EBioMedicine xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Genetic Factors of the Disease Course After Sepsis: Rare Deleterious Variants Are Predictive

Stefan Taudien [a,b,1], Ludwig Lausser [b,c,1], Evangelos J. Giamarellos-Bourboulis [a,d,2], Christoph Sponholz [a,b,e], Franziska Schöneweck [a,f], Marius Felder [b], Lyn-Rouven Schirra [c], Florian Schmid [c], Charalambos Gogos [g,2], Susann Groth [b], Britt-Sabina Petersen [h], Andre Franke [h], Wolfgang Lieb [i], Klaus Huse [b], Peter F. Zipfel [j,l], Oliver Kurzai [a,k], Barbara Moepps [m], Peter Gierschik [m], Michael Bauer [a,e], André Scherag [a,f], Hans A. Kestler [b,c,l,*,1], Matthias Platzer [b,**,1]

[a] Integrated Research and Treatment Center, Center for Sepsis Control and Care (CSCC), Jena University Hospital, Jena, Germany
[b] Leibniz Institute on Aging – Fritz Lipmann Institute, Jena, Germany
[c] Institute of Medical Systems Biology, Ulm University, Germany
[d] 4th Department of Internal Medicine, National and Kapodistrian University of Athens, Athens, Greece
[e] Department of Anaesthesiology and Intensive Care Therapy, Jena University Hospital, Jena, Germany
[f] Research group Clinical Epidemiology, CSCC, Jena University Hospital, Jena, Germany
[g] Department of Internal Medicine, University of Patras, Medical School, Greece
[h] Institute of Clinical Molecular Biology, Christian-Albrechts-Universität Kiel, Kiel, Germany
[i] Institute of Epidemiology, Christian-Albrechts-Universität Kiel, Kiel, Germany
[j] Leibniz Institute for Natural Product Research and Infection Biology – Hans-Knöll-Institute, Jena, Germany
[k] Septomics Research Center Jena, Leibniz Institute for Natural Product Research and Infection Biology – Hans-Knöll-Institute, Jena, Germany
[l] Friedrich Schiller University Jena, Jena, Germany
[m] Institute of Pharmacology and Toxicology, Ulm University Medical Center, Ulm, Germany

ARTICLE INFO

Article history:
Received 24 May 2016
Received in revised form 19 August 2016
Accepted 24 August 2016
Available online xxxx

Keywords:
Sepsis
Exome
Rare single nucleotide variation
Population stratification
Classification
Semantic set covering machine

ABSTRACT

Sepsis is a life-threatening organ dysfunction caused by dysregulated host response to infection. For its clinical course, host genetic factors are important and rare genomic variants are suspected to contribute. We sequenced the exomes of 59 Greek and 15 German patients with bacterial sepsis divided into two groups with extremely different disease courses. Variant analysis was focusing on rare deleterious single nucleotide variants (SNVs).
We identified significant differences in the number of rare deleterious SNVs per patient between the ethnic groups. Classification experiments based on the data of the Greek patients allowed discrimination between the disease courses with estimated sensitivity and specificity > 75%. By application of the trained model to the German patients we observed comparable discriminatory properties despite lower population-specific rare SNV load. Furthermore, rare SNVs in genes of cell signaling and innate immunity related pathways were identified as classifiers discriminating between the sepsis courses.
Sepsis patients with favorable disease course after sepsis, even in the case of unfavorable preconditions, seem to be affected more often by rare deleterious SNVs in cell signaling and innate immunity related pathways, suggesting a protective role of impairments in these processes against a poor disease course.

1. Introduction

According to the new definition (Seymour et al., 2016; Shankar-Hari et al., 2016; Singer et al., 2016), sepsis is a life-threatening organ dysfunction caused by dysregulated host response to infection. Host genetic factors are important for the clinical course (Sorensen et al., 1988; Petersen et al., 2010). Only a limited number of molecular genetic studies in sepsis have been conducted so far - mostly focusing on candidate genes with known methodological challenges (Sutherland and Walley, 2009). Three genome-wide association studies (GWAS) related to

* Correspondence to: H.A. Kestler, Institute of Medical Systems Biology, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany.
** Correspondence to: M. Platzer, Leibniz Institute on Aging – Fritz Lipmann Institute (FLI), Beutenbergstraße 11, 07745 Jena, Germany.
E-mail addresses: hans.kestler@uni-ulm.de (H.A. Kestler), matthias.platzer@leibniz-fli.de (M. Platzer).
[1] These authors are designated as co-first or co-senior authors.
[2] On behalf of the Hellenic Sepsis Study Group.

http://dx.doi.org/10.1016/j.ebiom.2016.08.037

sepsis have been performed focusing on different phenotypes (e.g. therapeutic response within a randomized controlled trial (Man et al., 2012) or 28-day mortality (Rautanen et al., 2015; Scherag et al., 2016) and aiming for the identification of common genomic variants. However, rare genomic variants are suspected to contribute to the so-called "missing heritability" (Manolio et al., 2009), and the rare protein-affecting ones - predominantly evolved recently - have a high potential of causing deleterious effects. For example, rare and low-frequency variants with large effects were recently proven to be associated with coronary artery disease (Helgadottir et al., 2016). Furthermore, disease-related genes contain a higher proportion of these deleterious variants than other genes (Fu et al., 2013; Tennessen et al., 2012). Altogether, this suggests that assessment of rare deleterious protein affecting variants is a promising approach for elucidating the genetic component of sepsis. The identified variants can be used as proxies for inferring causality, a key step in identification of novel therapeutic targets.

To assess these variants, whole-exome sequencing (WES) is a successful strategy even for complex diseases like schizophrenia, cardiomyopathy or inflammatory bowel disease (Christodoulou et al., 2012; Loohuis et al., 2015; Norton et al., 2012). WES delivers ten-thousands of variants which subsequently have to be functionally prioritized which is still a critical issue despite the availability of numerous tools (Calabrese et al., 2009; Gonzalez-Perez and Lopez-Bigas, 2011; Li et al., 2009; Reva et al., 2011; Schwarz et al., 2014; Shihab et al., 2013). Remarkably, a unified approach for testing the association between rare variants and phenotypes in sequencing association studies was proposed and evaluated using sepsis-associated acute-lung-injury WES data (Lee et al., 2012).

As sepsis is a complex disease depending on genetic, environmental and live-history traits, we used a classification experiment as proof of principle for the role of rare genetic variants in the disease course. To recruit two classes, we carefully selected the most extreme cases from >4000 sepsis patients showing either a favorable or adverse disease course. To improve robustness of our approach (i) training and validation cohorts for the classification experiment were selected from different European populations and (ii) different criteria for defining the extremes in the two patient repositories were applied. Altogether, our approach allowed discrimination between the disease courses with high sensitivity and specificity, indicating the relevance of rare deleterious variants for sepsis research and ultimately new clinical applications.

2. Materials & Methods

2.1. Patients and Samples

Two patient cohorts of different European ethnic background were collected. For the study only patients were considered with at least one sepsis-associated organ failure. Patients with blood cultures yielding isolates of coagulase-negative Staphylococcus spp. or skin commensals were excluded. All subjects or their legal representatives gave written informed consent.

Greek patients were derived from the biobank of the Hellenic Sepsis Study Group which is a collection of biomaterial from patients with sepsis, severe sepsis and septic shock conducted in 65 departments in Greece since May 2006 (www.sepsis.gr). The study protocol is reviewed and approved by the Ethics Committees of the participating study sites (approval 26 June 2006). The selection of eligible patients for WES was done in June 2013 when 3955 patients were enrolled. All patients had a bacteria-positive blood culture. Further selection for extreme clinical phenotypes was done by filtering the patients with two different sets of criteria:

Group A (N = 32): i) age ≥ 18 years; ii) survival after 28 days despite the administration of empirically administered inappropriate antimicrobials. The inappropriateness of antimicrobials was realized when the antibiogram became known;

Group B (N = 27): i) relatively young i.e. age between 18 and 60 years; ii) lack of any comorbidity or other medical condition predisposing to sepsis, iii) critically ill with high mortality rates despite receiving appropriate therapy.

German patients were treated on the same ICU at the University Hospital Jena, Germany (August 2008–May 2011). The study approval was given by the faculty ethics review board (3624-11/12, 2712-12/09, 2160-11/07). All patients presented in clinically bad condition with septic shock resulting from anastomosis insufficiency after major abdominal surgery. Selection of extreme phenotypes from a pool of 120 patients was based on the course of organ dysfunction (measured by Sequential Organ Failure Assessment (SOFA) Scoring) resulting from the same focus of sepsis within a period of five days after sepsis onset:

Group A (N = 5): Patients with fast resolution of organ dysfunction, defined as decreasing SOFA scores of≥4;

Group B (N = 10): Patients with considerable worsening organ dysfunction, defined as increasing SOFA scores of≥4.

Although the definitions of sepsis stages of the study protocol were those of 2003, retrospective evaluation showed that all patients met the new Sepsis-3 definition (Seymour et al., 2016; Shankar-Hari et al., 2016; Singer et al., 2016). Detailed description of sepsis patient's characteristics are given in Table S1. Peripheral blood samples were taken from patients under aseptic conditions and kept refrigerated at −80 °C into an EDTA-coated tube. For all 74 patients, genomic DNAs were prepared from 200 μl blood each using the QIAamp DNA Mini Kit (Qiagen).

WES data of 93 healthy German control individuals were generated at the University Kiel, Germany. These individuals (81/87.1% females; 12/12.9% males; median age: 66; quantiles Q1: 62, Q3: 69) are part of the population-based cohort POPGEN (Nothlings and Krawczak, 2012) and their WES data were recently used as control group data in an early-onset IBD case-control study (Kelsen et al., 2015).

2.2. Whole Exome Sequencing

2–3 μg genomic DNA per sepsis patient was fragmented on a Covaris M220 focused ultra-sonicator and exomes were enriched by use of Agilent SureSelect XT Human All Exon V5 + UTRs kit, targeting 74,856,280 bp encompassing the coding sequence and untranslated regions of 20,791 human genes. After sequence capture target enrichment, individual libraries were prepared which were quantified and checked for quality by Agilent High Sensitivity DNA chip. Six libraries were pooled each and sequenced on the Illumina HiSeq2500 platform (RapidRun, 2 × 100 bp Paired End). On average, 5.4×10^7 sequence pairs (10.8 Gb) per sample were generated, corresponding to a 215-fold mean depth of coverage per exome (Table S2). A mean of 21% duplicates was detected. DNAs from control individuals were sequenced at the University Kiel after enrichment using the same kit as for the sepsis patients.

2.3. Mapping and Variant Assessment

The Illumina paired-end sequences of the sepsis patients were mapped to the entire human reference genome version GRCh37/hg19 using the Burrows-Wheeler Aligner BWA (Li and Durbin, 2009) with the default settings. Data was processed using the Genome Analysis ToolKit GATK v2.5 (DePristo et al., 2011; McKenna et al., 2010). Regions with alignment gaps were realigned (GATK IndelRealigner), duplicate reads were marked using Picard Tools (http://picard.sourceforge.net) and all aligned read data was subjected to base quality recalibration (GATK BaseRecalibrator). Reads that did not align, or aligned outside of the target regions, were discarded. For the mapped reads we obtained an 87-fold mean depth of coverage, ranging from 40-fold to 155-fold

Please cite this article as: Taudien, S., et al., Genetic Factors of the Disease Course After Sepsis: Rare Deleterious Variants Are Predictive, EBioMedicine (2016), http://dx.doi.org/10.1016/j.ebiom.2016.08.037