Identification of immunodominant *Bartonella bacilliformis* proteins: a combined in-silico and serology approach

Alexander A Dichter, Tilman G Schultze, Anne Wenigmann, Wibke Ballhorn, Andreas Latz, Elif Schlüfter, Palmira Ventosilla, Humberto Guerra Allison, Cesar Ugarte-Gil, Pablo Tsukayama, Volkhard A J Kempf

Summary

Background ***Bartonella bacilliformis*** **is the aetiological agent of Carrión's disease, a biphasic and highly lethal illness formerly restricted to the South American Andes that is now spreading to adjacent areas. Reliable serodiagnostic approaches and vaccines are urgently needed. In this study, we aimed to identify immunodominant proteins of** ***B bacilliformis*** **and to establish novel and reliable serodiagnostic tools.**

Methods **We used a reverse vaccinology approach in combination with an analysis of heterologous genomic expression libraries to identify immunodominant proteins, on the basis of the genome sequences of** ***B bacilliformis*** **strains KC583 and KC584. Antigens were screened with serum samples collected from Peruvian patients with** ***B bacilliformis*** **infections and from German healthy blood donors without history of travel to South America. We further analysed immunoreactive proteins of** ***B bacilliformis*** **with immunoblotting and line blots. We used selected target proteins to develop a diagnostic ELISA. To assess the performance of this ELISA, we did receiver operating characteristic analyses to assess the area under the curve, cutoff values, sensitivities, and specificities with 95% CIs.**

Findings **We used serum samples obtained between Dec 23, 1990, and May 5, 2018, from 26 Peruvian patients with** ***B bacilliformis*** **infections and serum samples taken between Aug 28 and Aug 31, 2020, from 96 healthy German blood donors. 21 potentially immunodominant proteins were identified and recombinantly expressed, and their reactivity was assessed with immunoblotting and line blots. Of these 21 antigens, 14 were found to be immunoreactive. By using serum samples of Peruvian patients with Carrión's disease and of healthy German blood donors, we identified three antigens (porin B, autotransporter E, and hypothetical protein B) as suitable immunodominant antigens, and we applied them in a diagnostic ELISA using two different antigen combinations (porin B plus autotransporter E and porin B plus autotransporter E plus hypothetical protein B). For the combination of porin B and autotransporter E, with optical density measured at 450 nm (OD_{450}) cutoff value of 0·29, sensitivity was 80·8% (95% CI 60·7–93·5) and specificity was 94·8% (88·3–98·3) for all Peruvian patient samples. For a combination of porin B, autotransporter E, and hypothetical protein B, with an OD_{450} cutoff of 0·34, sensitivity was 76·9% (56·4–91·0) and specificity was 93·8% (86·9–97·7) for all Peruvian patient samples.**

Interpretation **This novel ELISA could represent a useful serodiagnostic tool for future epidemiological studies of** ***B bacilliformis*** **in endemic areas. Additionally, the immunodominant antigens we have identified could provide a first basis for future vaccine development to prevent the highly lethal Carrión's disease.**

Funding **DRUID (Novel Drug Targets against Poverty-Related and Neglected Tropical Infectious Diseases) Initiative and Robert Koch Institute.**

***Lancet Microbe* 2021**

Published **Online** September 10, 2021 https://doi.org/10.1016/S2666-5247(21)00184-1

For the Spanish translation of the abstract see **Online** for appendix 1

For the Quechua translation of the abstract see **Online** for appendix 2

University Hospital, Goethe-University, Institute for Medical Microbiology and Infection Control, Frankfurt, Germany (A A Dichter MSc, T G Schultze PhD, A Wenigmann BSc, W Ballhorn, Prof V A J Kempf MD); **Eurofins Technologies Animal Health Germany, Dietzenbach, Germany** (A Latz PhD, E Schlüfter BSc); **Instituto de Medicina Tropical Alexander von Humboldt, Universidad Peruana Cayetano Heredia, Lima, Peru** (P Ventosilla MSc, H Guerra Allison PhD, C Ugarte-Gil PhD, Prof P Tsukayama PhD)

Correspondence to: Prof Volkhard A J Kempf, University Hospital, Goethe-University, Institute for Medical Microbiology and Infection Control, D-60596 Frankfurt am Main, Germany **volkhard.kempf@kgu.de**

Introduction

Carrión's disease is a vector-borne biphasic illness restricted to the South American Andes, with Peru as the most affected country.[1] The causative agent, *Bartonella bacilliformis*, is transmitted to humans by sandflies (*Lutzomyia* spp).[2] In the acute phase (Oroya fever), *B bacilliformis* infects human erythrocytes, causing a severe haemolytic anaemia with a case-fatality rate of up to 88% if untreated (decreasing to 0·7–10% if adequately treated). This acute stage is often followed by a chronic phase (verruga peruana) characterised by subcutaneous nodules as a result of vasculoendothelial proliferations.[3]

Little is known about the immune response to *B bacilliformis* infections apart from some evidence that humoral and cellular mechanisms mediate a lifelong humoral immunity conferring partial protection.[4] The seroprevalence of *B bacilliformis* antibodies is up to 65% in endemic areas of Peru.[5] The rate of asymptomatic human carriers of *B bacilliformis* is 37% in post-outbreak areas and 52% in endemic areas; asymptomatic individuals represent the main reservoir of the pathogen.[6]

Oroya fever is mainly diagnosed on the basis of clinical symptoms and by the detection of Giemsa-stained intraerythrocytic bacteria in peripheral blood smears, with a sensitivity of 36% but a specificity of 96%.[7]

Research in context

Evidence before this study

We searched PubMed for articles published in any language with the terms ("*Bartonella bacilliformis*" OR "Oroya fever") AND ("ELISA" OR "immunodominant proteins") from database inception to June 28, 2021. We found 95 articles, of which ten pertained to the evaluation of human serum samples (nine articles) or serum from experimentally infected animals (one article). Four proteins were analysed in more detail for seroreactivity: GroEL (two studies), succinyl-CoA-synthetase (subunits α and β, one study), phage-associated protein (five studies), and lysin motif-containing protein (one study with one patient serum sample). In four studies, characterised serum samples (either from patients who were symptomatic or experimentally infected animals) were included. The use of systematical approaches for analysis of immunodominant targets has not been applied so far, but it is indispensable for defining diagnostic and vaccine proteins.

Added value of this study

To our best knowledge, this study is one of the first that systematically analysed immunoreactive proteins of *B bacilliformis*. By combining in-silico and in-vitro approaches, we identified 21 putative antigen candidates. Three of these antigens were successfully applied to serodiagnostic tools (line blots and ELISA), and these were evaluated by using serum samples from patients with Carrión's disease from one of the best characterised serum libraries.

Implications of all the available evidence

We developed serodiagnostic tools for *B bacilliformis* infections on the basis of genomic analyses, reverse vaccinology, and recombinant protein expression by using qualified patient serum samples. These assays can detect *B bacilliformis* antibodies and could be used to gain better understanding of the seroreactivity of particular antigens and of the epidemiology of Carrión's disease in South America. Furthermore, the detailed analysis of the pattern of immunoreactive antigens through line blots represents a first basis for future development of a vaccine.

Microbiological culture methods are limited by the slow growth of *B bacilliformis* (2–6 weeks), but several PCR-based diagnostic approaches that use peripheral blood have been established.[3,8] Although PCR-based diagnostics are highly sensitive, they do not allow the detection of past infections, limiting their use for epidemiological surveys and disease control. To overcome this constraint, solid serodiagnostic tests are needed. Only four immunoreactive proteins have been previously described: GroEL (65 kDa heat-shock protein), LysM (lysin domain-containing protein), succinyl-CoA synthetase (subunit SCS-α and SCS-β), and Pap31 (phage-associated protein); Pap31 has been considered as a potentially suitable IgG seromarker indicating past exposure.[6] Although efforts have been made to apply Pap31 and other candidate proteins (GroEL, SCS-α, and SCS-β) to an ELISA assay format, no approved diagnostic test is available yet.[9] Additionally, vaccines are not established for this disease, although they are urgently needed.[10] However, before vaccine development, immunodominant proteins need to be identified.

Experimental work applying molecular genetics to *B bacilliformis* is demanding due to the fastidious growth of the pathogen and the absence of suitable animal infection models. Additionally, patient specimens are rare, further hampering ex-vivo analyses. Therefore, novel concepts to overcome these limitations are necessary. Reverse vaccinology, a method for in-silico prediction of antigens based on pathogen's genomic information, has successfully been used for the identification of immunodominant vaccine candidates.[11,12] In this study, we aimed to systematically identify immunodominant proteins of *B bacilliformis* to establish a reliable serodiagnostic tool by combining the analysis of genomic *B bacilliformis* expression libraries with reverse vaccinology.

Methods

Study design, human samples, and bacteria and cell cultures

Serum samples from Peruvian patients with a clinical diagnosis of *B bacilliformis* infection were taken and collected at the Instituto de Medicina Tropical at Universidad Peruana Cayetano Heredia (UPCH), Lima, Peru, between Dec 23, 1990, and May 5, 2018, (appendix 3 pp 2–3) as part of clinical standard of care and de-identified. Serum samples were stored at –20°C and provided by UPCH according to a collaboration contract (UPCH 737-01-19). Serum samples were used to identify immunoreactive proteins of *B bacilliformis* strain KC583 (American Type Culture Collection [ATCC] 35685; ATCC, Mansassas, VA, USA) by a combined reverse vaccinology, genomic, and protein expression approach. Potential seroreactive proteins were further analysed with line blots for reactivity and specificity. Promising antigen candidates were evaluated in prototypic ELISAs. As negative controls, we used serum samples of de-identified healthy German blood donors without travel history to South America; these samples (from generic blood donations) were taken between Aug 28 and Aug 31, 2020, and collected by the German Red Cross blood service of Baden-Württemberg–Hesse (Frankfurt, Germany). Cell culture-derived antigen was generated with human HeLa-229 cervical epithelial cells (ATCC CCL-2.1). Experiments were approved by the Institutional

See Online for appendix 3

Ethics Committee at UPCH (SIDISI 200047) and by the University Hospital Frankfurt am Main (reference 423/11; Frankfurt, Germany).

Whole-genome sequence analysis of *B bacilliformis* and sequence analysis

We cultured *B bacilliformis* strain KC583 and extracted genomic DNA for whole-genome sequencing as described in the appendix 3 (p 2). For whole-genome sequencing, we used PacBio long-read sequencing (Norwegian Sequencing Centre, Oslo, Norway; appendix 3 p 2), and for de-novo assembly we used Unicycler version 0.4.8. For genomic alignment, we annotated fasta sequences of all strains with prokka 1.14.6.[13] We used a read-mapping-based approach against the previously available genome sequences of strain KC583 (NC_008783.1) for in-depth sequence analyses (GenBank accession numbers SRR13618139 for KC583 sequenced in this study, SRR2823708 for FDAARGOS_174 [NZ_CP014012.2], and SRR10344655 for KC584 [NZ_CP045671.1]). To assess a broad presence of immunodominant protein candidates in *B bacilliformis* and to evaluate homologies affecting potential cross-reactivities with other *Bartonella* spp, we analysed these deduced protein sequences with National Center for Biotechnology Information (NCBI) protein–protein Basic Local Alignment Search Tool in the genomes of *B bacilliformis* KC583 (NC_008783.1), KC583 (sequenced in this study), ATCC 35685, KC583 (NZ_CP014012), KC584 (NZ_CP045671.1), FDAARGOS_174 (NZ_CP014012.2), *B ancashensis* (NZ_CP010401.1), *B henselae* (NZ_CP020742.1), and *B quintana* (NZ_AP019773.1). Identity values were calculated by pairwise alignment analysis (appendix 3 p 2).

Reverse vaccinology prediction

For the identification of potentially immunoreactive proteins, we predicted putative antigenic proteins in silico from the genomic sequences of *B bacilliformis* KC583 and KC584 with the web-based software application Vaxign 1 (appendix 3 p 2),[14] applying the following criteria: surface exposed or secreted proteins, adhesin probability higher than 0·4, exclusion of proteins with more than one transmembrane helix, MHC class I and class II epitope binding, and no sequence similarities to human, pig, or mouse proteins.

Immunofluorescence assays for determination of anti-*B bacilliformis* IgG in patient serum samples

We assessed the reactivity of patient serum samples with immunofluorescence assays (IFA; with fluorescein isothiocyanate conjugated goat anti-human IgG antibodies; Euroimmun, Lübeck, Germany). We used human HeLa-229 cervical epithelial cells (ATCC CCL-2.1) infected with *B bacilliformis* KC583 as antigens, and we evaluated human serum samples at different dilutions (1:80, 1:160, and 1:320) by fluorescence microscopy (appendix 3 p 3). For technical control, a rabbit anti-*B bacilliformis* serum was applied. We considered serum positive at a titre of 1:320 or higher and negative at a titre lower than 1:80. We considered samples with a titre of 1:80 or higher and lower than 1:320 equivocal. Additionally, we analysed the reactivity of patient serum samples by immunoblotting with *B bacilliformis* KC583 whole cell lysate.

Generation of recombinant bacterial *B bacilliformis* protein expression libraries

We cloned DNA fragments (1–4 kbp) of *B bacilliformis* KC583 genomic DNA (obtained and processed as described in the appendix 3 p 3) in pET28 plasmids (Merck, Darmstadt, Germany) in three different reading frames (pET28a, pET28b, and pET28c) under transcriptional control of a T7 promoter in *Escherichia coli* BL21 (DE3; New England Biolabs, Ipswich, MA, USA; appendix 3 p 6). From each library, we randomly picked 32 transformants and analysed them by colony PCR with oligonucleotides for amplification of the insert region, demonstrating the insertion of various genomic fragments (appendix 3 pp 3, 7–8, 12).

Immunoscreening of *B bacilliformis* expression libraries and identification of DNA insertions

We screened the genomic expression libraries by colony immunoblotting with a preadsorbed pool of serum samples from five patients (those found to be most reactive by IFA and blotting). Protein expression was induced by isopropyl-β-D-thiogalactopyranoside, and reactive clones were identified by immunoblotting. We identified plasmid-inserted DNA sequences of reactive clones by pairwise alignment with the genomic sequence of *B bacilliformis* KC583.

For **NCBI** see https://www.ncbi.nlm.nih.gov/

For the **genome sequence of ATCC 35685** see https://www.atcc.org/products/35685

Cloning, expression, and analysis of immunodominant *B bacilliformis* candidate proteins

We heterologously expressed selected target genes and tested them for serum reactivity by immunoblotting. We assembled linearised pET28a vector DNA and PCR-amplified inserts with Gibson cloning (NEBuilder HiFi DNA Assembly Gibson cloning; New England Biolabs) and transformed them in *E coli* BL21 (DE3). We confirmed gene expression by anti-T7 tag (Abcam, Cambridge, UK) immunoblotting, including non-induced cultures as negative controls (appendix 3 p 4).

We did immunoblot analysis with *B bacilliformis* and *E coli* cells according to standard laboratory procedures using polyclonal goat anti-T7 antibodies (Abcam), rabbit anti-*B bacilliformis* serum (generated by immunising rabbits over 28 days with inactivated *B bacilliformis* KC583; appendix 3 pp 2–3), and human serum samples (appendix 3 p 6).

Purification of recombinantly expressed proteins and line blotting

We purified recombinantly expressed proteins by poly-histidine-tag affinity chromatography under denaturing

conditions. After lysis, we ultracentrifuged the suspension and solubilised the resulting pellet in 8 M urea, 10 mM tris (tris[hydroxymethyl]aminomethane or Trizma), 50 mM NaH_2PO_4, and 100 mM NaCl at pH 8·0. The supernatant was cleared from cellular debris by ultracentrifugation, transferred onto a column, washed, and purified.

We used line blots for standardised analysis of target protein seroreactivity with serum samples from patients and controls. We produced the line blots by printing purified antigens with defined protein amounts (2·1 ng, 10·5 ng, and 21 ng per antigen line) on nitrocellulose membranes. We cut air-dried membranes into 3 mm strips and incubated them with human serum samples (1:100 diluted) for 1 h followed by 30-min incubation with horseradish peroxidase-coupled protein A/G (1:18 000). We quantified signal strength from scanned line blots by densitometric analysis (appendix 3 p 4).

Development of a recombinant anti-*B bacilliformis* IgG ELISA

We coated 96-well ELISA plates with 0·1 µg of each recombinantly expressed antigen (porin B, autotransporter E, and hypothetical protein B as the most reactive antigens) and stored them at 4°C. We added 100 µL of serum from patients or controls (diluted 1:100) per well and incubated the plates for 1 h followed by a 30-min incubation with horseradish peroxidase-coupled protein A/G (1:18 000). Development was done with 3,3′,5,5′-tetramethylbenzidine for 15 min. We measured absorbance at 450 nm, with 620 nm as a reference in an ultraviolet spectrophotometer (appendix 3 p 5).

Statistical analysis

We analysed differences in IgG-reactivity to target proteins between experimental (serum samples from Peruvian patients) and control groups (serum samples from healthy German blood donors) using the Mann-Whitney *U* test. To assess the performance of *B bacilliformis* IgG ELISA to distinguish between positive and negative samples, we did receiver operating characteristic (ROC) analyses to assess the area under the curve (AUC), cutoff values, sensitivities, and specificities with 95% CIs. p values of less than 0·05 were considered statistically significant. Statistical analyses were done with Prism V6 (GraphPad Software, San Diego, CA, USA).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

In this study, we used serum samples obtained between Dec 23, 1990, and May 5, 2018, from 26 Peruvian patients with *B bacilliformis* infections and serum samples taken between Aug 28 and Aug 31, 2020, from 96 healthy German blood donors to identify potentially immunodominant proteins of *B bacilliforimis*.

For the in-silico identification of candidate immunodominant proteins, we first aimed to determine the exact genomic background of the bacterium. We compared the genome sequences of KC583 (sequenced in this study; 1 417 364 bp), FDAARGOS_174 (NZ_CP014012.2; 1 443 410 bp), ATCC 35685 (ATCC; 1 417 066 bp), and KC584 (NZ_CP045671.1; 1 411 655 bp) to the reference sequence NC_008783.1 (1 445 021 bp), and showed that both published NC_008783.1 and FDAARGOS_174 sequences contain a major repetitive sequence of approximately 30 kbp that is absent in the genome of our KC583 isolate and the sequence given by ATCC (appendix 3 p 13). Compared with KC583, strain KC584 had a deletion in a second site, resulting in a loss of BARBAKC583_RS1505 (*bbadB*) and a fusion of the adjacent genes BARBAKC583_RS1500 and BARBAKC583_RS1510. When compared with NC_008783.1, four additional proteins in KC583 and 22 proteins in KC584 (eg, Pap31) were found to be genetically affected, ranging from single-point mutations to frameshift mutations (appendix 3 pp 9–10).

We did a dynamic Vaxign target analysis for *B bacilliformis* KC583 and KC584 using 1209 protein sequences for KC583 and 1200 protein sequences for KC584. 20 potential immunodominant target proteins were identified (table), with 15 presumptive outer membrane proteins and five extracellular proteins. Protein length ranged from 124 amino acids for hypothetical protein A to 1259 amino acids for BbadA, and adhesion probability ranged from 0·401 for flagellin to 0·872 for hypothetical protein B. BbadB was not identified as a potential target protein because the software did not predict the subcellular location of the protein.

When tested by IFA, we found that rabbit anti-*B bacilliformis* serum (technical positive control) was reactive to titres of 1:1600 against infected HeLa-299 cells, whereas rabbit pre-immune serum samples and serum samples from healthy German blood donors were unreactive (ie, no specific fluorescence signals for titres lower than 1:80). 12 (46%) of 26 serum samples of Peruvian patients tested positive for IgG with a titre of 1:320 or higher (samples 1, 2, 4, 6–9, 16, 17, 19, 23, and 24; representative samples depicted in the appendix 3 p 14), 11 (42%) samples tested equivocal (titres ≥1:80 to lower than 1:320: samples 3, 5, 10, 11, 13–15, 18, 21, 25, and 26), and three (12%) tested negative (titres lower than 1:80: samples 12, 20, and 22). Patient samples were additionally evaluated by immunoblotting. The most reactive patient serum samples (1, 2, 4, 8, and 9; appendix 3 p 14) from both IFA and immunoblotting were pooled in equal amounts for later screening of the expression libraries and to analyse the immunoreactivity of the in-silico predicted (table) and recombinantly expressed antigens.

For the generation of recombinant *B bacilliformis* expression libraries, genomic DNA fragments were

	Accession number	Designation	Functional class	Cellular localisation probability	Adhesin probability	Transmembrane helices	Length, amino acids
1*	WP_005765706.1	AprI/Inh family metalloprotease inhibitor	Other outer membrane protein	OM, 0·992	0·437	0	177
2*	WP_005766217.1	BbadC	Autotransporter protein	OM, 1·000	0·643	0	551
3*	WP_005766273.1	Hypothetical protein A	Other outer membrane protein	OM, 0·992	0·698	1	124
4*	WP_005766494.1	Porin A	Porin	OM, 0·992	0·689	0	406
5*	WP_005766517.1	TonB-dependent haemoglobin–transferrin–lactoferrin family receptor	Other outer membrane protein	OM, 1·000	0·464	0	731
6*	WP_035453117.1	LPS-assembly protein LptD	Other outer membrane protein	OM, 1·000	0·735	1	791
7*	WP_005766622.1	Autotransporter outer membrane β-barrel domain-containing protein A	Autotransporter protein	OM, 0·983	0·652	0	800
8*	WP_005766636.1	LysM peptidoglycan-binding domain-containing protein	Other outer membrane protein	OM, 0·886	0·453	0	397
9*	WP_005767307.1	DUF1561 family protein	Other outer membrane protein	EC, 0·965	0·610	0	649
10*	WP_005767360.1	BbadA	Autotransporter protein	OM, 0·995	0·861	0	1259
11*	WP_011807398.1	Flagellin	Flagellum-associated protein	EC, 0·971	0·401	0	380
12*	WP_005767706.1	Autotransporter outer membrane β-barrel domain-containing protein B	Autotransporter protein	OM, 0·949	0·561	0	1193
13*	WP_005767730.1	Flagellar hook protein FlgE	Flagellum-associated protein	EC, 1·000	0·751	0	404
14*	WP_051007750.1	Hypothetical protein B	Other outer membrane protein	EC, 0·965	0·872	0	610
15*	WP_193741240.1	Autotransporter outer membrane β-barrel domain-containing protein C	Autotransporter protein	EC, 0·964	0·652	0	297
16*	WP_005767753.1	Autotransporter outer membrane β-barrel domain-containing protein D	Autotransporter protein	OM, 0·952	0·855	0	825
17*	WP_005767762.1	Flagellar basal body L-ring protein FlgH	Flagellum-associated protein	OM, 0·992	0·410	1	230
18*	WP_005767896.1	Porin B	Porin	OM, 1·000	0·424	0	284
19*	WP_005767899.1	Pap31	Porin	OM, 0·993	0·534	0	300
20*	WP_005767902.1	Porin C	Porin	OM, 1·000	0·493	0	281
21†	WP_011807413.1	Autotransporter E	Autotransporter protein	OM, 0·992	0·351	0	1058
22‡	WP_005766221.1	BbadB	Autotransporter protein	Prediction failed	0·670	1	1235

The 22 listed proteins were identified in silico by Vaxign (*) or experimentally by immunoscreening of genomic *B bacilliformis* expression libraries (†); additionally, BbadB (‡) was selected due to its high homology to the immunoreactive *B henselae* BadA. Accession number, designation, functional class, predicted cellular localisation, adhesion probability, number of transmembrane helices, and protein length are given for each protein. EC=extracellular. OM=outer membrane.

Table: **Immunodominant protein candidates of *Bartonella bacilliformis***

expressed in three different reading frames (*E coli* DH5α, average insert size of 1900 bp), resulting in approximately 24700 recombinant clones (pET28a, about 9100 clones; pET28b, about 5400 clones; and pET28c, about 10 200 clones), covering more than 99% of the *B bacilliformis* KC583 genome. We used a pooled plasmid preparation for each reading frame to create three expression libraries in *E coli* BL21 (DE3). Analysis of protein production by anti-T7 tag immunoblotting revealed an expression rate higher than 96% (321 of 332 colonies, according to the analysis of one colony blot agar plate). GroEL and flagellin were characterised as immunoreactive by screening approximately 2000 colonies with rabbit-anti *B bacilliformis* serum, confirming the functional capability of the libraries.

We screened the expression libraries with a pool of pre-adsorbed IFA-positive patient serum samples (appendix 3 p 15). Because of the limited amount of patient samples, about 7400 of 24700 colonies in total were analysed. The analysis yielded six consistently reactive colonies. Sequence analysis of the respective plasmids revealed that all six colonies contained gene fragments coding for autotransporter outer membrane β-barrel domain-containing protein (autotransporter E, protein accession number WP_011807413.1; table).

In total, 19 antigen targets from the Vaxign prediction, the experimentally identified autotransporter E (from the expression library), and BbadB (included on the basis of homology to BadA) were recombinantly expressed

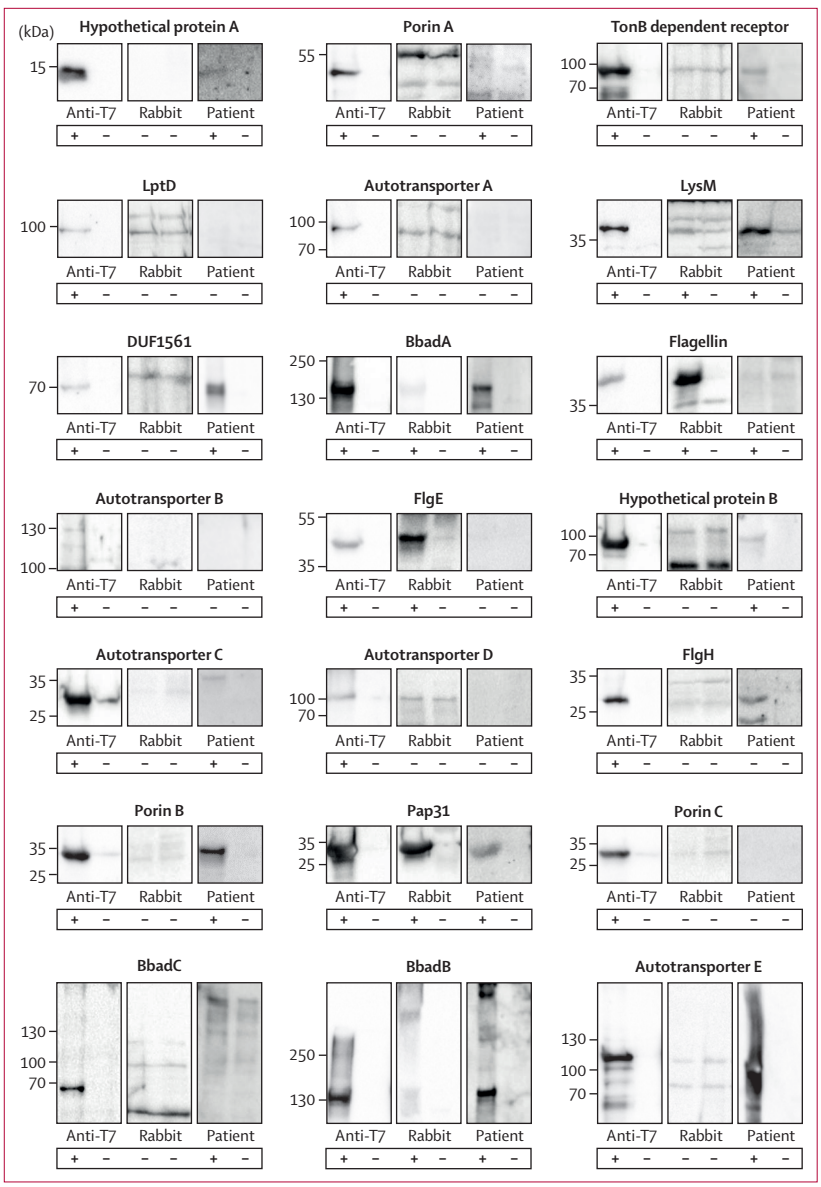

Figure 1: **Immunoblot analysis of target proteins**

All target proteins were produced in *Escherichia coli BL21* (DE3) and subjected to immunoblotting (left lanes show induced protein expression, right lanes show uninduced protein expression). Heterologous protein production was confirmed by T7 tag immunoblotting (anti-T7, technical control). For analysis of immunoreactivity, serum of a *Bartonella bacilliformis*-immunised rabbit and a pool of five IFA-positive and immunoblot-positive serum samples of Peruvian patients were used. Reactivity is given individually per target protein (bottom rows, indicated as a plus sign [reactive] and a negative sign [non-reactive]). IFA=immunofluorescence assay.

and analysed for protein production (with anti-T7-tag antibodies as technical control), reactivity with rabbit anti-*B bacilliformis* serum, and reactivity with the pool of five IFA-positive patient serum samples. Six proteins (LysM, BbadA, flagellin, FlgE, Pap31, and BbadB) were reactive with the rabbit serum and 14 proteins (hypothetical protein A, porin A, TonB-dependent receptor, LysM, DUF1561, BbadA, hypothetical protein B, autotransporter C, FlgH, porin B, Pap31, BbadC, BbadB, and autotransporter E) showed reactivity with Peruvian patient serum samples (figure 1). Recombinant expression of AprI/Inh family metalloprotease inhibitor was unsucessful; therefore, this protein was not included in further analyses.

Subsequently, 13 of 14 seroreactive proteins identified were printed on line blots (figure 2A, B). Because of its strong reactivity with human control serum samples, LysM was excluded from further analyses (data not shown). Human serum samples were categorised into three groups: anti-*B bacilliformis* IgG IFA-positive patient samples with an anti-*B bacilliformis* IgG IFA titre of 1:320 or higher (n=12); equivocal or negative patient samples with a titre lower than 1:320 (n=14); and, as a control group, samples from healthy German blood donors without travel history to South America (n=96). None of the tested human samples showed a specific signal for BbadC, porin A, TonB-dependent receptor, FlgH, hypothetical protein A, and autotransporter C, whereas DUF1561 was highly reactive with control samples (11 [11%] of 96). Both BbadA and BbadB reacted with one (8%) of 12 IFA-positive patient samples and were highly reactive with control samples (ten [10%] of 96 for BbadA and six [6%] for BbadB). Pap31 showed nearly no reactivity with IFA-positive samples (one [8%] of 12) nor with IFA-equivocal samples (one [7%] of 14), showing no significant differences from control samples (figure 2C). Hypothetical protein B reacted with two (17%) of 12 IFA-positive samples and porin B reacted with three (25%), showing significant differences from control samples (p=0·020 for hypothetical protein B and p=0·0008 for porin B). Autotransporter E displayed the strongest immunoreactivity among the tested antigens, with seven (58%) of 12 IFA-positive samples being reactive (p=0·0036), whereas none of the control samples showed a positive signal. Two of 12 IFA-positive samples (patient samples 16 and 17) were non-reactive to any antigen.

The most reactive antigens in the line blots (porin B, autotransporter E, and hypothetical protein B) were additionally tested in an ELISA format to detect anti-*B bacilliformis* IgG antibodies. We used two different antigen combinations (porin B plus autotransporter E and porin B plus autotransporter E plus hypothetical protein B) to test for IgG antibodies either with Peruvian patient samples that were IFA positive (n=12), samples that were equivocal or negative (n=14), and control samples (n=96). Both ELISAs showed significant detection (p<0·0001) of IFA-positive samples (figure 3A, B). A ROC analysis for the ELISA with a combination of porin B and autotransporter E yielded an AUC of 0·9335 (95% CI 0·8662–1·0010) when all patient samples were included in the analysis (figure 3C). At an optical density measured at 450 nm (OD_{450}) cutoff value of 0·38 (based on highest specificity), the sensitivity for IFA-positive samples was 100·0% (12 of 12 samples; 95% CI 73·5–100·0) and the specificity was 100·0% (96 of 96, 96·2–100·0). At the

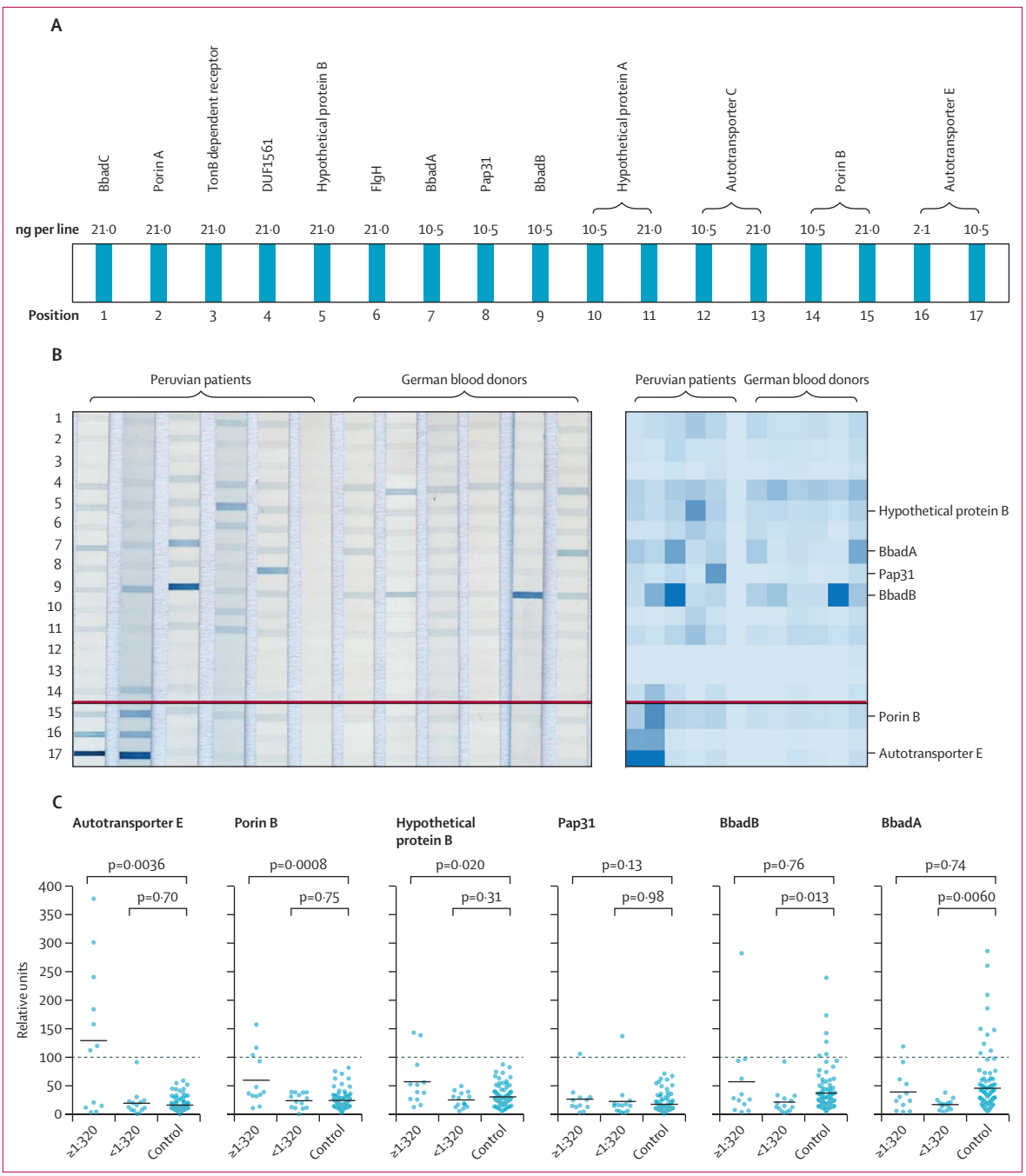

Figure 2: **Line blot analysis of serum samples from Peruvian patients and German controls using *B bacilliformis* seroreactive proteins**
(A) Design of the line blot: polyhistidine-tag purified antigens were printed on nitrocellulose membranes. (B) Line blot and densitometric analysis: line blots (left) were incubated either with human patient serum (IFA titre ≥1:320, six samples) or control serum (six samples) of German healthy blood donors; densitometric analysis of reactivity is depicted by heat maps (right); numbers refer to the position of each protein on the line blot; the most reactive proteins are labelled. (C) Analysis of antigen reactivity in line blots: results for the most reactive antigens (BbadA, BbadB, Pap31, hypothetical protein B, porin B, and autotransporter E) are presented for serum samples with high anti-*Bartonella bacilliformis* IFA titres from patients with *B bacilliformis* infection (n=12; titres ≥1:320), samples with IFA titres lower than 1:320 (n=14), and control samples (n=96); values higher than a band intensity of 100 units are considered positive; respective p values between sample categories are given. IFA=immunofluorescence assay.

same cutoff value and including all Peruvian patient samples, sensitivity was 69·2% (18 of 26 samples; 48·2–85·7) and specificity was 100·0% (96 of 96; 96·2–100·0). A lowered OD_{450} cutoff value of 0·29 (best trade-off between specificity and sensitivity) resulted in 100·0% sensitivity (12 of 12; 73·5–100·0) and 94·8%

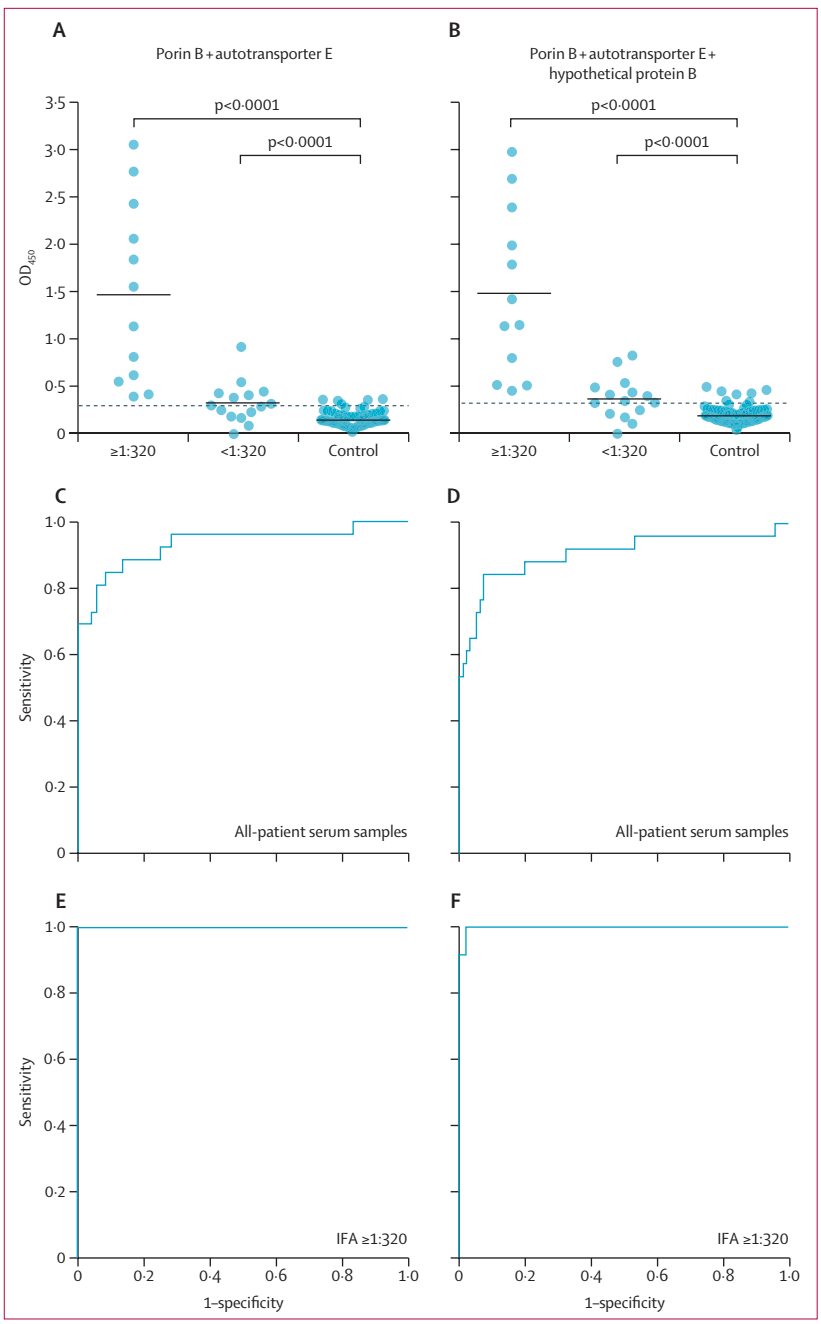

Figure 3: **Analysis of serum samples from Peruvian patients and German controls using *Bartonella bacilliformis* ELISA**

For the ELISA analysis (A, B), we used serum samples with high anti-*B bacilliformis* IFA titres from patients with *B bacilliformis* infection (n=12; ≥1:320), samples with an IFA titre lower than 1:320 (n=14), and control samples (n=96). OD_{450} values higher than 0·29 (A) or 0·34 (B) were considered positive. Respective p values between sample categories are given. We drew ROC curves using all serum samples from Peruvian patients (n=26; C, D) and using samples of Peruvian patients with an IFA titre of 1:320 or higher (n=12; E, F). Values are depicted in a ROC diagram according to sensitivity (true-positive rate) and 1 – specificity (false-positive rate). IFA=immunofluorescence assay. OD_{450}=optical density of 450 nm. ROC=receiver operating characteristic.

specificity (91 of 96; 88·3–98·3) for IFA-positive samples and 80·8% sensitivity (21 of 26; 60·7–93·5) and 94·9% specificity (91 of 96; 88·3–0·98·3) for all Peruvian patient samples. For a combination of porin B, autotransporter E, and hypothetical protein B, the AUC score was 0·9083 (0·8254–0·9911; figure 3D). At an OD_{450} cutoff of 0·34 (best trade-off between specificity and sensitivity), we found 100·0% sensitivity (12 of 12; 73·5–100·0) and 93·8% specificity (90 of 96; 86·9–97·7) for IFA-positive samples and 76·9% sensitivity (20 of 26; 56·4–91·0) and 93·8% specificity (90 of 96; 86·9–97·7) for all Peruvian patient samples. The ROC curves estimated by using exclusively IFA-positive patient samples yielded AUC scores of 1·0000 (1·0000–1·0000) for the ELISA with a combination of porin B and autotransporter E and 0·9983 (0·9936–1·0030) for the ELISA with porin B, autotransporter E, and hypothetical protein B (figure 3E, F). In-silico analysis of the genomic sequences of porin B, autotransporter E, and hypothetical protein B revealed a 99·8–100% homology of the deduced proteins from all analysed genomes of *B bacilliformis*, whereas homologies were much lower for *B ancashensis*, *B henselae*, and *B quintana* (35·2–35·8% for autotransporter E, 41·5–45·3% for porin B, and 30·4–41·8% for hypothetical protein B; appendix 3 p 11).

Discussion

By using a combination of an in-silico reverse vaccinology approach and large-scale immunoscreening of expression libraries, we identified 22 potentially immunodominant *B bacilliformis* proteins, of which 21 were recombinantly expressed. Seven proteins were not immunoreactive with serum samples of Peruvian patients. The remaining 14 seroreactive antigens were analysed with line blots. Autotransporter E, porin B, and hypothetical protein B were the best candidates for a diagnostic ELISA with a calculated sensitivity of 77% and specificity of 94% (combination of all three proteins) and sensitivity of 81% and specificity of 95% (combination of porin B and autotransporter E) for all Peruvian patient samples. These data suggest that the combination of porin B and autotransporter E is sufficient to detect human anti-*B bacilliformis* IgG-antibodies with ELISA.

In our study, we did a dynamic Vaxign analysis for *B bacilliformis* strains KC583 and KC584, resulting in 20 surface-exposed target proteins (flagellum-associated proteins, autotransporter proteins, porins, and other outer membrane proteins; table). The immunoreactivity is known for some homologous proteins from infections with other pathogens. Flagellum-associated proteins (eg, FlgE and p41 flagellin of *Borrelia burgdorferi*) are often reactive with serum samples of patients with Lyme borreliosis.[15,16] For unclear reasons, *B bacilliformis* FlgE and flagellin were not recognised by patient serum samples and FlgH showed only weak reactivity; therefore, they are not suitable serodiagnostic markers. Autotransporter proteins are the largest family of secreted proteins in Gram-negative bacteria and are known as key immunodominant vaccine components (eg, pertactin of *Bordetella pertussis*).[17,18] In our study,

autotransporter E showed the strongest serum reactivity among all tested proteins, suggesting its potential use in serodiagnostics and as a vaccine candidate. Trimeric autotransporter adhesins (TAAs) are often immunodominant and considered as promising vaccine candidates (eg, *Acinetobacter* trimeric autotransporter [Ata] and *Haemophilus influenzae* adhesin [Hia]).[19] TAAs of *B bacilliformis* were already considered as potential vaccine targets due to their homology to *Neisseria meningitidis* NadA, which is included in the Bexsero meningococcal group B vaccine (GlaxoSmithKline, Brentford, UK).[10,11] However, the reactivity of BbadA and BbadB with approximately 10% of the control serum samples excludes their potential use as specific immunoreactive proteins. Porins are not used in human vaccines, but their potential use was shown by porins from *Salmonella enterica* serotype Typhi that induced a strong antibody response in mice.[20] The Vaxign software predicted several porins (porin A, B, and C and Pap31) as immunodominant targets. Moreover, porin A, porin B, and Pap31 were recognised by serum samples of Peruvian patients and a detailed analysis underlined that porin B might represent a valuable serological marker and potential vaccine target. From the remaining outer membrane proteins (TonB-dependent receptor, LysM, DUF1561, and hypothetical protein B), only hypothetical protein B showed a significant difference in reactivity to control samples, and was thus included in our ELISA.

B bacilliformis outbreaks usually occur in Andean communities, but they are now expanding to adjacent areas.[21–23] The implementation of serodiagnostic tools allowing broad epidemiological monitoring of Carrión's disease and the availability of a vaccine would be highly desirable; however, only a few immunogenic targets have been analysed.[6,10] Novel sequencing techniques and bioinformatics have revolutionised clinical microbiology and vaccine development.[11] Recent approaches combine reverse vaccinology with comparative multigenome analyses to identify species-conserved antigens.[12] Accordingly, we included five genomic sequences of *B bacilliformis* in our strategy, revealing several genomic discrepancies (appendix 3 pp 9–10). In particular, the loss of the TAA BbadB in strain KC584 and the diversity of Pap31 restrict their use as diagnostic or vaccine targets.[24]

There are limitations in the application of reverse vaccinology approaches. When comparing several reverse vaccinology software programs against a set of confirmed bacterial protective antigens, 75% of these proteins were detected in silico in the best case.[25] To minimise risks of missing antigens, we combined the in-silico Vaxign prediction with an immunoscreening of *B bacilliformis* expression libraries,[26] resulting in the finding of immunodominant autotransporter E. Screening of the expression libraries with rabbit serum retrieved two further immunodominant antigens (GroEL and flagellin). The reasons for these discrepancies in antigen prediction might be caused by the search algorithms of the Vaxign software and by the different exposure routes to the pathogen (natural infection in humans, and immunisation with inactivated bacteria in rabbits).

Prior to this study, four immunoreactive proteins of *B bacilliformis* had been described (GroEL, LysM, Pap31, and SCS-α–SCS-β). In case of GroEL, specific antibodies were shown to inhibit GroEL-mediated mitogenicity on endothelial cells, suggesting its use as a vaccine compound.[27,28] Despite studies evaluating its serodiagnostic use, GroEL was not predicted by the Vaxign software nor was it detected with *B bacilliformis* genomic expression libraries.[6,27] LysM was shown to react with serum samples of patients diagnosed with bartonellosis.[29] This protein was also identified as a potential target protein by the Vaxign software and exhibited a good reactivity with patient serum samples (figure 1). However, LysM was also highly reactive with healthy donor serum samples (92%). Because of its reactivity with IgM and IgG antibodies, Pap31 was suggested as a potential seromarker for both acute and past *B bacilliformis* infections.[6,9] Surprisingly, our data did not confirm these previous findings when using linear epitopes in line blots because only two of 26 patient samples were reactive (figure 2), suggesting that Pap31 might be more reactive as a conformational epitope. An in-silico study published in 2020 described a high genetic variability of the *pap31* gene in *B bacilliformis* strains, which was supported by our findings (appendix 3 pp 9–10).[24] This variability, potentially involved in evading the host's immune response, restricts the use of Pap31 for further serodiagnostic and vaccination strategies. Finally, SCS-α and SCS-β subunits of the succinyl-CoA synthetase were previously identified as immunodominant proteins, but showed only moderate suitability in an ELISA-based approach.[6]

Clearly, our study is limited by the small number of included serum samples from patients with Carrión's disease, although the number of defined patient samples used herein is one of the highest reported. However, the use of these carefully collected serum samples of Peruvian patient was the only possible way to evaluate immunodominant targets, although they need to be validated in larger studies (eg, to assess potential cross-reactivities).[30] Further work will focus on the production of readily usable ELISA kits through a professional manufacturing platform, so that they can be used for seroepidemiological analyses in Peru and other South American countries. The results from such surveys will offer a much better epidemiological understanding of *B bacilliformis* infections and will lead to a solid knowledge base of which proteins might be included in a vaccine against *B bacilliformis* able to confer long-term immunity against Oroya fever.

Contributors

VAJK conceived the study. AAD, AW, WB, ES, AL, PT, and VAJK did experimental laboratory work, data collection and analysis, data

interpretation, and writing. TGS and PT did genome sequence analyses. PV was involved in the collection of patient serum samples. HGA and CU-G contributed with data interpretation and critical review of the manuscript. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication. AAD, AW, TGS, WB, and VAJK verified the underlying data of the study. All authors approved the final manuscript.

Declaration of interests

VAJK reports grants from the State of Hesse (LOEWE center DRUID [Novel Drug Targets against Poverty-Related and Neglected Tropical Infectious Diseases]) and from the Robert Koch-Institute (Berlin, Germany); and has received grants from Deutsche Forschungsgemeinschaft, Bundesministerium für Bildung und Forschung, and the EU, outside of the submitted work. VAJK, AAD, and AW are inventors on a pending patent (a method for inducing an immune response against *B bacilliformis* in a patient; application EP21158777.9). All other authors declare no competing interests.

Data sharing

Genomic information is deposited in public databases as indicated in the Methods section (NCBI sequence read archive accession number SRR13618139 for KC583 and SRR10344655 for KC584). Expression libraries, recombinant clones, and rabbit antibodies will be provided upon request to the corresponding author.

Acknowledgments

This work was supported by the LOEWE Center DRUID (project C2) and the Robert Koch-Institute, Berlin, Germany (*Bartonella* Consiliary Laboratory, 1369-354). The sequencing service was provided by the Norwegian Sequencing Centre, a national technology platform hosted by the University of Oslo (Oslo, Norway) and supported by the Functional Genomics and Infrastructure programmes of the Research Council of Norway. We thank Michael Schmidt and Erhard Seifried (German Red Cross blood service Baden-Württemberg-Hesse, Frankfurt, Germany) for providing blood donor serum samples and Eugenia Álvarez for the Quechua translation of the Summary.

References

1 Gomes C, Pons MJ, Del Valle Mendoza J, Ruiz J. Carrion's disease: an eradicable illness? *Infect Dis Poverty* 2016; **5:** 105.

2 Battisti JM, Lawyer PG, Minnick MF. Colonization of *Lutzomyia verrucarum* and *Lutzomyia longipalpis* sand flies (*Diptera: Psychodidae*) by *Bartonella bacilliformis*, the etiologic agent of Carrión's disease. *PLoS Negl Trop Dis* 2015; **9:** e0004128.

3 Minnick MF, Anderson BE, Lima A, Battisti JM, Lawyer PG, Birtles RJ. Oroya fever and verruga peruana: bartonelloses unique to South America. *PLoS Negl Trop Dis* 2014; **8:** e2919.

4 Pons MJ, Gomes C, Aguilar R, et al. Immunosuppressive and angiogenic cytokine profile associated with *Bartonella bacilliformis* infection in post-outbreak and endemic areas of Carrion's disease in Peru. *PLoS Negl Trop Dis* 2017; **11:** e0005684.

5 Knobloch J, Solano L, Alvarez O, Delgado E. Antibodies to *Bartonella bacilliformis* as determined by fluorescence antibody test, indirect haemagglutination and ELISA. *Trop Med Parasitol* 1985; **36:** 183–85.

6 Gomes C, Palma N, Pons MJ, et al. Succinyl-CoA synthetase: new antigen candidate of *Bartonella bacilliformis*. *PLoS Negl Trop Dis* 2016; **10:** e0004989.

7 Ellis BA, Rotz LD, Leake JAD, et al. An outbreak of acute bartonellosis (Oroya fever) in the Urubamba region of Peru, 1998. *Am J Trop Med Hyg* 1999; **61:** 344–49.

8 del Valle Mendoza J, Silva Caso W, Tinco Valdez C, et al. Diagnosis of Carrion's disease by direct blood PCR in thin blood smear negative samples. *PLoS One* 2014; **9:** e92283.

9 Angkasekwinai N, Atkins EH, Romero S, Grieco J, Chao CC, Ching WM. An evaluation study of enzyme-linked immunosorbent assay (ELISA) using recombinant protein Pap31 for detection of antibody against *Bartonella bacilliformis* infection among the Peruvian population. *Am J Trop Med Hyg* 2014; **90:** 690–96.

10 Henriquez-Camacho C, Ventosilla P, Minnick MF, Ruiz J, Maguiña C. Proteins of *Bartonella bacilliformis*: candidates for vaccine development. *Int J Pept* 2015; **2015:** 702784.

11 Pizza M, Scarlato V, Masignani V, et al. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science* 2000; **287:** 1816–20.

12 Maione D, Margarit I, Rinaudo CD, et al. Identification of a universal Group B streptococcus vaccine by multiple genome screen. *Science* 2005; **309:** 148–50.

13 Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30:** 2068–69.

14 He Y, Xiang Z, Mobley HLT. Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol* 2010; **2010:** 297505.

15 Jwang B, Dewing P, Fikrig E, Flavell RA. The hook protein of *Borrelia burgdorferi*, encoded by the *flgE* gene, is serologically recognized in Lyme disease. *Clin Diagn Lab Immunol* 1995; **2:** 609–15.

16 Luft BJ, Dunn JJ, Dattwyler RJ, Gorgone G, Gorevic PD, Schubach WH. Cross-reactive antigenic domains of the flagellin protein of *Borrelia burgdorferi*. *Res Microbiol* 1993; **144:** 251–57.

17 Wells TJ, Tree JJ, Ulett GC, Schembri MA. Autotransporter proteins: novel targets at the bacterial cell surface. *FEMS Microbiol Lett* 2007; **274:** 163–72.

18 Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D. Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* 2004; **68:** 692–744.

19 Thibau A, Dichter AA, Vaca DJ, Linke D, Goldman A, Kempf VAJ. Immunogenicity of trimeric autotransporter adhesins and their potential as vaccine targets. *Med Microbiol Immunol (Berl)* 2020; **209:** 243–63.

20 Secundino I, López-Macías C, Cervantes-Barragán L, et al. *Salmonella* porins induce a sustained, lifelong specific bactericidal antibody memory response. *Immunology* 2006; **117:** 59–70.

21 Silva-Caso W, Mazulis F, Weilg C, et al. Co-infection with *Bartonella bacilliformis* and *Mycobacterium* spp in a coastal region of Peru. *BMC Res Notes* 2017; **10:** 656.

22 Maco V, Maguiña C, Tirado A, Maco V, Vidal JE. Carrion's disease (*Bartonellosis bacilliformis*) confirmed by histopathology in the High Forest of Peru. *Rev Inst Med Trop São Paulo* 2004; **46:** 171–74.

23 Sanchez Clemente N, Ugarte-Gil C, Solorzano N, Maguiña C, Moore D. An outbreak of *Bartonella bacilliformis* in an endemic Andean community. *PLoS One* 2016; **11:** e0150525.

24 Ruiz J, Gomes C. In silico analysis of Pap31 from *Bartonella bacilliformis* and other *Bartonella* spp. *Infect Genet Evol* 2020; **84:** 104482.

25 Dalsass M, Brozzi A, Medini D, Rappuoli R. Comparison of open-source reverse vaccinology programs for bacterial vaccine antigen discovery. *Front Immunol* 2019; **10:** 113.

26 Miller L, Richter M, Hapke C, Stern D, Nitsche A. Genomic expression libraries for the identification of cross-reactive orthopoxvirus antigens. *PLoS One* 2011; **6:** e21950.

27 Knobloch J, Schreiber M. Bb65, a major immunoreactive protein of *Bartonella bacilliformis*. *Am J Trop Med Hyg* 1990; **43:** 373–79.

28 Minnick MF, Smitherman LS, Samuels DS. Mitogenic effect of *Bartonella bacilliformis* on human vascular endothelial cells and involvement of GroEL. *Infect Immun* 2003; **71:** 6933–42.

29 Padmalayam I, Kelly T, Baumstark B, Massung R. Molecular cloning, sequencing, expression, and characterization of an immunogenic 43-kilodalton lipoprotein of *Bartonella bacilliformis* that has homology to NlpD/LppB. *Infect Immun* 2000; **68:** 4972–79.

30 Knobloch J. Analysis and preparation of *Bartonella bacilliformis* antigens. *Am J Trop Med Hyg* 1988; **39:** 173–78.