

OPEN

Common variation in *FAM155A* is associated with diverticulitis but not diverticulosis

Matthias C. Reichert^{1,9*}, Juozas Kupcinskas^{2,9}, Antje Schulz³, Christoph Schramm⁴, Susanne N. Weber¹, Marcin Krawczyk^{1,5}, Christoph Jüngst^{1,6}, Markus Casper¹, Frank Grünhage¹, Beate Appenrodt¹, Vincent Zimmer¹, Algimantas Tamelis⁷, Jaune I. Lukosiene², Neringa Pauziene⁸, Gediminas Kiudelis², Laimas Jonaitis², Tobias Goesser⁴, Maciej Malinowski³, Matthias Glanemann³, Limas Kupcinskas² & Frank Lammert¹

Colonic diverticulosis is a very common condition. Many patients develop diverticulitis or other complications of diverticular disease. Recent genome-wide association studies (GWAS) consistently identified three major genetic susceptibility factors for both conditions, but did not discriminate diverticulitis and diverticulosis in particular due the limitations of registry-based approaches. Here, we aimed to confirm the role of the identified variants for diverticulosis and diverticulitis, respectively, within a well-phenotyped cohort of patients who underwent colonoscopy. Risk variants rs4662344 in Rho GTPase-activating protein 15 (*ARHGAP15*), rs7609897 in collagen-like tail subunit of asymmetric acetylcholinesterase (*COLO*) and rs67153654 in family with sequence similarity 155 A (*FAM155A*) were genotyped in 1,332 patients. Diverticulosis was assessed by colonoscopy, and diverticulitis by imaging, clinical symptoms and inflammatory markers. Risk of diverticulosis and diverticulitis was analyzed in regression models adjusted for cofactors. Overall, the variant in *FAM155A* was associated with diverticulitis, but not diverticulosis, when controlling for age, BMI, alcohol consumption, and smoking status (OR_{adjusted} 0.49 [95% CI 0.27–0.89], $p = 0.002$). Our results contribute to the assessment specific genetic variants identified in GWAS in the predisposition to the development of diverticulitis in patients with diverticulosis.

Colonic diverticulosis is a widespread gastrointestinal condition described as formation of diverticula, which are sac-like protrusions of mucosa and submucosa through muscularis externa^{1,2}. Although diverticulosis is predominantly asymptomatic, many patients develop symptoms^{3,4}. These manifestations are classified as uncomplicated and complicated diverticular disease (DD), an inflammation of the tissue (diverticulitis) or diverticular bleeding. The prevalence of the disease and its complications is constantly increasing due to the aging populations^{5,6}. Therefore, the identification of underlying pathobiological mechanisms of the disease is relevant for improving everyday clinical practice. It is hypothesized that environmental factors together with structural alteration of the colonic wall and remodeling of the enteric nervous system are all predisposing components in the development of colonic diverticula⁵. In recent years, epidemiological twin data have suggested that genetic factors are another major cofactor in the pathogenesis of DD^{7,8}. Up to date, there has been relatively little research effort to determine these genes and their sequence variants that are associated with the development of DD⁹. Recently an Icelandic study group published a genome-wide association study (GWAS) searching for sequence variants that affect the

¹Department of Medicine II, Saarland University Medical Center, Saarland University, Homburg, Germany.

²Department of Gastroenterology and Institute for Digestive Research, Lithuanian University of Health Sciences, Kaunas, Lithuania. ³Department of General, Visceral, Vascular and Pediatric Surgery, Saarland University Medical Center, Homburg, Germany. ⁴Clinic for Gastroenterology and Hepatology, University Hospital of Cologne, Cologne, Germany.

⁵Laboratory of Metabolic Liver Diseases, Department of General, Transplant and Liver Surgery, Medical University of Warsaw, Warsaw, Poland. ⁶Department of Gastroenterology and Hepatology, University Hospital Zurich, and University of Zurich, Zurich, Switzerland. ⁷Department of Surgery, Lithuanian University of Health Sciences, Kaunas, Lithuania. ⁸Institute of Anatomy, Lithuanian University of Health Sciences, Kaunas, Lithuania.

⁹These authors contributed equally: Matthias C. Reichert and Juozas Kupcinskas. *email: matthias.reichert@uks.eu

risk of developing DD in the Icelandic population, containing a replication cohort of Danish individuals with DD¹⁰. The initial analysis included 15,220 Icelanders who were tested for associations with DD (5,426 cases) and its more severe form diverticulitis (2,764 cases). Subsequently, after applying weighted thresholds¹¹, 16 sequence variants identified in the GWAS were followed up in a DD sample from Denmark with 5,970 cases and 3,020 controls. In the combined analysis of these sample sets, three genetic loci that show genome wide-significance and may be associated with the risk of DD and/or diverticulitis were identified: intronic variants at the *ARHGAP15* (Rho GTPase-activating protein 15), *COLQ* (collagen-like tail subunit of asymmetric acetylcholinesterase) and *FAM155A* (family with sequence similarity 155 A) loci were significantly associated with DD. The second GWAS¹² included 27,444 patients from the European component of the UK Biobank resource and compared them with 382,284 controls. Overall, 154 associated variants were further tested in 31,221 patients from the Michigan Genomics Initiative, finally confirming 42 associated variants including the three previously identified variants¹⁰. Most recently, a third European GWAS¹³ containing 451,099 patients in addition to identification of further loci, also confirmed these three major loci.

Notably, the dissection of the specific phenotypes diverticulosis and diverticulitis was incomplete in the GWAS, since the assessment of diverticulosis or diverticulitis is based on the ICD code. The studies applied the corresponding ICD codes in the ICD10 (K572-K579) and ICD9 (K562.10-13) systems, which also encompasses patients with diverticulitis (as well as diverticular bleeding), among the patients with diverticulosis. Even though ICD codes can identify patients with diverticulosis and DD, they do not discriminate between diverticulosis and diverticulitis¹⁴. A subanalysis of the patients from the Icelandic subcohort of the Icelandic/Danish GWAS included only patients with either surgically treated or complicated diverticulitis¹⁰, whereas mild and often also outpatient-treated cases of diverticulitis were probably mostly missed. This information is not available in the Danish subcohort at all. Additionally, patients with asymptomatic diverticulosis were not analyzed in these GWAS studies. A part of our samples were used in the GWAS from Schafmayer *et al.*¹³, and after adding further samples re-analyzed using additional clinical covariates from the database as outlined below and specifically focusing on patients with diverticulosis no prior diverticulitis as controls and (endpoint diverticulitis) and healthy with no diverticula (endpoint diverticulosis).

In this study, we therefore aimed to evaluate the associations between three major SNPs reported in the Icelandic GWAS applying weighted thresholds and confirmed in the North-American¹² and European^{11,13} GWAS: *ARHGAP15* (rs4662344), *COLQ* (rs7609897) and *FAM155A* (rs67153654). The risk of developing diverticulosis and diverticulitis, respectively, was determined in a Caucasian (german/lithuanian) cohort phenotypically characterized for the specific phenotypes diverticulosis and diverticulitis. Due to the similar genetic background of Germans and Lithuanians¹⁵, a combined analysis was performed.

Patients and Methods

Study population. All patients taking part in the study were recruited at the Department of Medicine II, Saarland University Medical Center, Homburg, the Clinic for Gastroenterology and Hepatology, University Hospital of Cologne, Cologne, Germany, and the Department of Gastroenterology at the Lithuanian University of Health Sciences, Kaunas in Lithuania between 2012 and 2016 from patients referred for colonoscopy. A part of the samples from the Lithuanian cohorts came from our previous studies on colonic diseases and diverticulosis^{16–18} and were also used in a previous GWAS with less clinical information¹³. All patients and controls were of self-reported Caucasian ancestry (including grandparents). Risk factors, epidemiological and baseline data were assessed using a structured interview, performed by a physician assisting the patients with the questionnaires. The presence of diverticula was assessed by colonoscopy in all patients, which is the most widely accepted standard to detect diverticula. Only patients with complete colonoscopy including inspection of the cecum and at least adequate preparation, as assessed by the physician performing colonoscopy, were included in the study. All colonoscopies were performed using digital video endoscopes (high-resolution scopes Olympus CF 160, 180 or 190) by a senior gastroenterologist. Intestinal lavage for endoscopic examination was performed using 2 liters of a solution containing polyethylene glycol. Patients with inherited connective tissue disorders such as Ehlers-Danlos- or Marfan syndrome, non-Caucasian ethnicity or relatives of included patients were also excluded. The diagnosis of diverticulitis was established according to the current classifications for DD^{19,20}. It was based on imaging by either computed tomography and/or ultrasound imaging as well corresponding clinical (pain in the lower left abdomen) and laboratory characteristics (increased serum inflammation markers). Suspected complicated diverticulitis was assessed with computed tomography in all cases. The study protocol was approved by the Research Ethics Committee of the Saarland University (approval 63/11), the Research Ethics Committee of the University of Cologne (approval 16–397) and the Regional Kaunas Ethics Committee (protocol No BE-10-2). The study was performed according to the Declaration of Helsinki. All patients have signed an informed consent form to participate in the study. For the purpose of this study, cases were defined as patients with diverticulosis or diverticulitis, respectively.

DNA extraction and genotyping. Genomic DNA was extracted from peripheral blood mononuclear cells using the DNeasy Blood& Tissue Kit (Qiagen, Hilden, Germany). DNA concentrations were measured using a NanoDrop spectrophotometer. DNA samples were stored at -20°C until analysis. Genotyping of the three genetic polymorphisms (rs4662344, rs7609897 and rs67153654) with Taqman assays was performed in 856 patients with diverticulosis and 479 controls of Caucasian descent in our accredited laboratory (DIN EN ISO 15189) in Homburg by a technician blinded to the phenotype of the patients. The fluorescence data was analyzed with allelic discrimination 7500 software v.2.0.6.

Statistical analysis. Statistical Package for the Social Sciences (SPSS, Version 20, IBM, Munich, Germany) was used for statistical analysis. Power calculations were performed using PS (<http://biostat.mc.vanderbilt.edu/>)

Parameter	Diverticulosis (n = 858)	Controls (n = 474)	Diverticulitis (n = 198)	Total (n = 1332)	P value*	P value**
Age, years	67 (59–74)	57 (46–66)	61 (53–71)	64 (55–72)	<0.001	<0.001
BMI, kg/m ²	27.9 (25.2–31.6)	26.7 (23.7–30.3)	27.1 (24.8–30.5)	27.5 (24.8–31.2)	<0.001	0.20
Gender, men/women	424 (49.4%)	210 (44.3%)	105 (53.0%)	634 (47.6%)	0.08	0.26
Smoking ever, yes/no	300 (35.6%)	168 (36.7%)	86 (43.9%)	468 (36.0%)	0.72	0.006
Alcohol (daily), yes/no	50 (6.0%)	21 (4.6%)	19 (9.9%)	71 (5.5%)	0.31	0.01

Table 1. Baseline data of study population. Values are given as median and interquartile range (IQR), or frequencies and percentages. BMI = body mass index; Significant P values are highlighted in bold. *Diverticulosis versus healthy controls, **Diverticulitis versus Diverticulosis with no prior diverticulitis.

Gene	A _{min} /A _{maj}	MAF (%)	CT (%)	TT (%)	OR	P _{trend}	OR (95% CI)	P _{allelic}	OR (95% CI)	P _{genotypic}
ARHGAP15 (rs4662344)										
Diverticulitis*	T/C	19.5	32.8	3.1	1.20	0.16	1.27 (0.91–1.76)	0.16	1.38 (0.67–2.84)	0.38
Diverticulosis**	T/C	19.5	29.7	4.7	1.30	0.02	1.28 (1.00–1.63)	0.05	1.85 (0.97–3.50)	0.06
Controls _(no diverticulosis)	T/C	15.9	26.3	2.7						
Controls _(no prior diverticulitis)	T/C	18.8	28.8	4.4						
Iceland GWAS ¹⁰ _{Diverticular disease***}	T/C	20.9							1.23 (1.17–1.29)	<0.001
Iceland GWAS ¹⁰ _{Controls***}	T/C	17.7								
MGI GWAS ¹² _{****Diverticular disease}	T/C	19.1/20.0							0.88/0.89	<0.001/0.07
MGI GWAS ¹² _{****Controls}	T/C	17.3/18.0								
European GWAS ¹³ _{Diverticular disease}	T/C	17.9*****								<0.001
European GWAS ¹⁴ _{Controls}	T/C	17.9*****								
FAM155A (rs67153654)										
			TA (%) AA (%)							
Diverticulitis*	A/T	14.1	21.9	3.1	0.68	0.01	0.66 (0.47–0.92)	0.01	0.45 (0.19–1.09)	0.07
Diverticulosis**	A/T	22.8	35.4	5.1	1.01	0.90	1.01 (0.81–1.27)	0.91	1.02 (0.61–1.72)	0.93
Controls (no diverticulosis)	A/T	22.6	35.2	5.1						
Controls (no prior diverticulitis)	A/T	24.3	37.1	5.8						
Iceland GWAS ¹⁰ _{Diverticular disease***}	A/T	17.0							0.87 (0.83–0.91)	<0.001
Iceland GWAS ¹⁰ _{Controls***}	A/T	18.6								
MGI GWAS ¹² _{****Diverticular disease}	A/T	17.0/17.8							0.87/0.90	<0.001/0.024
MGI GWAS ¹² _{****Controls}	A/T	19.0/19.5								
European GWAS ¹³ _{Diverticular disease}	A/T	19.6*****								
European GWAS ¹³ _{Controls}	A/T	19.6*****								

Table 2. Genotypic and allelic frequencies in *ARHGAP15* and *FAM155A* of the combined German/Lithuanian cohort and published data from published GWAS. Genotypic and allelic frequencies of the variants. Values are given as count and percentage. A_{maj} = major allele; A_{min} = minor allele; CI = confidence interval; MAF minor allele frequency; MGI Michigan Genomics Initiative; OR = odds ratio. P Values calculated for *Diverticulitis versus Diverticulosis with no prior diverticulitis; **Diverticulosis versus healthy controls; ***refers to combined Icelandic/Danish sample; ****Data from variant rs6734367 which is in LD ($r^2 = 0.82$) with variant rs4662344 in caucasians³¹, data from UK biobank/MGI; *****Data from the variant rs11619840 which is in LD ($r^2 = 0.93$) with variant rs67153654 in caucasians³¹, data from UK biobank/MGI; *****Data only combined cases/controls available.

wiki/Main/PowerSampleSize) to detect a significantly increased OR of 2 with a power of 80%, based on the corresponding frequencies of the risk alleles in rs4662344, rs7609897 and rs67153654, and assuming type I error rates of 0.05. Quantitative data were expressed as medians and ranges. Comparisons of frequencies of genotypes at the three loci were performed in 3×2 contingency tables listing cases and controls. Genotypic and allelic association tests were performed using χ^2 -square or Fisher's exact tests (<https://ihg.gsf.de/cgi-bin/hw/hwa1.pl>). Due to the few homozygous mutants of the risk alleles we applied a dominant model. Genotype association analysis between SNPs and diverticulosis was performed using multiple logistic regression models adjusted for age, BMI, smoking status, and alcohol consumption, assuming log-additive effects. P-values < 0.05 were considered statistically significant. Results are expressed as odds ratios (OR) and 95% confidence intervals (CI). Pairwise linkage disequilibrium (r^2) was calculated utilizing LDpair²¹ using a caucasian reference population (CEU) of the identified variants in all three GWAS^{10,12,13}.

Results

Patient characteristics. In total, 1,332 patients (634 men, 47.6%) were included. Table 1 summarizes the baseline data of this study cohort. Frequency of diverticulosis in our cohort was similar to prior data²⁰. The median age was 64 years (IQR 55–72). Characteristics of the subgroup of patients with diverticulitis are presented

	Variant present with significant association in GWAS (Reference number)	ARHGAP15		FAM155A				
		rs4662344	rs6734367	rs9520344	rs11619840	rs67153654	rs9555371	rs9520339
ARHGAP15	<u>rs4662344</u>	^{10,13}	0.82	0.001	0.019	0.014	0.009	0.001
	rs6734367	^{12,13}		0.001	0.035	0.028	0.020	0.001
FAM155A	rs9520344	^{12,13}			0.0021	0.0008	0.0043	1.0
	rs11619840	^{12,13}				0.93	0.83	0.0021
	<u>rs67153654</u>	^{10,13}					0.90	0.0008
	rs9555371	¹³						0.0043
	rs9520339	¹³						

Table 3. Pairwise linkage disequilibrium (r^2) calculated using LDpair (Machiela *et al.*³¹) with a caucasian reference population (CEU) of the identified variants in all three GWAS^{10,12,13}. Bold values indicate where r^2 is >0.8 representing strong LD. Variants from the initial GWAS¹⁰ are marked in underline.

Parameter	Adjusted OR* (95% CI)	P value
ARHGAP15 rs4662344:T (CC + TC vs TT)	1.22 (0.93–1.61)	0.15
ARHGAP15 rs4662344:TC (CT vs TT + CC)	1.14 (0.86–1.52)	0.35
ARHGAP15 rs4662344:TT (TT vs TC + CC)	1.89 (0.84–4.29)	0.13
FAM155A rs67153654:T (TA + AA vs TT)	1.09 (0.84–1.00)	0.53
FAM155A rs67153654:AT (TA vs TT + AA)	1.02 (0.78–1.30)	0.89
FAM155A rs67153654:TT (TT vs AT + AA)	1.19 (0.67–2.10)	0.56

Table 4. Multivariate analysis of factors associated with diverticulosis versus controls. CI = confidence interval. OR = odds ratio. *Adjusted for age, BMI, alcohol and smoking status.

Parameter	Adjusted OR* (95% CI)	P value
ARHGAP15 rs4662344:T (CC + TC vs TT)	1.43 (1.00–2.03)	0.05
ARHGAP15 rs4662344:TC (CT vs TT + CC)	1.35 (0.44–1.94)	0.11
ARHGAP15 rs4662344:TT (TT vs TC + CC)	1.40 (0.61–3.21)	0.43
FAM155A rs67153654:T (TA + AA vs TT)	0.70 (0.49–0.99)	0.04
FAM155A rs67153654:AT (TA vs TT + AA)	0.73 (0.51–1.06)	0.10
FAM155A rs67153654:TT (TT vs AT + AA)	0.68 (0.27–1.69)	0.41

Table 5. Multivariate analysis of factors associated with diverticulitis in patients with diverticulosis. CI = confidence interval, OR = odds ratio. *Adjusted for age, BMI, alcohol and smoking status. Significant P values are highlighted in bold.

in Supplementary Table 1. The call-rate for the variations were $>95\%$ for all variants. The genotype frequencies (cut-off $P > 0.05$) were in Hardy-Weinberg equilibrium (HWE) in all controls for the variants in *ARHGAP15* and *FAM155A*. The HWE for the variant in *COLQ* deviated in controls (both in diverticulosis and diverticulitis-analyses $p < 0.001$), and was not included in further analysis. The minor allele frequencies (MAF) (Table 2) were similar to prior data^{10,12,13}. In comparison of patients with diverticulosis and controls, patients with diverticulosis were significantly ($P < 0.001$) older and more obese ($P < 0.001$) than individuals with no diverticulosis. When comparing patients with diverticulitis, with diverticulosis and no prior diverticulitis, patients with diverticulitis were significantly younger ($P < 0.001$), more often smokers ($P = 0.006$), and more frequently current alcohol drinkers ($P = 0.001$). No association was detected for BMI ($P = 0.20$) and gender ($P = 0.26$). Table 3 presents the data on linkage disequilibrium (LD) of all variants identified in the GWAS^{10,12,13}.

Associations of variants and diverticulosis. Table 2 presents the allelic and genotypic frequencies comparing patients with diverticulosis to controls. MAF of the variant in *ARHGAP15* (rs4662344) was increased compared to controls, as described in the GWAS^{10,12,13}. The major (T) allele of rs4662344 in *ARHGAP15* was significantly (OR 1.28; 95% CI 1.00–1.63) associated with diverticulosis. This association did not withstand after adjusting for corresponding environmental cofactors though (OR 1.22; 95% CI = 0.93–1.61) (Table 4). Neither was the MAF of the variant in rs67153654 in *FAM155A* different between cases and controls, nor could an association with diverticulosis be detected (Table 4).

Associations of SNPs and diverticulitis. The minor allele of the variant rs4662344 in *ARHGAP15* was more frequent in diverticulitis cases in comparison to controls with diverticulosis and no prior diverticulitis, as similarly described previously in the GWAS^{10,12,13}. The MAF of the major (A) allele of rs67153654 in *FAM155A* was markedly (OR 0.66; 95% CI 0.47–0.92) reduced in patients with prior diverticulitis compared to controls (Table 2) as also previously described^{10,12,13}. These associations remained significant after adjusting for

environmental cofactors (Table 5). The variant rs4662344 in *ARHGAP15* was borderline significantly (OR 1.43; 95% CI 1.00–2.03; $P = 0.05$) associated with diverticulitis after adjusting for the corresponding cofactors. Even though hampered by small sample size, similar results ($n = 64$) were obtained when analyzing patients with surgically treated diverticulitis (Supplementary Tables 2 and 3).

Discussion

The aim of our present study was to assess the role of genetic variations consistently identified in the three large recent GWAS^{10,12,13} for the specific risks for diverticulosis and diverticulitis, respectively. Our results are in line with previous data concerning the association of diverticulosis with age and BMI as risk factors (diverticulosis)^{22–24}, as well as alcohol consumption^{25,26} and smoking status^{27–29} as risk factors for diverticulitis.

Our major finding is that the rs67153654 risk allele in *FAM155* is significantly associated with diverticulitis after adjusting for cofactors, but not with diverticulosis. The data on the risk variant rs4662344 in *ARHGAP15* was less consistent, it was borderline significantly associated with both diverticulosis and diverticulitis, but for confirmation additional larger studies are warranted. MAF of the analyzed SNPs was similar to previous data^{10,12,13,30}. Even though the variants in *ARHGAP15* and *FAM155A* initially discovered in the Icelandic/Danish GWAS¹⁰ are at partially different genetic positions on the same genes compared to the variants identified in the following GWAS^{12,13}, their LD indicates their common heritability. Therefore, an analysis using the SNPs initially identified in the Icelandic GWAS¹⁰, which applied weighted thresholds, is justified.

All of the analyzed SNPs are located in introns, supporting a molecular mechanism at the level of RNA-expression in the surrounding gene or LD to another, yet unidentified causal variant.

One of the major strengths of our study is the availability of clinical and endoscopic data and covariates, allowing the exact separation of patients with uncomplicated diverticulosis from patients developing diverticulitis. Patients treated for diverticulitis as outpatients were also included in our analysis. Our study adds new insights into the susceptibility for diverticulitis, specifically attributing the association with the risk variant in *FAM155A* (protective effect) to diverticulitis, but not diverticulosis. Our study has certain limitations though, that have to be acknowledged. Due to the retrospective design of the study, we could not investigate the outcomes of diverticulosis and diverticulitis, including DD-associated mortality. Even though a similar genetic background is shared between Germans and Lithuanians¹⁵ our results can not necessarily be transferred to non-Caucasians and have to be validated in other ethnicities ethnicity-specific analysis was not feasible due to sample size. Furthermore, several other variants associated to DD in the prior GWAS were not assessed in our analysis, and could also contribute significantly to the genetic risk to develop diverticulitis in patients with diverticulosis. Ultimately, the development of a genome-wide polygenic score³¹ should be strived for in both clinical entities diverticulosis and diverticulitis. The role of the risk variants in complicated DD could also not be explored, as our sample size for these additional subgroups was too small. Additionally, to the best of our knowledge only one study from 2010 assessing the application of ICD codes for the discrimination of diverticulosis and diverticulitis is available. Therefore, confirmatory investigations are necessary. Further studies are also needed to understand which of the variants in *FAM155A* are the causal mutations. Furthermore, as the function of the involved genes is largely known, further elucidation of the molecular background of *FAM155A* deficiency in the pathogenesis of diverticulitis is warranted.

Conclusions

Our results indicate, that the variant in *FAM155A* is associated with diverticulitis, but not diverticulosis in Caucasians, whereas a risk variant in *ARHGAP15* might be associated with both diverticulosis and diverticulitis. Our results contribute to the assessment of these genetic variants identified in GWAS in the predisposition to the development of diverticulitis in patients with diverticulosis.

Data availability

The datasets generated during and/or analyzed during the current study are available on request as permitted by data protection laws and patients consent.

Received: 24 April 2019; Accepted: 14 January 2020;

Published online: 03 February 2020

References

1. Tursi, A. Diverticulosis today: unfashionable and still under-researched. *Therap. Adv. Gastroenterol.* **9**, 213–218 (2016).
2. Pfützner, R. H. & Kruis, W. Management of diverticular disease. *Nat. Rev. Gastroenterol. Hepatol.* **12**, 629–638 (2015).
3. Strate, L. L. *et al.* Diverticular disease is associated with increased risk of subsequent arterial and venous thromboembolic events. *Clin. Gastroenterol. and Hepatol.* **2014** **12**, 1695–1701 (2014).
4. Sheth, A. A., Longo, W. & Floch, M. H. Diverticular disease and diverticulitis. *Am. J. Gastroenterol.* **103**, 1550–1556 (2008).
5. von Rahden, B. H. & Germer, C. T. Pathogenesis of colonic diverticular disease. *Langenbecks Arch. Surg.* **397**, 1025–1033 (2012).
6. Everhart, J. E. *et al.* Burden of digestive diseases in the United States part II: lower gastrointestinal diseases. *Gastroenterology.* **136**, 741–754 (2009).
7. Strate, L. L. *et al.* Heritability and familial aggregation of diverticular disease: a population-based study of twins and siblings. *Gastroenterology.* **144**, 736–742 (2013).
8. Granlund, J. *et al.* The genetic influence on diverticular disease—a twin study. *Aliment. Pharmacol. Ther.* **35**, 1103–1107 (2012).
9. Reichert, M. C. & Lammert, F. The genetic epidemiology of diverticulosis and diverticular disease: Emerging evidence. *United European Gastroenterol. J.* **3**, 409–418 (2015).
10. Sigurdsson, S. *et al.* Sequence variants in *ARHGAP15*, *COLQ* and *FAM155A* associate with diverticular disease and diverticulitis. *Nat. Commun.* **8**, 15789 (2017).
11. Sveinbjornsson, G. *et al.* Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nat. Genet.* **48**, 314–317 (2016).

12. Macguire, L. H. *et al.* Genomewide association analyses identify 39 new susceptibility loci for diverticular disease. *Nat. Genet.* **50**, 1359–1365 (2018).
13. Schafmayer, C. *et al.* Genome-wide association analysis of diverticular disease points towards neuromuscular, connective tissue and epithelial pathomechanisms. *Gut*. **5**, 854–865 (2019).
14. Erichsen, R. *et al.* Positive predictive values of the International Classification of Disease, 10th edition diagnoses codes for diverticular disease in the Danish National Registry of Patients. *Clin. Exp. Gastroenterol.* **3**, 139–142 (2010).
15. Nelis, M. *et al.* Genetic structure of Europeans: a view from the North-East. *PLoS One*, **4**(5), e5472 (2009).
16. Kupcinskas, J. *et al.* Lack of association between miR-27a, miR-146a, miR-196a-2, miR-492 and miR-608 gene polymorphisms and colorectal cancer. *Sci. Rep.* **4**, 5993 (2014).
17. Kupcinskas, J. *et al.* Common Genetic Variants of PSCA, MUC1 and PLCE1 Genes are not Associated with Colorectal Cancer. *Asian Pac. J. Cancer Prev.* **16**, 6027–6032 (2015).
18. Reichert, M. C. *et al.* A Variant of COL3A1 (rs3134646) Is Associated With Risk of Developing Diverticulosis in White Men. *Dis. Colon. Rectum.* **61**, 604–611 (2018).
19. Stollman, N. *et al.* American Gastroenterological Association Institute Guideline on the Management of Acute Diverticulitis. *Gastroenterology*. **149**, 1944–1949 (2015).
20. Kruis, W. *et al.* Diverticular disease: guidelines of the German Society for Gastroenterology, Digestive and Metabolic Diseases and the German Society for General and Visceral Surgery. *Digestion*. **90**, 190–207 (2014).
21. Machiela, M. J. & Chanock, S. J. LDassoc: an online tool for interactively exploring genome-wide association study results and prioritizing variants for functional investigation. *Bioinformatics*. **33**, 887–888 (2018).
22. Rosemar, A., Angeras, U. & Rosengren, A. Body mass index and diverticular disease: a 28-year follow-up study in men. *Dis. Colon Rectum.* **51**, 450–455 (2008).
23. Mashayekhi, R. *et al.* Obesity, but Not Physical Activity, Is Associated With Higher Prevalence of Asymptomatic Diverticulosis. *Clin. Gastroenterol. and Hepatol.* **16**, 586–587 (2018).
24. Ma, W. *et al.* Association Between Obesity and Weight Change and Risk of Diverticulitis in Women. *Gastroenterology*. **155**, 58–66 (2018).
25. Aldoori, W. H. *et al.* A prospective study of alcohol, smoking, caffeine, and the risk of symptomatic diverticular disease in men. *Ann. Epidemiol.* **5**, 221–228 (1995).
26. Tonnesen, H., Engholm, G. & Moller, H. Association between alcoholism and diverticulitis. *Br. J. Surg.* **86**, 1067–1068 (1999).
27. Hjern, E., Wolk, A. & Hakansson, N. Smoking and the risk of diverticular disease in women. *Br. J. Surg.* **98**, 997–1002 (2011).
28. Usai, P. *et al.* Cigarette smoking and appendectomy: effect on clinical course of diverticulosis. *Dig. Liver Dis.* **43**, 98–101 (2011).
29. Humes, D. J., Ludvigsson, J. F. & Jarvholm, B. Smoking and the Risk of Hospitalization for Symptomatic Diverticular Disease: A Population-Based Cohort Study from Sweden. *Dis. Colon Rectum.* **59**, 110–114 (2016).
30. Auton, A. *et al.* A global reference for human genetic variation. *Nature*. **526**, 68–74 (2015).
31. Khera, A. V. *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet.* **50**, 1219–1224 (2018).

Acknowledgements

The authors would like to thank Annika Bohner, Irina Nowak and Friederike Reuner (all Homburg) for their outstanding work in analyzing the blood samples. This study was supported by a grant from the Faculty of Medicine, Saarland University (HOMFOR grant T201000747 to Matthias C. Reichert) and a grant of the Research Council of Lithuania No. SEN-06/2015/PRM15-135.

Author contributions

M.C.R., J.K. and F.L. designed the study; M.C.R., J.K., F.G., V.Z., B.A., A.T., J.I.L., M.K., M.C., C.J., A.S., M.M., M.G., C.S., G.K., L.J., N.P., S.N.W., T.G. and L.K. participated in acquisition and collection of the data, drafted the manuscript, and together with F.L. analyzed the data and finalized the manuscript, which was then revised by all authors. The final draft of the manuscript has been approved by all authors. The contents of this manuscript are our original work and have not been published, in whole or in part, prior to or simultaneous with our submission of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-58437-1>.

Correspondence and requests for materials should be addressed to M.C.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020