylori* infection.[41]

In light of animal data[33,34] and new data generated by this study, one hypothesis is that the protective minor allele–associated haplotype might confer less effective anti-inflammatory TLR1-TLR2 signaling. The single nonsynonymous SNP rs4833095 in *TLR1* identified as associated with *H pylori* seroprevalence represents a putative candidate for such a predicted modification of TLR1-TLR2 function, because the position of the Asn248Ser amino acid substitution is located within the immediate vicinity of the TLR1 ligand binding and dimerization site.[32] This hypothesis could be tested in a recently generated knockout animal.[42] In the present study, the allele associated with higher anti–*H pylori* IgG titers was also associated with higher *TLR1* expression. Whether the nonsynonymous SNP rs4833095 is associated with *H pylori* treatment failure or recurrence requires further investigation, because the cross-sectional design of this population-based study does not allow an evaluation of these parameters.

In the GWAS meta-analysis, the *1q23.3* region was also significantly associated with the *H pylori* phenotype. The minor allele of the top-ranked SNP, rs368433, was associated with low anti–*H pylori* IgG titers, decreased blood levels of *HSPA6* and *FCGR2A*, and increased expression levels of *FCGR2B*. The *FCGR2B*-encoded Fcy receptor IIB also seems plausibly related to seroprevalence, because genetic variations affecting the receptor's affinity for IgG subclass 2 (IgG2) have been reported.[43] Furthermore, it has been observed that neutrophils from individu-

als who are homozygous for one of these SNPs can phagocytose IgG2-opsonized bacteria more effectively.[44] On the other hand, the increased expression of FCGR2B on the surface of immune cells might simply result in more pronounced cellular titration of IgGs and, consequently, decreased titers of soluble IgG. Additional studies are required to validate the association and, if validated, to explore the underlying biological mechanisms related to the observed effects.

At this time, the clinical implications of the current findings are unknown. Based on these data, genetic testing to evaluate *H pylori* susceptibility outside of research projects would be premature. This study has several limitations. First, the study was conducted among participants phenotyped only for seroprevalence and not symptomatic *H pylori* infection. In agreement with previous studies, this study shows a quantitative correlation for *H pylori* serology and *H pylori* fecal antigen; however, the correlation is weak (Spearman $\rho=0.59$). Therefore, additional studies are needed to determine if the observed associations are present in populations phenotyped for clinically significant infection. Second, the validity of the current results are restricted to individuals of European ancestry. Additional GWASs of cohorts with non-European ancestry will be required to determine the extent to which the results can be generalized to other ethnic groups.

Third, it is also not clear if the results would be reproducible in a cohort exposed to a significantly higher pathogen pressure. A significant proportion of the study population had no definitive serological evidence of *H pylori* exposure, nor were study population–specific data available for *H pylori* exposure rates. It is technically possible that the observed associations are related to effective exposure to *H pylori*, rather than serologic conversion. Fourth, the *FCGR2A/B* locus showed genome-wide significant association in the meta-analysis under a fixed-effects model only and requires replication in independent, ethnically diverse, cohorts.

GWAS meta-analysis identified an association between *TLR1* and *H pylori* seroprevalence, a finding that requires replication in other independent populations. If confirmed, genetic variations in *TLR1* may help explain some of the observed variation in individual risk for *H pylori* infection.

Author Contributions: Drs Lerch and Kuipers had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Drs Mayerle, den Hoed, and Stolk and Ms Schurmann contributed equally to the study. Drs Lerch and Kuipers share senior authorship.

Study concept and design: Mayerle, den Hoed, Homuth, Capelle, Völzke, de Vries, Völker, Teumer, Hofman, Kroemer, Uitterlinden, Lerch, Kuipers.

Acquisition of data: Mayerle, den Hoed, Capelle, Zimmermann, Rivadeneira, Gruska, Völzke, van Meurs, Nauck, Ernst, Prokisch, Uitterlinden, Kuipers.

Analysis and interpretation of data: Mayerle, den Hoed, Schurmann, Stolk, Homuth, Peters, Capelle, Zimmermann, Völker, Teumer, van Meurs, Steinmetz, Weiss, Zenker, Prokisch, Uitterlinden, Lerch, Kuipers.

Drafting of the manuscript: Mayerle, den Hoed, Schurmann, Stolk, Homuth, Peters, Capelle, Gruska, Uitterlinden, Lerch, Kuipers.

Critical revision of the manuscript for important intellectual content: den Hoed, Schurmann, Stolk, Peters, Zimmermann, Rivadeneira, Völzke, de Vries, Völker, Teumer, van Meurs, Steinmetz, Nauck, Ernst, Weiss, Hofman, Zenker, Kroemer, Prokisch, Uitterlinden, Lerch, Kuipers.

Statistical analysis: den Hoed, Schurmann, Stolk, Peters, Capelle, Teumer, Ernst.

Obtained funding: Mayerle, Homuth, Völzke, Völker, Hofman, Kroemer, Uitterlinden, Lerch, Kuipers.

Administrative, technical, or material support: Mayerle, den Hoed, Homuth, Capelle, Zimmermann, Völzke, de Vries, Völker, Steinmetz, Nauck, Prokisch, Uitterlinden.

Study supervision: Mayerle, Homuth, de Vries, Völker, van Meurs, Nauck, Ernst, Weiss, Hofman, Uitterlinden, Lerch, Kuipers.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Völzke reported receiving a grant from the Federal Ministry of Education and Research (BMBF). Dr Völker reported receiving a grant from BMBF. Dr Nauck reported receiving grants from BMBF (01ZZ9603), the Ministry of Cultural Affairs (01ZZ0103), and the Social Ministry of the Federal State of Mecklenberg-West Pomerania (01ZZ0403). Dr Weiss reported receiving a grant pending from EFRE/TBI. Dr Mayerle reported providing expert testimony on behalf of AstraZeneca and Novartis; and receiving payment for lectures from the Falk Foundation.

Funding/Support: This study was funded by BMBF, the Center of Knowledge Interchange program of the Siemens AG, and the Caché-Campus program of the InterSystems GmbH. The Rotterdam Study is supported by the Netherlands Organisation of Scientific Research (NWO) Investments, the Research Institute for Diseases in the Elderly, and the Netherlands Genomics Initiative/NWO. Further funding and support is described in the eAppendix.

Role of the Sponsors: The funders had no role in the design and conduct of the study; the collection, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.

Online-Only Material: The eAppendix, eTables 1-6, and eFigures 1-3 are available at http://www.jama.com.

REFERENCES

1. Amieva MR, El-Omar EM. Host-bacterial interactions in *Helicobacter pylori* infection. *Gastroenterology*. 2008;134(1):306-323.

2. Suerbaum S, Michetti P. *Helicobacter pylori* infection. *N Engl J Med*. 2002;347(15):1175-1186.

3. Kusters JG, van Vliet AH, Kuipers EJ. Pathogenesis of *Helicobacter pylori* infection. *Clin Microbiol Rev*. 2006;19(3):449-490.

4. Montecucco C, Rappuoli R. Living dangerously: how *Helicobacter pylori* survives in the human stomach. *Nat Rev Mol Cell Biol*. 2001;2(6):457-466.

5. Kuipers EJ, Peña AS, van Kamp G, et al. Seroconversion for *Helicobacter pylori*. *Lancet*. 1993;342(8867):328-331.

6. Malaty HM, El-Kasabany A, Graham DY, et al. Age at acquisition of *Helicobacter pylori* infection: a follow-up study from infancy to adulthood. *Lancet*. 2002;359(9310):931-935.

7. Kuipers EJ, Uyterlinde AM, Peña AS, et al. Long-term sequelae of *Helicobacter pylori* gastritis. *Lancet*. 1995;345(8964):1525-1528.

8. Bardhan PK. Epidemiological features of *Helicobacter pylori* infection in developing countries. *Clin Infect Dis*. 1997;25(5):973-978.

9. Graham DY, Malaty HM, Evans DG, Evans DJ Jr, Klein PD, Adam E. Epidemiology of *Helicobacter pylori* in an asymptomatic population in the United States: effect of age, race, and socioeconomic status. *Gastroenterology*. 1991;100(6):1495-1501.

10. Malaty HM, Engstrand L, Pedersen NL, Graham DY. *Helicobacter pylori* infection: genetic and environmental influences: a study of twins. *Ann Intern Med*. 1994;120(12):982-986.

11. Falush D, Wirth T, Linz B, et al. Traces of human migrations in *Helicobacter pylori* populations. *Science*. 2003;299(5612):1582-1585.

12. Queiroz DM, Saraiva IE, Rocha GA, et al. IL2-330G polymorphic allele is associated with decreased risk of *Helicobacter pylori* infection in adulthood. *Microbes Infect*. 2009;11(12):980-987.

13. Zambon CF, Basso D, Navaglia F, et al. Pro- and anti-inflammatory cytokines gene polymorphisms and *Helicobacter pylori* infection: interactions influence outcome. *Cytokine*. 2005;29(4):141-152.

14. Zheng Z, Jia Y, Hou L, et al. Genetic variation in a4GnT in relation to *Helicobacter pylori* serology and gastric cancer risk. *Helicobacter*. 2009;14(5):120-125.

15. El-Omar EM, Carrington M, Chow WH, et al. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. *Nature*. 2000;404(6776):398-402.

16. El-Omar EM, Rabkin CS, Gammon MD, et al. Increased risk of noncardia gastric cancer associated with proinflammatory cytokine gene polymorphisms. *Gastroenterology*. 2003;124(5):1193-1201.

17. Völzke H, Alte D, Schmidt CO, et al. Cohort profile: the Study of Health In Pomerania. *Int J Epidemiol*. 2011;40(2):294-307.

18. Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol*. 2009;24(9):553-572.

19. Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: objectives and design update. *Eur J Epidemiol*. 2007;22(11):819-829.

20. Hofman A, van Duijn CM, Franco OH, et al. The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol*. 2011;26(8):657-686.

21. Rehnberg-Laiho L, Salomaa A, Rautelin H, Koskela P, Sarna S, Kosunen TU. Accelerated decline in *Helicobacter pylori* seroprevalence rate during the screen and treat project in Vammala, Finland, as demon-

strated in 29- to 45-year-old pregnant women. *APMIS*. 2004;112(1):34-38.
22. Hanvivatvong O, Pongpanich A, Thong-Ngam D, Thammacharoenrach N, Kullavanijaya P. Evaluation of commercial immunoassays for detection of antibody against *Helicobacter pylori* in Thai dyspeptic patients. *Clin Diagn Lab Immunol*. 2004;11(3):618-620.
23. Tonkic A, Tonkic M, Lehours P, Mégraud F. Epidemiology and diagnosis of *Helicobacter pylori* infection. *Helicobacter*. 2012;17(suppl 1):1-8.
24. Teumer A, Rawal R, Homuth G, et al. Genome-wide association study identifies four genetic loci associated with thyroid volume and goiter risk. *Am J Hum Genet*. 2011;88(5):664-673.
25. Estrada K, Abuseiris A, Grosveld FG, Uitterlinden AG, Knoch TA, Rivadeneira F. GRIMP: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics*. 2009;25(20):2750-2752.
26. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190-2191.
27. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. 2010;36(3):1-48.
28. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575.
29. Schurmann C, Heim K, Schillert A, et al. Analyzing Illumina gene expression microarray data from different tissues: methodological aspects of data analysis in the metaxpress consortium. *PLoS One*. 2012;7(12):e50938.
30. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557-560.
31. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*. 2008;24(24):2938-2939.
32. Guan Y, Omueti-Ayoade K, Mutha SK, Hergenrother PJ, Tapping RI. Identification of novel synthetic toll-like receptor 2 agonists by high throughput screening. *J Biol Chem*. 2010;285(31):23755-23762.
33. Rad R, Ballhorn W, Voland P, et al. Extracellular and intracellular pattern recognition receptors cooperate in the recognition of *Helicobacter pylori*. *Gastroenterology*. 2009;136(7):2247-2257.
34. Rad R, Brenner L, Krug A, et al. Toll-like receptor–dependent activation of antigen-presenting cells affects adaptive immunity to *Helicobacter pylori*. *Gastroenterology*. 2007;133(1):150-163.
35. Gautam JK, Ashish, Comeau LD, Krueger JK, Smith MF Jr. Structural and functional evidence for the role of the TLR2 DD loop in TLR1/TLR2 heterodimerization and signaling. *J Biol Chem*. 2006;281(40):30132-30142.
36. Jin MS, Kim SE, Heo JY, et al. Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell*. 2007;130(6):1071-1082.
37. Suda Y, Ogawa T, Kashihara W, et al. Chemical structure of lipid A from *Helicobacter pylori* strain 206-1 lipopolysaccharide. *J Biochem*. 1997;121(6):1129-1133.
38. Taylor BD, Darville T, Ferrell RE, Kammerer CM, Ness RB, Haggerty CL. Variants in toll-like receptor 1 and 4 genes are associated with *Chlamydia trachomatis* among women with pelvic inflammatory disease. *J Infect Dis*. 2012;205(4):603-609.
39. Wong SH, Gochhait S, Malhotra D, et al. Leprosy and the adaptation of human toll-like receptor 1. *PLoS Pathog*. 2010;6:e1000979.
40. Kao JY, Zhang M, Miller MJ, et al. *Helicobacter pylori* immune escape is mediated by dendritic cell–induced Treg skewing and Th17 suppression in mice. *Gastroenterology*. 2010;138(3):1046-1054.
41. Zhang JY, Liu T, Guo H, et al. Induction of a Th17 cell response by *Helicobacter pylori* urease subunit B. *Immunobiology*. 2011;216(7):803-810.
42. Miranda-Hernandez S, Gerlach N, Fletcher JM, et al. Role for MyD88, TLR2 and TLR9 but not TLR1, TLR4 or TLR6 in experimental autoimmune encephalomyelitis. *J Immunol*. 2011;187(2):791-804.
43. Lehrnbecher T, Foster CB, Zhu S, et al. Variant genotypes of the low-affinity Fcgamma receptors in two control populations and a review of low-affinity Fcgamma receptor polymorphisms in control and disease populations. *Blood*. 1999;94(12):4220-4232.
44. Salmon JE, Edberg JC, Brogle NL, Kimberly RP. Allelic polymorphisms of human Fc gamma receptor IIA and Fc gamma receptor IIIB: independent mechanisms for differences in human phagocyte function. *J Clin Invest*. 1992;89(4):1274-1281.