European Journal of Human Genetics (2016) 24, 1041–1048

www.nature.com/ejhg

ARTICLE

Polymorphisms of cystathionine beta-synthase gene are associated with susceptibility to sepsis

Christoph Sponholz*[,1,2,3], Marcel Kramer[1,2], Franziska Schöneweck[1,4], Uwe Menzel[5], Kolsoum Inanloo Rahatloo[2,6], Evangelos J Giamarellos-Bourboulis[1,7], Vassileios Papavassileiou[8], Korina Lymberopoulou[9], Maria Pavlaki[10], Ioannis Koutelidakis[11], Ioannis Perdios[12], André Scherag[1,4], Michael Bauer[1,3], Matthias Platzer[2] and Klaus Huse[2]

Sepsis is the systemic inflammatory host response to infection. Cystathionine beta-synthase (CBS)-dependent homocysteine (Hcy) pathway was demonstrated to affect disease severity and mortality in patients with severe sepsis/septic shock. Independent studies identified a single-nucleotide polymorphism (SNP, rs6586282, hg19 chr21:g.44478497C>T) in intron 14 of the CBS-coding gene (*CBS*) associated with Hcy plasma levels. We aimed to describe the association of this SNP and variants of a splice donor-affecting variable-number tandem repeat (VNTR, NG_008938.1:g.22763_22793[16_22]) 243 bp downstream of rs6586282 with severe human sepsis. We analyzed the VNTR structure and genotyped variants of rs6586282 and a neighboring SNP (rs34758144, hg19 chr21:g.44478582G>A) in two case–control studies including patients with severe sepsis/septic shock from Germany ($n=168$) and Greece ($n=237$). In both studies, we consistently observed an association of CBS VNTR alleles with sepsis susceptibility. Risk linearly increased with number of tandem repeats (per allele odds ratio in the adjusted analysis 1.34; 95% confidence interval (CI) = 1.17–1.55; $P<0.001$). Association had also been shown for rs34758144 whose risk allele is in linkage disequilibrium with one long VNTR allele (19 repeat). In contrast, we observed no evidence for an effect on 28-day survival in patients with severe sepsis/septic shock (per allele hazard ratio in the adjusted analysis for VNTR 1.10; 95% CI = 0.95–1.28; $P=0.20$). In a minigene approach, we demonstrated alternative splicing in distinct VNTR alleles, which, however, was independent of the number of tandem units. In conclusion, there is no ordinary conjunction between human CBS and severe sepsis/septic shock, but CBS genotypes are involved in disease susceptibility.
European Journal of Human Genetics (2016) **24,** 1041–1048; doi:10.1038/ejhg.2015.231; published online 28 October 2015

INTRODUCTION

Sepsis is an impetuous inflammatory response of the organism after the immune system has been challenged by inflammatory agents and the complex regulation of its components is inappropriate or fails. At least in part, susceptibility to infectious diseases but more importantly the course of the corresponding response and the clinical outcome depend on the genetic make-up of an individual.[1,2] Identifying genetic factors for complex diseases such as sepsis is complicated by the potentially large number of involved genes and by environmental factors that greatly influence the outcome such as medical treatment.

Beside inflammation-associated processes, multiple alterations in different metabolic pathways have been reported for sepsis.[3–6] Changes in the concentration of amino acids including the sulfur containing ones are known for a long time.[7] Especially, redox stress during infection increases the need for synthesis of glutathione and its precursor cysteine.[8–10] A major source of cysteine is transsulfuration from homocysteine (Hcy), a pathway which is also affected in sepsis.[11,12] A recent study focusing on the plasma metabolome in septic patients revealed correlation of alpha-n-aminobutyric acid with the patients Sequential Organ Failure Assessment (SOFA) Score and cystathionine levels.[13] Both cystathionine and alpha-n-aminobutyric acid are metabolites of Hcy downstream of the cystathionine beta-synthase (CBS, EC 4.2.1.22) reaction. CBS has a crucial role in the regulation of plasma Hcy, provision of cysteine and glutathione and in the formation of the gasotransmitter hydrogen sulfide[14] (Figure 1). The latter function is of interest with respect to the emerging role of hydrogen sulfide in physiological and pathological processes[15–17] – especially in sepsis.[18] Several mutations and polymorphisms in the *CBS* gene (NG_008938 at chromosome 21: hg19.44473301_44496040) have been associated with CBS function and Hcy plasma concentrations.[19–25] Two independent studies identified a common single-nucleotide polymorphism (SNP, rs6586282, hg19 chr21: g.44478497C>T) in intron 14 of *CBS* to be associated with higher Hcy plasma level.[25,26] However, no obviously functional difference between the SNP alleles exists. At the exon 15–intron 15 border (exon–intron numbering from NG_0089381.1) downstream (243 bp) of rs6586282 a 31-bp variable-number tandem repeat (VNTR, NG_008938.1:g.22763_22793[16_22]) was identified. In each of the

[1]Integrated Research and Treatment Center, Center for Sepsis Control and Care (CSCC), Jena University Hospital, Jena, Germany; [2]Genome Analysis, Leibniz Institute for Age Research—Fritz Lipmann Institute, Jena, Germany; [3]Department of Anaesthesiology and Intensive Care Therapy, Jena University Hospital, Jena, Germany; [4]Research group Clinical Epidemiology, CSCC, Jena University Hospital, Jena, Germany; [5]Systems Biology and Bioinformatics Group, Leibniz Institute for Natural Product Research and Infection Biology—Hans Knöll Institute, Jena, Germany; [6]School of Biology, College of Science, University of Tehran, Tehran, Iran; [7]4th Department of Internal Medicine, University of Athens, Medical School, Athens, Greece; [8]Department of Internal Medicine, Larissa University Hospital, Larissa, Greece; [9]2nd Department of Internal Medicine, Sismanogleion General Hospital, Athens, Greece; [10]Department of Internal Medicine, Argos General Hospital, Argos, Greece; [11]2nd Department of Surgery, University of Thessaloniki, Medical School, Thessaloniki, Greece; [12]1st Department of Internal Medicine, 'G. Gennimatas' General Hospital, Athens, Greece
*Correspondence: Dr C Sponholz, Department of Anaesthesiology and Critical Care Medicine, Friedrich-Schiller-University, Erlanger Allee 101, Jena D-07747, Germany. Tel: +49 3641 9322225; Fax: +49 3641 9323112; E-mail: christoph.sponholz@med.uni-jena.de
Received 31 March 2015; revised 14 September 2015; accepted 29 September 2015; published online 28 October 2015

VNTR repeat units, a putative donor splice site is located indicative of a splice donor array (Figure 2) with potentially functional implications. So far, this VNTR was associated with plasma Hcy levels following postmethionine load.[27,28] Thus, it seems reasonable to investigate linkage disequilibrium (LD) of the VNTR alleles with those of rs6586282 given that the VNTR might be the functional variation causing the observed CBS associations.

We hypothesize that either one or both variations in *CBS* affect Hcy metabolism and thereby confer susceptibility for severe sepsis/septic shock. Therefore, we perform two case–control studies with severe sepsis/septic shock patients from Germany and Greece for potential allelic association of the *CBS* VNTR marker alleles. Second, we assess LD between the *CBS* VNTR and two SNP (rs6586282 and a neighboring SNP rs34758144) alleles. Third, we investigate the prognostic value of the VNTR alleles for patient 28-day survival in sample of our patients. Finally, we search for molecular mechanisms underlying the observed association by analyzing potential VNTR allele-specific splicing in a minigene approach.

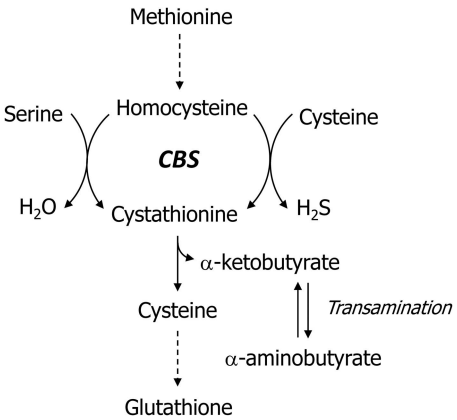


Figure 1 Cystathionine beta-synthase (CBS) as a key enzyme in the homocysteine metabolism.

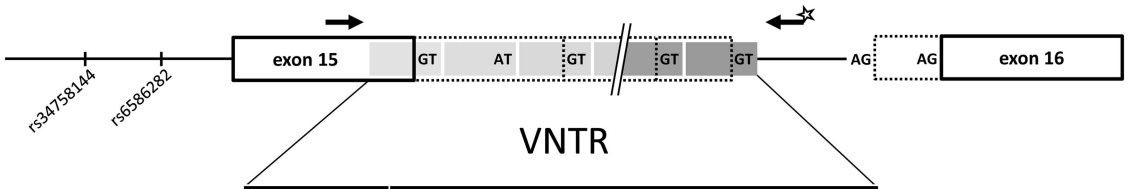
MATERIALS AND METHODS

CBS nomenclature

Lievers *et al* introduced a nomenclature for the sequence content of individual repeats in the *CBS* VNTR facilitating the description of haplotypes.[28] Based on this, we use a slightly different nomenclature relying on the sequence of the 18-repeat allele present in the hg19/build37 assembly of the human genome (chr21:44477721_44478278) avoiding different designations for the same repeat sequence (Supplementary Table 1). For example, 14 different repeat units (A–N, Figure 2) constitute the hg19 allele in the order ABCDEFGHIJJ-KIGCLMN (Table 1).

Sample characteristics

Cases. The severe sepsis/septic shock cases were ascertained in two settings. Although the German patients were collected in the multidisciplinary surgical intensive care unit (ICU) of Jena University Hospital (JUH), Greek patients arose from a multicentric patient cohort from the Hellenic Sepsis Study Group (HSSG, http://hwww.sepsis.gr) with a focus on internal medicine departments. Regarding the Greek patients no specific stratification was followed. Study sites provided samples of patients enrolled in the study in a sequential row. All patients admitted to the ICU were screened within 24 h after admission for clinical signs of systemic inflammatory response syndrome resulting from possible or proven infection. When sepsis was diagnosed, all adult patients with organ dysfunction according to the ACCP/SCCM criteria[29] were eligible for study inclusion. Blood samples were collected within 24 h after clinical diagnosis of severe sepsis/septic shock. After approval by the ethics committees,



Repeat unit	Repeat sequence
A	GTCATCTACAAGCAGTTCAAACAG**GT**ACCCA
B	GTCACCTACAGGCAGCTCAAACAG**AT**GCGCA
C	GTCACCTACAGGCAGCTCAAACAG**GT**GCCCG
D	GTCACCTACACGCAGCTCAAACAG**GT**GCGCG
E	GTCACCTACAGGCAGCTCAAACAG**GT**GAGCG
F	GTCACCTACAGGCAGCTCAAACAG**GT**ACCCG
G	GTCACCTACAGGCAGCTCAAACAG**GT**GCGCG
H	GTCAGCTACAGGCAGCTCAAGCGG**GT**GCGCG
I	GTCACCTACAGGCAGCTCAAACGG**GT**GCCCA
J	GTCACCTACAGGCAGCTCAAACGG**GT**GCCCG
K	GTCACCTACAGGCAGCTCAAACGG**GT**GCGCG
L	GTCAGCTACAGGCAGCTCAAACAG**GT**GCGCG
M	GTCACCTACACGCAGTTCAAACAG**GT**GCGCA
N	GTCACCTACACGCGGCTCAAACAG**GT**ACTCA
O	GTCACCTACAGGCAGCTCAAAAGG**GT**GCCCG
P	GTCAGCTACACGCAGCTCAAACAG**GT**GCCCG

Figure 2 Human *CBS* gene structure (intron 14 to exon 16) and analyzed variations in the context of the exon 15–intron 15 border. Position of the VNTR in intron 15 is shown with the relative positions of the primers used for genotyping (arrows). PCR primers in exon 15 (upstream of donor array) and intron 15 (downstream of donor array) give amplicons of different sizes, which can be separated by capillary electrophoresis. The fluorescence-labeled primer is marked by a star. Identified repeat units and their sequences are given below: donor sites are in bold and the *Kpn*I sites in repeat A and B are underlined. Repeats present in the hg19 18-repeat allele (ABCDEFGHIJJKIGCLMN) are shaded. The alternative acceptors of exon 16 referred to in the text are depicted by AG.

Table 1 Nomenclature and composition of CBS alleles

Nomenclature		*Number of repeat*		*Order of repeat units*		
Allele	*Naming*	*Length*	*Units*	*Common start*	*Unique center*	*Common end*
hg19 allele	H1	18	14	ABCD	**EFGHIJJKIGC**	LMN
Alternative allele	H2	18	14	ABCD	**FGHJJJKOGCP**	LMN
Allele-17		17	14	ABCD	**EFGHJKIGCP**	LMN
Allele-19		19	15	ABCD	**EFGHJJIKIKJP**	LMN
Allele-21		21	14	ABCD	**EFGHJJJKJKJKCP**	LMN

Bold letters indicate the specific order of repeat units within the different CBS alleles.
Underlined letters indicate *Kpn*I restriction site.

all patients or legal surrogates gave informed consent for genetic analyses and data collection.

Controls. Patients admitted to the University Hospital Jena for elective abdominal surgery and healthy volunteers from Jena (Germany) and Greece served as controls. All Greek study sites provided control samples. General inclusion criteria were European origin and lack of any type of familial relationship with other patients/control individuals participating in this study. Blood samples were taken after reviewing the inclusion and exclusion criteria by a questionnaire. All patients gave written informed consent for genetic analyses and data collection.

Genotyping was performed for all individuals, that is, cases and controls, in the same laboratory by personnel blinded to the phenotype information. The study was conducted in accordance with the Declaration of Helsinki (World Medical Association 2001).

All variants and individuals were submitted to the Leiden Open Variation Database, refer to http://www.lovd.nl/cbs for details (patient IDs: 00046568, 00046569 and 00047820).

DNA extraction

Genomic DNA from the collected EDTA blood and lymphoblastoid cell lines was extracted and purified using the QIAmp DNA Blood Kit (Qiagen, Hilden, Germany).

Cell lines and cell culture

Lymphoblastoid cell lines were obtained from collections at Coriell Cell Repositories (Camden, NJ, USA) and grown in RPMI 1640 medium (Gibco/BRL, Eggenstein, Germany) supplemented with 15% fetal calf serum (Gibco/BRL) and L-glutamine (2 mM, Gibco/BRL) at 37 °C, 5% CO_2 and 95% humidity.

HEK293-EBNA cells (kindly provided by C Kaether, Leibniz Institute for Age Research, Jena, Germany) were cultured in Dulbecco's MEM medium (PAA Laboratories, Paschberg, Austria) supplemented with 10% fetal calf serum at 5% CO_2 atmosphere, 95% humidity and 37 °C.

PCR amplification

PCR was performed with 'BioMix white' (Bioline, Randolph, MA, USA) according to the manufacturer's protocol in a total volume of 25 µl with 10 pmol of each primer. PCR reaction started with an initial denaturation at 93 °C for 1 min, followed by 25 cycles of 30 s; denaturation at 95 °C for 30 s; annealing at 59 °C and 1 min elongation at 72 °C. Primers (all from Metabion, Planegg/Steinkirchen, Germany) for genotyping the VNTR and for SNP typing are depicted in Supplementary Table 6 of the supplement. The amplicon for SNP typing covers positions of 20 known SNPs (dbSNP141) of which rs6586282, hg19 chr21:g.44478497C>T and, in addition, the nearby variant rs34758144, hg19 chr21:g.44478582G>A were called.

Sequencing

PCR products were precipitated by 2 µl of 7.5 mM ammonium acetate and 85 µl of ethanol (96%, v/v) washed twice with ethanol (70%, v/v) and Sanger sequenced with the amplification primers using Dye Terminator 3.1 chemistry and a 3730×l DNA Analyzer (Applied Biosystems, Darmstadt, Germany).

Capillary electrophoresis—laser-induced fluorescence analysis

Capillary electrophoresis was conducted following the protocol previously described.[30] PCR amplification was carried out with CBSx15donf and the 5'-6-carboxyfluorescein-labeled CBSx15donr primer (Supplementary Table 6). PCR products were appropriately diluted (up to 1/40) and 1 µl was supplemented with 10 µl of formamide (Roth) and 0.5 µl of GeneScan ROX 500 (Applied Biosystems) size standard. The mixture was denatured at 94 °C for 3 min, and subsequently cooled on ice. The denatured products were separated on an ABI 3730×l capillary sequencer and analyzed with the GeneMapper 4.0 software (Applied Biosystems). VNTR repeat lengths were identified according to the size of the PCR products (Supplementary Figure 1A). Genotyping by capillary electrophoresis was confirmed by sequencing of cloned PCR products of 18-, 19- and 21-repeat length and H1/H2 haplotypes as controls.

Haplotyping

PCR products obtained with primers CBSx15donf and 5'-6-carboxyfluorescein-labeled CBSx15donr were digested by incubation of 10 µl amplicon with *Kpn*I restriction enzyme in the appropriate buffer in 20 µl total volume for 2 h at 37 °C. 2 µl of digested samples were subjected to capillary electrophoresis as aforementioned (Supplementary Figure 1B).

Splicing analysis in a *CBS* minigene

Minigenes were constructed exhibiting the complete VNTR-containing intron 15 flanked by parts of exon 15 and exon 16. Therefore, PCR products obtained from a pool of human genomic DNA (Roche) with primers CBSmgx15f and CBSmgx16r (Supplementary Table 6) were cloned into pcDNA3.1-TOPO vector (Clontech, Saint-Germain-en-Laye, France). Plasmids containing the VNTR length alleles 17, 18, 19 and 21 and haplotypes H1 and H2 of 18-repeat alleles were identified and verified by Sanger sequencing. Besides its specific VNTR length, the 21-repeat minigene differs at three intronic nucleotide positions by carrying the haplotype G-T-C from the common variations rs234702-rs1005584-rs1005585 (hg19 chr21:g.[44477543G>C; 44477095C>T; 44477033T>C]). All other minigenes exhibit the haplotype C-C-T.

HEK293-EBNA cells were transfected with *CBS* minigenes using Lipofectamine 2000 (Invitrogen, Darmstadt, Germany) according to the manufacturer's instructions. After 24 h, cells were harvested, total RNA was isolated using RNAeasy (Qiagen) and cDNA synthesized with Maxime RT Random Primer PreMix Kit according to the manufacturer's instructions (iNtRON Biotechnology, via HiSS Diagnostiks, Freiburg, Germany). Minigene-driven transcripts were amplified (for detailed primer information see Supplementary Table 6) and splice isoform fractions were calculated subsequent to capillary electrophoresis with laser-induced fluorescence analysis (GeneMapper 4.0 software, Applied Biosystems).

For sequence verification of splice isoforms, RT-PCR was performed with unlabeled primers. PCR products were cloned into pCR2.1-TOPO vector (Clontech) and Sanger sequenced.

Statistical analyses

Clinical characteristics of the 405 patients with severe sepsis/septic shock are summarized as absolute/relative frequencies or means ± standard deviation. Accordingly we compared German and Greek cases by (generalized) Fisher's

Table 2 Characteristics of the severe sepsis/septic shock cases

	All cases (n = *405*)	*German cases (JUH;* n = *168)*	*Greek cases (HSSG;* n = *237)*	*Comparison of case groups* (P-*value*[a])
Mean age [years] ± SD[b]	67 ± 12.2	64 ± 10.9	70 ± 12.6	<0.001
Gender, male	249 (61.5%)	113 (67.3%)	136 (57.4%)	0.049
Mean APACHE-II on admission ± SD[b]	21 ± 8.2	22 ± 8.7	21 ± 7.7	0.021
Mean SOFA Score ± SD[b]	10 ± 3.8	11 ± 3.4	9 ± 3.6	<0.001
28-Day mortality, n_{deaths}	156 (38.7%)	48 (28.6%)	108 (46.8%)	<0.001
Missing	0	0	2	
Site of infection				
Urogenital	53 (13.1%)	2 (1.2%)	51 (21.6%)	<0.001
Pneumonia	222 (55.0%)	73 (43.5%)	149 (63.1%)	
Abdominal	70 (17.3%)	66 (39.3)	4 (1.7%)	
Bacteremia	42 (10.4%)	10 (6.0%)	32 (13.6%)	
Soft tissue	12 (3.0)	12 (7.1%)	0 (0.0%)	
Endocarditis	5 (1.2%)	5 (3.0%)	0 (0.0%)	
Missing	1	0 (0.0%)	1	

[a]We compared the Greek and the German cases by Student's *t*-test for continuous data or (generalized) Fishers' exact test for count data (*P*-value derived from Monte Carlo simulations with 2000 replicates); missing values were excluded from the analyses.
[b]Standard deviation.

Table 3 Genotype distribution of the severe sepsis/septic shock cases

	All cases (severe sepsis patients; n = *405), (%)*	*German cases (JUH;* n = *168), (%)*	*Greek cases (HSSG;* n = *237), (%)*	*Comparison of case groups* (P-*value*[a])
Genotype distribution of *CBS* VNTR				0.118
17-17	0 (0.0)	0 (0.0)	0 (0.0)	
17-18	33 (8.1)	20 (11.9)	13 (5.5)	
17-19	7 (1.7)	4 (2.4)	3 (1.3)	
17-21	2 (0.5)	0 (0.0)	2 (0.8)	
18-18	235 (58.0)	88 (52.4)	147 (62.0)	
18-19	82 (20.2)	36 (21.4)	46 (19.4)	
18-21	36 (8.9)	14 (8.3)	22 (9.3)	
19-19	4 (1.0)	3 (1.8)	1 (0.4)	
19-21	5 (1.2)	3 (1.8)	2 (0.8)	
21-21	1 (0.2)	0 (0.0)	1 (0.4)	
Genotype distribution of *CBS* SNP rs34758144				0.609
G-G	247 (75.3)	116 (73.4)	131 (77.1)	
A-G	78 (23.8)	40 (25.3)	38 (22.4)	
A-A	3 (0.9)	2 (1.3)	1 (0.6)	
Missing	77	10	67	
Genotype distribution of *CBS* SNP rs6586282				0.195
C-C	216 (65.9)	97 (61.4)	119 (70.0)	
C-T	106 (32.3)	57 (36.1)	49 (29.8)	
T-T	6 (1.8)	4 (2.5)	2 (1.2)	
Missing	77	10	67	

[a]We compared the Greek and the German cases by generalized Fishers' exact test for count data (*P*-value derived from Monte Carlo simulations with 2000 replicates); missing values were excluded from the analyses.

exact or Student's *t*-test in Table 2, respectively. The non-parametric Wilcoxon–Mann–Whitney tests for continuous variables were applied as sensitivity analyses (as results were similar data are not shown). All genotype distributions were checked for evidence for a deviation from Hardy–Weinberg-equilibrium using exact or Monte Carlo exact tests (2000 replicates)[31,32] in cases and controls separately for the VNTR and SNP alleles; all nominal *P*-values ≥ 0.01; Table 3).

We compared cases and controls using logistic regression with and without (data not shown) adjustment for covariates (sex, age (linear) and country) and report estimated adjusted odds ratios (ORs) and 95% confidence intervals (95% CIs) for the genotype effect and the respective nominal two-sided *P*-values in Table 4. Similarly, we performed time-to-event analyses by Cox regression. Survival was defined as the interval from time of diagnosis of severe sepsis/septic shock until death. Patients alive after the 28-day follow-up were regarded as censored. Kaplan–Meier estimators were used to display the overall 28-day survival data for the VNTR and SNP genotypes (see Supplementary Figures 2–4). We report adjusted (sex, age (linear), country (Greece, Germany) and APACHE-II as covariates) estimated hazard ratios (HR), 95% CIs and nominal two-sided *P*-values in Table 5. Next, we recoded the VNTR genotypes numerically starting with 0 for 17-17 to 8 for 21-21 to estimate the (linear)

Table 4 Results of the adjusted logistic regression analyses comparing severe sepsis cases and controls: odds ratio point estimates, 95% CIs and *P*-values (two-sided) are reported

	All cases and controls[a]		*German cases and controls*[b]		*Greek cases and controls*[b]	
CBS marker genotype	*Odds ratio (95% CI)*	P-*value*	*Odds ratio (95% CI)*	P-*value*	*Odds ratio (95% CI)*	P-*value*
VNTR						
18-18	1	—	1	—	1	—
18-19	2.03 (1.36;3.06)	<0.001	2.26 (1.23;4.29)	0.01	1.88 (1.11;3.24)	0.02
18-21	1.61 (0.94;2.79)	0.08	1.87 (0.78;4.70)	0.17	1.49 (0.76;3.00)	0.25
17-18	0.64 (0.40;1.02)	0.06	1.00 (0.51;1.92)	0.99	0.41 (0.20;0.80)	0.01
Other	1.33 (0.67;2.68)	0.42	1.08 (0.43;1.02)	0.87	1.94 (0.65;1.30)	0.25
Per VNTR repeat unit	1.34 (1.17;1.55)	<0.001	1.26 (1.02;1.56)	0.03	1.42 (1.17;1.73)	<0.001
SNP rs34758144						
G-G	1	—	1	—	1	—
A-G/A-A	2.81 (1.82;4.44)	<0.001	2.44 (1.35;4.58)	0.004	3.31 (1.75;6.58)	<0.001
SNP rs6586282						
C-C	1	—	1	—	1	—
C-T/T-T	1.48 (1.05;2.09)	0.03	1.68 (1.03;2.77)	0.04	1.29 (0.80;2.11)	0.30

[a]Adjusted for sex, age (linear), country (Germany, Greece).
[b]Adjusted for sex and age (linear).

Table 5 Results of the adjusted Cox regression analyses in severe sepsis cases: hazard ratio point estimates, 95% CIs and *P*-values (two-sided) are reported

	All cases[a]		*German cases*[b]		*Greek cases*[b]	
CBS marker genotype	*Hazard ratio (95% CI)*	P-*value*	*Hazard ratio (95% CI)*	P-*value*	*Hazard ratio (95% CI)*	P-*value*
VNTR						
18-18	1	—	1	—	1	—
18-19	1.22 (0.81;1.83)	0.34	1.22 (0.61;2.45)	0.58	1.19 (0.72;1.96)	0.50
18-21	1.21 (0.67;2.18)	0.52	0.75 (0.22;2.50)	0.64	1.74 (0.88;3.46)	0.11
17-18	0.69 (0.35;1.38)	0.30	0.94 (0.36;2.49)	0.91	0.50 (0.18;1.39)	0.19
Other	1.24 (0.54;2.84)	0.62	2.11 (0.72;6.15)	0.17	0.80 (0.19;3.27)	0.75
Per VNTR repeat unit	1.10 (0.95;1.28)	0.20	0.99 (0.76;1.29)	0.93	1.21 (1.00;1.46)	0.04
SNP rs34758144						
G-G	1	—	1	—	1	—
A-G/A-A	1.38 (0.91;2.10)	0.13	1.37 (0.72;2.60)	0.35	1.34 (0.76;2.36)	0.31
SNP rs6586282						
C-C	1	—	1	—	1	—
C-T / T-T	1.29 (0.87;1.91)	0.21	1.20 (0.66;2.17)	0.55	1.24 (0.73;2.12)	0.43

[a]Adjusted for sex, age (linear), country (Germany, Greece), APACHE-II (linear).
[b]Adjusted for sex and age (linear), APACHE-II (linear).

length–dosage effect (abbreviated as 'per allele' in Tables 3 and 4). Finally, we performed regression model diagnostics including graphical and formal checks and applied an exploratory two-sided significance level of 5% throughout. All analyses were run using PLINK (version 1.07) or R (version 3.1.1.; packages glm, gap and survival) if not otherwise specified.

RESULTS

Allele complexity of *CBS* intron 15 splice donor VNTR

To setup methods and gain first insights into the *CBS* VNTR allele heterogeneity, we analyzed the *CBS* locus DNA on 16 human-derived lymphoblastoid cell lines and two family members (CEPH/UTAH PEDIGREE 1362 and 1463). PCR-based re-sequencing identified four cell lines comprising an 18-repeat allele homozygous across the entire 558 nucleotides. One of them (GM11994) was homozygous for the hg19 allele, whereas the three others (GM10860, GM12716, GM15386) were homozygous for an alternative allele differing at 20 nucleotide positions. This alternative allele lacks repeat units E and I and contains two novel ones, O and P (see Table 1 for details). We further sequenced 43 individual samples of the German control group (JUH) carrying two 18-repeat length alleles. Thereof, 13 genomes were homozygote for the first (hg19) and 8 for the alternative allele. All other sequence patterns could be explained by heterozygosity composed of the two alleles. Sequencing of cloned 17-, 19- and 21-repeat length alleles (4 each) showed haplotypes with differing structures as depicted in Table 1.

Only repeat units A and F contain a *Kpn*I recognition sequence (GGTACC; Figure 2) and constitute two haplotypes by their relative positioning. Although A is in all alleles of the VNTR at the first position, repeat F may reside at positions 6 (further on called haplotype H1; eg, the hg19 allele) or 5 (H2; eg, the alternative allele-18, see Table 1). Therefore, *Kpn*I digestion enables a high-throughput identification of H2 alleles (Supplementary Figure 1B). Applying haplotyping by *Kpn*I digestion to all samples of our study, 792 (97.8% of 810) successfully digested samples were informative and identified 478 (59%) as H2 carriers (Supplementary Table 3).

VNTR and SNP alleles associated with susceptibility to severe sepsis/septic shock

To describe possible association of the CBS VNTR alleles and susceptibility to severe sepsis/septic shock, we analyzed two independent case–control studies with patients (cases) from Germany ($n=168$) and Greece ($n=237$). Patient characteristics are summarized in Table 2. Although Greek patients were older and included a higher proportion of females, German patients had a higher illness severity at ICU admission, measured by APACHE-II[33] and displayed a higher rate of organ dysfunction, measured by mean SOFA scores. Nevertheless, German patients had a lower 28-day mortality rate compared with Greek patients. However, a striking difference between the two patient groups was observable for the infection sites, which reflects the different settings of surgical or non-surgical ICUs (for convenience see Table 2). For both case groups, we selected non-septic controls from Germany and Greece with similar age and sex distributions (Germany: $n=168$, mean age 64±10.9 years in septic cases *vs* 64±10.9 years in controls ($P=0.998$), 67.3% *vs* 66.7% male gender ($P=1.0$); Greece: $n=237$, mean age 70±12.6 years in septic cases *vs* 69±12.4 years in controls ($P=0.684$), 57.4% *vs* 58.6% male gender ($P=0.852$)).

Among all 405 septic patients, *CBS* VNTR length ranged from 17 to 21 repeats. A 20-repeat allele was not detected (see Table 3 for details). To evaluate the allelic association of *CBS* polymorphisms (VNTR length as well as SNPs rs34758144 and rs6586282 genotypes) and susceptibility to sepsis, we conducted a logistic regression analysis comparing cases and controls. In this approach, susceptibility to sepsis significantly increased with length of VNTR alleles (per allele OR in the adjusted analysis 1.34; 95% CI = 1.17–1.55; $P<0.001$). Moreover, carriers of the A allele at rs34758144 (adjusted OR = 2.81; 95% CI = 1.82–4.44; $P<0.001$) and carriers of the T-alleles at rs6586282 (adjusted OR = 1.48; 95% CI = 1.05–2.09; $P<0.001$) were more frequently found among the sepsis patients. Despite the setting differences of patient recruitment, the results were consistently found in both Greece and German case–control studies (for details see Table 4).

LD between VNTR and SNP alleles

We calculated LD between *CBS* VNTR and SNP alleles for all genotyped individuals (Supplementary Table 4). The largest r^2 of 0.85 (Greek: 0.87; German: 0.84) resulted for the VNTR marker allele 19, which was in LD with the A allele of SNP rs34758144. The second largest r^2 of 0.60 (Greek: 0.58; German: 0.62) resulted for the VNTR marker allele 18, which was in LD with the C allele of SNP rs6586282. This difference in LD highlights the value of genotyping the more complex multi-allelic VNTR instead of the bi-allelic SNPs alone and explains the difference in the ORs of both SNPs (Table 4).

Prognostic value of the VNTR alleles for 28-day survival

To address the potential prognostic value of *CBS* markers, we evaluated the association of 28-day survival in severe sepsis/septic shock patients by VNTR or SNP genotype (Table 5 and Supplementary Figures 2-4) or by the number of VNTR length alleles. We found no evidence that supported a prognostic value of the *CBS* marker genotypes/alleles.

Allele-specific splicing is not VNTR length dependent

Searching for the molecular mechanisms underlying the observed association, we examined splicing of *CBS* transcripts for potential VNTR allele specificity. Each repeat unit (besides B) harbors a putative splice donor site. To first gain insights into the phylogenetic history of the VNTR and therewith into the impact of natural selection on potentially functional relevant genetic structures, we inspected available sequences of chimpanzee, gorilla, gibbon and rhesus monkey (Supplementary Table 2) for the structure of the respective donor of intron 15. For all species, we identified multiple donor repeats (chimpanzee 25, gorilla 21, rhesus 4, gibbon 2) with intact consensus splice site GT in each repeat unit including the second one (Supplementary Table 2). Therefore, G>A exchange in repeat unit B is most probably acquired in the human lineage preventing potential alternative splicing at this site only in this species.

To experimentally study alternative splicing within the donor array, we tested minigene constructs containing all four haplotype H1 length alleles (17, 18, 19 and 21 repeats) and the H2 18-repeat allele. Usage of the splice site in repeat unit A was the major donor event (96%) in H1 constructs and exclusively used in the H2 construct (Supplementary Figure 5A). The only functional alternative donor in the VNTR was detected in unit E, specific for haplotype H1, with a frequency of 4% independent of the allele length (see Supplementary Figure 5B). This is consistent with splice site predictions[28] yielding the highest splice site score among VNTR donors for unit E (Supplementary Table 5). Subsequently, the alternative transcript contains a 124-nt insertion of four 31-nt-repeat units causing a frameshift changing the C-terminus of the encoded CBS isoform c.517_518insNG_008938.1:g.22787_22910 (p.(Gln489_Ile490ins41;Ile490fs*48)). In a small fraction (~2%) of the transcripts, an alternative intron 15 acceptor was utilized (CAG, hg19 chr21:g44477234_44477236).

DISCUSSION

The phenotypic heterogeneity of sepsis may be responsible for the current lack of robust genetic associations, although a strong genetic predisposition for this multifactorial infectious disease is likely.[1] Manifold changes in metabolism occur in sepsis that further expand the complexity beyond the underlying infection. Transsulfuration as a central amino-acid pathway with CBS as a key enzyme serves diverse function by linking the metabolism of sulfur-containing amino acids with more than hundred *S*-adenosyl methionine-dependent methylation reactions, redox-control, the folate cycle and the metabolism of signaling molecules. CBS function is regulated by diverse metabolites (heme, CO, NO) and hormones, the levels of which change during sepsis. CBS degrades Hcy by condensation with either serine or cysteine. In the latter reaction hydrogen sulfide is formed, a gasotransmitter involved in many inflammatory processes. Manipulation of endogenous hydrogen sulfide is even considered a potential therapeutic intervention for sepsis.[34] Recently, data were presented suggesting the possibility of an independent regulation of hydrogen sulfide formation and canonical transsulfuration fueled by CBS.[35] Therefore, even subtle differences in CBS functions might have a conspicuous pleiotropic impact on complex phenotypes as the host response in sepsis. We hypothesized that genetic variations of *CBS* are associated with sepsis susceptibility. A 31-bp VNTR and a nearby SNP (rs6586282, within intron 14, 243 nt upstream of the splice donor site

of the first VNTR repeat) were both associated with alterations in Hcy metabolism and therefore studied in the present work. We developed a high-throughput assay to determine the genotypes of one of these variants, the multi-allelic VNTR originally described by Kraus *et al*.[36] Unfortunately, there is no common definition of the repeat unit in this VNTR, which leads to apparent differences of VNTR genotype annotation in several reports (Supplementary Table 1). In our study, we used the repeat definition of Lievers *et al*.[28] We found the 18/18-repeat homozygotes and the 18-repeat allele, respectively, as being the most prevalent. 20-repeat alleles were not found, neither in the JUH nor in the HSSG cohort. Long alleles of this VNTR are known to be associated with elevated Hcy levels and lower CBS activity indicating a disturbed Hcy homeostasis.[27,28]

We performed two genetic association studies with three *CBS* markers in case–control samples from different European countries, which was suggested to be advantageous.[37] We demonstrated an association of *CBS* VNTR alleles at the exon 15–intron 15 boundary with susceptibility to severe sepsis/septic shock in both German and Greek samples. Differences in the genetic landscape of these European regions have been demonstrated.[38] Importantly, although the JUH and HSSG cohorts also differed regarding age, mortality and source of sepsis, both displayed a similar shift toward longer *CBS* VNTR alleles compared with healthy controls. An allelic association was also found for two SNPs (rs6586282 and rs34758144), which we demonstrated to be in LD with the VNTR alleles. Accordingly, the distribution of the SNPs and VNTR length alleles and the respective genotypes all differed significantly between septic patients and control individuals.

Neither allele of the intronic SNP rs6586282 carries a sequence feature, which makes it an obvious causative variation, although an association with Hcy levels has been described previously. One should rather take a possible functional effect of the linked VNTR on *CBS* mRNA processing into account. Yang *et al*.[39] postulated that alternative splicing is unlikely due to the lack of a donor GT within the second repeat unit and the fact that all other donors are too far from exonic splice regulatory sequences. On the other hand, Lievers *et al*.[28] provided experimental evidence by RT-PCR for the retention of intronic sequence in *CBS* transcripts derived from cultured fibroblasts. Appearance and conservation of tandem repeat arrays from rhesus to human indicates functional importance, respectively. The *CBS* transcript levels are low in most tissues. Sufficient transcription of homozygous alleles is, however, a prerequisite to study VNTR allele-specific splicing. Therefore, we analyzed splicing of intron 15 in narrowed sequence context of a minigene. For haplotype H1, alternative splicing indeed occurs at the splice donor sites of repeats A and E, whereas in haplotype H2, repeat A exclusively provides the donor splice site. Another alternative transcript is derived from a splice acceptor site in intron 15. However, the observed transcript isoforms are formed at low frequency (below 5%). Neither the alternative donor nor the alternative splice acceptor usage in *CBS* are correlated with VNTR allele length. Therefore, the identified alternative splice events are not sufficient to explain the observed correlations of *CBS* length alleles and sepsis susceptibility. This, however, does not exclude the possibility that other splicing events are dependent on the VNTR length or that repeat unit numbers influence mRNA stability in their native context.

Shortcomings of our study are rather low number of well-characterized patient samples as well as the inclusion of only two European ethnicities. However, as we have analyzed not only patients from different geographic locations with well-characterized disease states, but also cases arising from either surgical or medical ICUs with varying infection sites and individual attributes. Observing consistent findings despite this increased heterogeneity may thus be regarded as a sign of robustness of our results.

Taken together, in our case–control studies, we found an association of the *CBS* VNTR with susceptibility to severe sepsis/septic shock in patients with very different clinical characteristics. The prognostic impact of the *CBS* VNTR should be examined in larger patient cohorts calculating sample sizes in advance given that this part of our analysis was likely the most underpowered. Thus, *CBS* remains a promising target in sepsis research.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Beate Szafranski for skillful technical assistance. This work was supported by the Federal Ministry of Education and Research (BMBF), Germany, FKZ: 01EO1002 and the Hellenic Sepsis Study Group.

1 Sørensen TI, Nielsen GG, Andersen PK, Teasdale TW: Genetic and environmental influences on premature death in adult adoptees. *N Engl J Med* 1988; **318**: 727–732.
2 Petersen L, Andersen PK, Sørensen TIA: Genetic influences on incidence and case-fatality of infectious disease. *PLoS One* 2010; **5**: e10603.
3 Biolo G, Toigo G, Ciocchi B *et al*: Metabolic response to injury and sepsis: changes in protein metabolism. *Nutr Burbank Los Angel Cty Calif* 1997; **13**: 52S–57S.
4 Plank LD, Hill GL: Sequential metabolic changes following induction of systemic inflammatory response in patients with severe sepsis or major blunt trauma. *World J Surg* 2000; **24**: 630–638.
5 Say J: The metabolic changes associated with trauma and sepsis. *Nurs Crit Care* 1997; **2**: 83–87.
6 Wu G, Fang Y-Z, Yang S, Lupton JR, Turner ND: Glutathione metabolism and its implications for health. *J Nutr* 2004; **134**: 489–492.
7 Freund H, Atamian S, Holroyde J, Fischer JE: Plasma amino acids as predictors of the severity and outcome of sepsis. *Ann Surg* 1979; **190**: 571–576.
8 Biolo G, Antonione R, De Cicco M: Glutathione metabolism in sepsis. *Crit Care Med* 2007; **35**: S591–S595.
9 Malmezat T, Breuillé D, Pouyet C, Mirand PP, Obled C: Metabolism of cysteine is modified during the acute phase of sepsis in rats. *J Nutr* 1998; **128**: 97–105.
10 Malmezat T, Breuillé D, Capitan P, Mirand PP, Obled C: Glutathione turnover is increased during the acute phase of sepsis in rats. *J Nutr* 2000; **130**: 1239–1246.
11 Malmezat T, Breuillé D, Pouyet C *et al*: Methionine transsulfuration is increased during sepsis in rats. *Am J Physiol Endocrinol Metab* 2000; **279**: E1391–E1397.
12 Semmler A, Smulders Y, Struys E *et al*: Methionine metabolism in an animal model of sepsis. *Clin Chem Lab Med* 2008; **46**: 1398–1402.
13 Chiarla C, Giovannini I, Siegel JH: Characterization of alpha-amino-n-butyric acid correlations in sepsis. *Transl Res J Lab Clin Med* 2011; **158**: 328–333.
14 Beard RS, Bearden SE: Vascular complications of cystathionine β-synthase deficiency: future directions for homocysteine-to-hydrogen sulfide research. *Am J Physiol Heart Circ Physiol* 2011; **300**: H13–H26.
15 Li L, Moore PK: Putative biological roles of hydrogen sulfide in health and disease: a breath of not so fresh air? *Trends Pharmacol Sci* 2008; **29**: 84–90.
16 Paul BD, Snyder SH: H_2S signalling through protein sulfhydration and beyond. *Nat Rev Mol Cell Biol* 2012; **13**: 499–507.
17 Whiteman M, Le Trionnaire S, Chopra M, Fox B, Whatmore J: Emerging role of hydrogen sulfide in health and disease: critical appraisal of biomarkers and pharmacological tools. *Clin Sci Lond Engl* 1979 2011; **121**: 459–488.
18 Norris EJ, Larion S, Culberson CR, Clemens MG: Hydrogen sulfide differentially affects the hepatic vasculature in response to phenylephrine and endothelin 1 during endotoxemia. *Shock Augusta Ga* 2013; **39**: 168–175.
19 Aras O, Hanson NQ, Yang F, Tsai MY: Influence of 699C–>T and 1080C–>T polymorphisms of the cystathionine beta-synthase gene on plasma homocysteine levels. *Clin Genet* 2000; **58**: 455–459.
20 De Stefano V, Dekou V, Nicaud V *et al*: Linkage disequilibrium at the cystathionine beta synthase (CBS) locus and the association between genetic variation at the CBS locus and plasma levels of homocysteine. The Ears II Group. European Atherosclerosis Research Study. *Ann Hum Genet* 1998; **62**: 481–490.
21 Kraus JP, Janosík M, Kozich V *et al*: Cystathionine beta-synthase mutations in homocystinuria. *Hum Mutat* 1999; **13**: 362–375.
22 Kruger WD, Wang L, Jhee KH, Singh RH, Elsas LJ: Cystathionine beta-synthase deficiency in Georgia (USA): correlation of clinical and biochemical phenotype with genotype. *Hum Mutat* 2003; **22**: 434–441.
23 Mayfield JA, Davies MW, Dimster-Denk D *et al*: Surrogate genetics and metabolic profiling for characterization of human disease alleles. *Genetics* 2012; **190**: 1309–1323.

24 Tsai MY, Bignell M, Yang F, Welge BG, Graham KJ, Hanson NQ: Polygenic influence on plasma homocysteine: association of two prevalent mutations, the 844ins68 of cystathionine beta-synthase and A(2756)G of methionine synthase, with lowered plasma homocysteine levels. *Atherosclerosis* 2000; **149**: 131–137.
25 Paré G, Chasman DI, Parker AN *et al*: Novel associations of CPS1, MUT, NOX4, and DPEP1 with plasma homocysteine in a healthy population: a genome-wide evaluation of 13 974 participants in the Women's Genome Health Study. *Circ Cardiovasc Genet* 2009; **2**: 142–150.
26 Tanaka T, Scheet P, Giusti B *et al*: Genome-wide association study of vitamin B6, vitamin B12, folate, and homocysteine blood concentrations. *Am J Hum Genet* 2009; **84**: 477–482.
27 Lievers KJA, Kluijtmans LAJ, Blom HJ, Wilson PW, Selhub J, Ordovas JM: Association of a 31bp VNTR in the CBS gene with postload homocysteine concentrations in the Framingham Offspring Study. *Eur J Hum Genet* 2006; **14**: 1125–1129.
28 Lievers KJ, Kluijtmans LA, Heil SG *et al*: A 31bp VNTR in the cystathionine beta-synthase (CBS) gene is associated with reduced CBS activity and elevated post-load homocysteine levels. *Eur J Hum Genet* 2001; **9**: 583–589.
29 Bone RC, Balk RA, Cerra FB *et al*: Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. *American College of Chest Physicians/Society of Critical Care Medicine. 1992. Chest* 2009; **136**: e28.
30 Kramer M, Quickert S, Sponholz C *et al*: Alternative Splicing of SMPD1 in Human Sepsis. *PLoS One* 2015; **10**: e0124503.
31 Wigginton JE, Cutler DJ, Abecasis GR: A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 2005; **76**: 887–893.
32 Guo SW, Thompson EA: Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 1992; **48**: 361–372.
33 Knaus WA, Draper EA, Wagner DP, Zimmerman JE: APACHE II: a severity of disease classification system. *Crit Care Med* 1985; **13**: 818–829.
34 Rivers JR, Badiei A, Bhatia M: Hydrogen sulfide as a therapeutic target for inflammation. *Expert Opin Ther Targets* 2012; **16**: 439–449.
35 Yadav PK, Xie P, Banerjee R: Allosteric communication between the pyridoxal 5'-phosphate (PLP) and heme sites in the H2S generator human cystathionine β-synthase. *J Biol Chem* 2012; **287**: 37611–37620.
36 Kraus JP, Oliveriusová J, Sokolová J *et al*: The human cystathionine beta-synthase (CBS) gene: complete sequence, alternative splicing, and polymorphisms. *Genomics* 1998; **52**: 312–324.
37 Kurreeman FAS, Stahl EA, Okada Y *et al*: Use of a multiethnic approach to identify rheumatoid- arthritis-susceptibility loci, 1p36 and 17q12. *Am J Hum Genet* 2012; **90**: 524–532.
38 Cruciani F, La Fratta R, Trombetta B *et al*: Tracing past human male movements in northern/eastern Africa and western Eurasia: new clues from Y-chromosomal haplogroups E-M78 and J-M12. *Mol Biol Evol* 2007; **24**: 1300–1311.
39 Yang F, Hanson NQ, Schwichtenberg K, Tsai MY: Variable number tandem repeat in exon/intron border of the cystathionine beta-synthase gene: a single nucleotide substitution in the second repeat prevents multiple alternate splicing. *Am J Med Genet* 2000; **95**: 385–390.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)