The identification of novel immunogenic antigens as potential *Shigella* vaccine components

Ruklanthi de Alwis, Li Liang, Omid Taghavian, Emma Werner, Hao Chung The, Trang Nguyen Hoang Thu, Vu Thuy Duong, D. Huw Davies, Philip L. Felgner and Stephen Baker

**Abstract**

**Background:** *Shigella* is a major diarrheal pathogen for which there is presently no vaccine. Whole genome sequencing provides the ability to predict and derive novel antigens for use as vaccines. Here, we aimed to identify novel immunogenic *Shigella* antigens that could serve as *Shigella* vaccine candidates, either alone, or when conjugated to *Shigella* O-antigen.

**Methods:** Using a reverse vaccinology approach, where genomic analysis informed the *Shigella* immunome via an antigen microarray, we aimed to identify novel immunogenic *Shigella* antigens. A core genome analysis of *Shigella* species, pathogenic and non-pathogenic *Escherichia coli*, led to the selection of 234 predicted immunogenic *Shigella* antigens. These antigens were expressed and probed with acute and convalescent serum from microbiologically confirmed *Shigella* infections.

**Results:** Several *Shigella* antigens displayed IgG and IgA seroconversion, with no difference in sero-reactivity across by sex or age. IgG sero-reactivity to key *Shigella* antigens was observed at birth, indicating transplacental antibody transfer. Six antigens (FepA, EmrK, FhuA, MdtA, NlpB, and CjrA) were identified in in vivo testing as capable of producing binding IgG and complement-mediated bactericidal antibody.

**Conclusions:** These findings provide six novel immunogenic *Shigella* proteins that could serve as candidate vaccine antigens, species-specific carrier proteins, or targeted adjuvants.

**Background**

*Shigella* is the causative agent of shigellosis, a severe acute gastrointestinal infection that frequently presents as bloody diarrhea, fever, and severe abdominal pain [1]. In 2016, *Shigella* was estimated to cause > 250 million cases and > 200,000 deaths globally [2]. Higher income countries experience *Shigella* infections among travelers, aging populations, deployed military personnel, and men who have sex with men (MSM) [2, 3]. However, the preponderance of the *Shigella* disease burden is in children aged under 5 years residing in low-middle income countries (LMICs). Infection in this vulnerable group can also result in significant long-term consequences such as severe stunting and impaired cognitive development [2, 4]. The global epidemiology of *Shigella* is worsened by the emergence and spread of multi- and extensively drug resistant (MDR and XDR) variants, making infections increasingly difficult to treat [5]. The principal method of *Shigella* prevention has been improvements in water, sanitation, and hygiene (WASH) [6]. However, due to the low infectious dose, the standard of WASH required to break transmission is difficult to attain in many LMICs [7]. Furthermore, recent application of molecular techniques to
identify 

Shigella infections found a severe underestimation of the global Shigella burden [8, 9], highlighting the need for new low-cost prevention techniques.

There is currently no licensed vaccine against Shigella [10]. However, studies in animals and controlled human infection models (CHIMs) have shown that protection through immunization is feasible [11, 12]. Natural disease epidemiology in humans and non-human primate infection studies show complete protection from re-infection with a homologous Shigella species. Long-term homologous protection has been attributed to serotype-specific systemic (serum IgG) and mucosal (IgA) antibody responses [12, 13]. The most immunodominant target of the Shigella IgG and IgA response is the O-antigen component of lipopolysaccharide (LPS) [14, 15]. LPS/O-antigen-specific antibodies elicit protection through antibody-mediated opsonization, phagocytosis, and intracellular cytotoxicity [13]. However, antibodies against Shigella O-antigen are highly specific for the infecting species only [11, 13], and do not provide protection against heterologous Shigella species. Since the Shigella genus consists of four species and >50 serotypes, a lack of cross-protection against heterologous species and serotypes poses a major challenge for vaccine development [16]. This challenge may be overcome by a vaccine that elicits either a broadly reactive immune response or numerous species-specific responses against the globally dominant Shigella species (i.e., S. flexneri and S. sonnei) [17].

The primary strategy of developing an efficacious Shigella vaccine has been to elicit antibody responses targeting Shigella O-antigen [10]. Live-attenuated vaccines with genetically attenuated Shigella [18-20], or the expression of Shigella O-antigen on live-attenuated vectors [21], have been shown to induce good antibody responses against O-antigen. Multivalent killed-vaccines also induce high titers of serum IgG and mucosal IgA targeting Shigella O-antigens and have shown protection in early clinical development [22, 23]. Recombinant forms of Shigella O-antigen have also been pursued as vaccine candidates [10], with a Shigella O-antigen conjugated to a carrier protein [24-26], which engages T cell help and produces a longer lasting antibody response to the polysaccharide antigen [27]. Various immunogenic proteins, such as the toxins from other pathogens, have been used as carrier proteins [28, 29]. However, protein antigens from Shigella have not been evaluated as a carrier protein for Shigella O-antigen.

Whole genome sequencing provided the ability to predict and derive novel antigens for use as vaccines, and this approach ultimately gave rise to the meningococcal B vaccine [30-32]. Further technological advances in immunology and protein engineering to study the interaction of pathogens with the immune system can aid in reverse engineering of protective immunogens [33-35]. Here, we aimed to identify novel immunogenic Shigella antigens that could serve as Shigella vaccine candidates, either alone, or when conjugated to Shigella O-antigen. Therefore, we conducted immunogen prediction using bioinformatic analysis, then created a protein microarray of predicted immunogenic Shigella antigens. These expressed antigens were screened for immunogenicity using polyclonal antibodies from patients who recovered from confirmed Shigella infections, to identify a novel set of proteins which may facilitate the development of novel Shigella vaccines.

**Methods**

**Ethics**

Human serum samples for the purposes of this investigation were collected from an observational study of children with diarrheal disease and a cohort study of healthy infants, both conducted in Ho Chi Minh City (HCMC), Vietnam [36, 37]. Both studies were approved by the institutional review boards of collaborating institutions HCMC and the Oxford Tropical Research Ethics Committee (OxTREC No. 1045-13) in the UK. Written informed consent by a legal guardian was a prerequisite for enrolment into the studies.

**Serum samples**

Paired acute-convalescent serum samples were collected as a component of a prospective, observational, multicenter, cross-sectional study conducted in HCMC, Vietnam. The clinical and microbiological data from this study has been published previously [37]. For this investigation, we utilized acute serum samples collected from microbiologically confirmed Shigella (n = 33) or Salmonella (n = 24) cases when they first presented at hospital with acute diarrheal disease (patients bled prior to diagnostic testing) (Table 1). Convalescent (follow-up) serum samples collected at a follow-up visit, 4 weeks (+1 week) after being enrolled in the study. Umbilical cord blood was collected from a large prospective birth cohort study, where healthy pregnant mothers who visited Hung Vuong Obstetrics Hospital in HCMC, Vietnam, were recruited prior to birth and cord blood sampled after delivery [36, 38]. Serum extracted from maternal blood during pregnancy has previously been subjected to ELISA to measure S. sonnei O-antigen IgG [39]; the cord blood samples screened here were from mothers with high (n = 45) and low (n = 40) S. sonnei O-antigen IgG titers (Table 1).

**Bioinformatic analysis**

The complete chromosomal sequences of 10 Shigella and 47 Escherichia coli (E. coli) were retrieved from GenBank (accessed in July 2014 using an in-house script
Table 1: Summary of genomes compared, serum samples tested, Shigella antigens studied, and in vivo immunogenicity testing conducted in present “reverse vaccinology 2.0” study

| Category                | Number |
|-------------------------|--------|
| Genome comparison       |        |
| S. sonnei               | 2      |
| S. flexneri             | 4      |
| S. dysenteriae          | 2      |
| S. boydii               | 2      |
| Pathogenic E. coli      | 32     |
| Non-pathogenic E. coli  | 15     |
| Number of genomes       | 57     |
| No. of subjects         |        |
| Shigella patients       | 33     |
| S. flexneri             | 2      |
| S. sonnei               | 31     |
| Salmonella patients     | 24     |
| Newborn (cord serum)    | 85     |
| Low LPS titer           | 40     |
| High LPS titer          | 45     |
| Antigen microarray      |        |
| Number of expressed antigens | 234  |
| S. flexneri-specific    | 22     |
| S. sonnei-specific      | 8      |
| Shigella orthologs      | 102 (×2) |
| Immunogenicity in vivo  |        |
| Number of antigens tested | 8    |

Proteins selected through the bioinformatics pipeline are shown in Additional file 5 Table S3. These targets were expressed, printed, and probed as described previously for other protein microarray projects [44–46]. Briefly, the corresponding coding sequences from selected Shigella proteins were amplified, cloned into a pXT7 vector, and expressed using a high-throughput in vitro transcription/translation (IVTT) E. coli system (BiotechRabbit, GmbH). Controls lacking DNA were included to account for background reactivity with E. coli, where IVTT was conducted without plasmid DNA. Expressed Shigella antigens from IVTT reactions were printed onto nitrocellulose-coated glass GraceBio slides using an Omni Grid 100 microarray printer (Genomic Solutions). LPS from Shigella (Sigma) was also printed on the microarray slides to act as positive control. Slides (with E. coli lysate (McLab) at a final concentration of 1 mg/ml) were probed with human serum (diluted 1:200), followed by biotin-conjugated secondary antibodies specific for human chromosomal genetic makeup. Hence, the number of proteins that fulfilled the above criteria (i.e., in groups 1 to 3) was limited (all are shown in Additional file 3 Table S2) and was not sufficient to develop downstream immunogenic assays.

Proteins which may show potential immunogenicity (i.e., due to their location on the outer membrane, cell wall, possessing transmembrane domain(s), or possessing a signal peptide) were retained. Unannotated proteins identified by Pfam as bacteriophage-related were discarded.

### Protein microarray

Proteins selected through the bioinformatics pipeline are shown in Additional file 5 Table S3. These targets were expressed, printed, and probed as described previously for other protein microarray projects [44–46]. Briefly, the corresponding coding sequences from selected Shigella proteins were amplified, cloned into a pXT7 vector, and expressed using a high-throughput in vitro transcription/translation (IVTT) E. coli system (BiotechRabbit, GmbH). Controls lacking DNA were included to account for background reactivity with E. coli, where IVTT was conducted without plasmid DNA. Expressed Shigella antigens from IVTT reactions were printed onto nitrocellulose-coated glass GraceBio slides using an Omni Grid 100 microarray printer (Genomic Solutions). LPS from Shigella (Sigma) was also printed on the microarray slides to act as positive control. Slides (with E. coli lysate (McLab) at a final concentration of 1 mg/ml) were probed with human serum (diluted 1:200), followed by biotin-conjugated secondary antibodies specific for human chromosomal genetic makeup. Hence, the number of proteins that fulfilled the above criteria (i.e., in groups 1 to 3) was limited (all are shown in Additional file 3 Table S2) and was not sufficient to develop downstream immunogenic assays.
IgM, IgG, and IgA (Jackson Immunoresearch). Binding antibody was detected using streptavidin-conjugated SureLightH P-3 (Columbia Biosciences), measured using Perkin Elmer ScanArray Express HT microarray scanner. Spot intensities were quantified using the ScanArray software.

Data analysis
Fold-over-control (FOC) normalizations were conducted to reduce assay to assay variation by dividing the mean spot intensities for each antigen by the intensity for the no DNA control IVTT. Positive serum reactivity to an antigen was defined as a FOC > 2 (i.e., > 2-fold increase in the mean intensity over the background control). Log2-transformed FOC values from paired acute and convalescent samples were compared using a Bayes regularized t test adapted from Cyber-T for protein arrays [47–49]. p values were subjected to Benjamini and Hochberg (BH) correction to control for false discovery rate [50]. Data were graphed using the R statistical software (http://www.r-project.org) and packages “Superheat,” “ggplot2,” “rgl,” and “frmsb.”

Protein immunization
His-tagged variants of selected proteins (NmpC, FepA, HtrB, EmrK, NlpB, FluA, Cjra, and MdtA) (Additional file 6 Table S4) were successfully expressed in a BacPowerTM E. coli protein expression system and purified using nickel affinity chromatography (GenScript Limited, Hong Kong). Four months old, male New Zealand rabbits (n = 2 per protein) were immunized with 0.2 mg of the successfully expressed and purified protein, and serum drawn at 1-week post-immunization of the third dose. The immunogenicity of each protein was assayed by testing the pre-immune and post-immune rabbit sera for sero-positivity using indirect enzyme-linked immunosorbent assay (ELISA) and immunoblot. For ELISA, plates were coated with protein at 4 μg/ml, blocked, incubated with sera (at 1 mg/ml IgG concentration), and detected using anti-rabbit IgG Fc-HRP secondary antibody. For western blots, 50 ng/well of purified proteins was run on SDS-PAGE, transferred to nitrocellulose membrane, blocked, probed with pre-immune and post-immune rabbit sera, and detected with goat anti-rabbit IgG-IRDye800cw secondary antibody.

Serum bactericidal assay
Purified serum antibody from immunized rabbits was tested for serum bactericidal activity (SBA) against S. flexneri 2a (strain EG 0478), S. sonnei (strain DE 1404, containing a cat chloramphenicol resistance gene on the virulence plasmid), and S. Typhimurium (strain ATCC 14028) using a previously described SBA protocol [51, 52]. Heat-inactivated sera were serially diluted from 50 to 0.07 μg/ml, then combined with bacteria (250 CFU/well) and 5 μl of baby rabbit complement and incubated at 37 °C for 90 min. Viable bacterial cells were estimated at time 0 min (T0) and at 90 min post-incubation (T90) by plating on nutrient agar plates. Bactericidal activity was calculated as a ratio of CFU at T90 over T0, from which SBA titers were estimated at 50% bactericidal activity. Convalescent immune serum from a confirmed Shigella-infected patient was used as a positive control [38]. All serum samples were tested in triplicate and the SBA titers averaged. The SBA assays with S. sonnei DE 1404 were performed with and without the supplementation of 10 μg/ml chloramphenicol. S. sonnei has the propensity to lose the virulence plasmid and O-antigen culture during culture, and the addition of 10 μg/ml chloramphenicol was to ensure the maintenance of plasmid and O-antigen during the SBA assay via the added cat gene. These data were compared to assess potential killing differences between plasmid+ and plasmid– organisms.

Results
Bioinformatic analysis identifies potential immunogenic Shigella core antigens
Genomic comparison of Shigella and E. coli core genomes was conducted with the aim of selecting both species-specific and species cross-reactive Shigella proteins common to the most globally dominant species, S. flexneri and S. sonnei. Protein sequences were extracted from the annotated chromosomes of various Shigella species (n = 10), pathogenic (n = 32) and non-pathogenic (n = 15) E. coli (Table 1 and Additional File 2 Table S1), and a list of potentially immunogenic antigens was selected using bioinformatic comparison, a list of potentially immunogenic antigens was selected (Table 2 and Additional File 5 Table S3). Our analysis was restricted to chromosomal proteins to identify novel immunogenic targets, as several proteins on Shigella virulence plasmids have already been extensively studied for their immunogenicity [53]. The analysis identified 22 S. flexneri-specific, 8 S. sonnei-specific, and 2 Shigella-specific proteins (IpaH3.1, IpaH4.5). Another 100 potentially immunogenic orthologs, from both S. sonnei and S. flexneri, were further included to expand the downstream immunogenic assays. Shigella LPS (O-antigen) was included as a positive control. This resulted in a total of 235 proteins that were expressed in vitro, and successfully printed on an antigen microarray for downstream analysis (Additional File 5 Table S3).

Antigen microarray reveals broad seroconversion following Shigella infections
The Shigella antigen microarray allowed us to assess the IgM, IgA, and IgG responses against the selected antigens following symptomatic Shigella infections. The antigen microarray was probed for sero-reactivity with pairs of
Shigella infection led to widespread seroconversion in all measured antibody isotypes (IgG, IgA, and IgM) among individuals and across multiple tested antigens, as observed by the increase in sero-reactivity (i.e., measured as fold change over control greater than 2, FOC > 2) from acute to early convalescence (Fig. 1). IgG sero-reactivity (i.e., FOC > 2) analysis showed that at least one individual (3% of all patients) during acute and convalescent phase reacted to a maximum of 35 and 166 antigens, respectively (Fig. 1a). Additionally, 50% of the individuals (i.e., 17 patients) in acute and early convalescent phase sero-reacted to IgG to 2 and 7 antigens, respectively (Fig. 1a).

The IgA responses were markedly higher than the IgG responses, with at least one individual generating a detectable IgA response to 209 and 227 antigens during acute disease and convalescence, respectively (Fig. 1b). We also observed that >50% of cases in acute and convalescence produced a reactive IgA response to 3 and 44 antigens, respectively. IgM against the antigens also detected in at least 1 patient and 9 to 38 reactive antigens in >50% of patients (Fig. 1c).

We additionally observed that antibody responses were on average higher in patients with inflammatory (i.e., bloody diarrhea) as compared to non-inflammatory (i.e., Watery diarrhea) disease (as can be seen in the scatter plots comparing mean IgG, IgA, and IgM responses at early convalescence, Additional File 7 Fig. S1). As a control for the assay, the antigen array was probed with paired acute and convalescent serum (n = 24) from diarrheal patients infected with an alternative genus of diarrheal pathogen, Salmonella. Notably, there were no significant increases in IgG, IgA, or IgM responses between the acute and convalescence in Salmonella-infected diarrheal cases, indicating that the antibody reactivity observed with the serum from the Shigella-infected patients was specific to Shigella (Additional File 8 Fig. S2).

**Table 2** Description of the protein subsets of interests studied during chromosomal genome comparison between S. sonnei, S. flexneri, and pathogenic and non-pathogenic E. coli

| Chromosome-encoded proteins | Protein subset of interest | Excluded proteins | Primary result | Following first filtrationa | Final selected proteinsb |
|-----------------------------|----------------------------|-------------------|---------------|---------------------------|--------------------------|
| Proteins found in S. sonnei | 1                          | Excluding proteins found in non-pathogenic E. coli | 63            | 37                        | 11                       |
|                             | 2                          | Excluding proteins found in any other Shigella or E. coli spp. | 25            | 16                        | 8                        |
| Proteins found in S. flexneri | 3                          | Excluding proteins found in non-pathogenic E. coli | 48            | 29                        | 20                       |
|                             | 4                          | Excluding proteins found in any other Shigella or E. coli spp. | 25            | 25                        | 18                       |
| Proteins shared between S. sonnei and S. flexneri | 5                          | Excluding proteins found in non-pathogenic E. coli chromosomes | 8            | 1                        | 1                        |
| Other predicted immunogenic Shigella proteins | 6                          |                    |               |                           | 114                      |

*a*First filtration: removal of insertion sequence (IS) elements, transposases, transposons, and bacteriophage-related proteins etc.

*b*Second filtration: removal of unspecific proteins (according to BLASTn), cytoplasmic and hypothetical proteins with bacteriophage domain
observed a statistically significant ($p < 0.05$) increase in mean sero-reactivity between acute to early convalescence in all three antibody isotypes (Fig. 2). Furthermore, > 50% of the *Shigella*-infected cases at convalescence (or follow-up) had positive IgG, IgA, and IgM (i.e., FOC > 2) responses to all selected antigens, with the exception of SSON_mdtA and SSON_fhuA (Fig. 3).

*Shigella* poses the greatest health burden in children under the age of 2 years; therefore, an effective vaccine against *Shigella* needs to induce a protective immune response in young children. Hence, we assessed the antibody responses to the 12 selected antigens in children aged < 2 years ($n = 8$) in comparison to those aged > 2 years ($n = 26$). At early convalescence, there were no significant differences in both the mean antibody (IgG and IgA) sero-reactivity (Additional File 9 Fig. S3A and B) or the percentage of cases that showed positive antibody (IgG and IgA) responses (Fig. S3), when compared between age groups. These data demonstrate that younger children generate antibody responses to the 12 selected antigens. Additionally, we also compared sero-reactivity between male and females; all selected antigens...
elicited similar antibody responses (IgG and IgA) at early convalescence in both female ($n=20$) and male ($n=14$) patients with diarrhea (Fig. S3Ca and D).

Transplacentally transferred antibodies

A potential mechanism for protecting young children and neonates is the prenatal vaccination of pregnant mothers [54]. Prenatal vaccines require the mother to mount a protective IgG response that can be efficiently transferred transplacentally to the unborn fetus [54]. Therefore, we measured the transplacental transfer of IgG as an indirect assessment of whether the selected immunogenic antigens could serve as potential prenatal vaccine candidates. We compared the IgG sero-reactivity at birth (cord blood serum) between infants from mothers with high ($n=45$) and low ($n=40$) antibody titers (to Shigella O-antigen) (Fig. 3). In general, infants from mothers with high antibody titers demonstrated greater IgG sero-reactivity against tested Shigella antigens in comparison to infants born to mothers with low IgG titers (Fig. 3). Specifically, with the exception of antigen SSON_mdtA, IgG sero-reactivity for the remaining 11 selected antigens was
significantly greater ($p < 0.05$) in infants born to mothers with high antibody titers, than those with low antibody titers (Fig. 3c). These data show that 11/12 selected *Shigella* antigens are capable of generating IgG antibodies that can successfully transfer across the placenta in a concentration-dependent manner and may serve as good candidates for prenatal *Shigella* vaccines.

**Immunogenic testing of *Shigella* antigens**

We then proceeded to conduct preclinical testing of the top sero-reactive antigens. A BacPowerTM *E. coli* protein expression system was used to express SF_nmpC, SF_fepA, and SSON_htrB, SSON_emrK, SSON_nlpB, SSON_fhuA, SSON_cjrA, SSON_mdtA, SSON_sbmA, SSON_mviN, SSON_pldA, and SSON_3803. We were unsuccessful in expressing SSON_sbmA, SSON_mviN, SSON_3803, and SSON_pldA using this system (Additional File 6 Table S4). Therefore, we proceeded to test the remaining eight antigens (i.e., SF_nmpC, SF_fepA, and SSON_htrB, SSON_emrK, SSON_nlpB, SSON_fhuA, SSON_cjrA, and SSON_mdtA) in rabbits to ensure they could induce an antibody response when immunized as recombinant antigens. Rabbits were immunized separately with the eight antigens, and the serum was screened to measure binding IgG antibodies. All eight of the purified antigens induced robust IgG responses in vivo, which was
detectable by immunoblot and ELISA (Fig. 4). We next tested the immunized rabbit sera for the potential to induce bacterial killing in an antibody-dependent complement-mediated serum bactericidal assay. We measured the killing potential of the purified rabbit serum against *S. flexneri*, *S. sonnei*, and *Salmonella Typhimurium* to ensure that the serum bactericidal activity (SBA) induced by the immunized antigens was *Shigella* specific. The SBA assay was unable to detect bactericidal activity below 50 μg/ml against *S. sonnei* (Table 1 and Additional File 10 Table S5) whether grown with or without chloramphenicol, which was used to maintain the virulence plasmid and O-antigen cluster in *S. Sonnei* due to the plasmid containing a *cat* resistance gene. However, six of the eight *Shigella* immunogens (i.e., SF_fepA, SSON_cjrA, SSON_emrK, SSON_fhuA, SSON_mdtA, and SSON_nlpB) induced antibody responses with strong serum bactericidal activity against *S. flexneri* (i.e., 50% SBA titer < 50 μg/ml) (Table 1); the lowest 50% SBA titer was induced by the immunogen SSON_cjrA (Table 3). Our results show that the antigen SSON_cjrA was the most potent inducer of bactericidal antibodies against *Shigella* and the best antigen to take forward into future testing as a potential vaccine antigen.

**Discussion**

In the current study, we exploited reverse vaccinology 2.0 to integrate both comparative genomics and human immuno-proteome analysis to identify novel immunogenic chromosomal *Shigella* proteins. Genomic comparison and bioinformatic analysis of 57 *Shigella* and *E. coli* genomes allowed us to narrow down to 235 predicted immunogenic antigens. The predicted antigens were then expressed and printed onto a microarray, probed with a panel of sera from *Shigella*-infected individuals, to narrow the selection to 12 highly sero-reactive antigens. We confirmed that antibody responses to these 12 antigens were similar across sex (i.e., males and females) and two age groups (i.e., < 2 years and > 2 years). Using cord blood samples, we additionally observed that IgG responses to 11 of these 12 antigens could be transmitted transplacentally, hence suggesting the possible application of these antigens as prenatal vaccine candidates. Among the 12 antigens, 8 were successfully expressed as recombinant proteins. Six of these antigens were both immunogenic in animal models and generated functionally protective antibody responses against *Shigella*. These were SF_fepA, SSON_cjrA, SSON_emrK, SSON_fhuA, SSON_mdtA, and SSON_nlpB.

**Table 3** Immunization with highly reactive *Shigella* proteins elicits protective antibody responses. Rabbits were immunized with 0.2 mg of *Shigella* proteins, and sera (at 1 week post-immunization) were tested in vitro for serum bactericidal activity (SBA) against *Shigella flexneri*, *Shigella sonnei*, and *Shigella Typhimurium*. SBA titers presented below are estimated at 50% bactericidal activity.

| Post-immunization SBA titer (μg/ml) | SBA titer (dilution) |
|-------------------------------------|---------------------|
| fepA      | nmpC    | cjrA    | emrK   | fhuA   | htrB   | mdtA   | nlpB   | *Shigella*-immune sera |
|---|---|---|---|---|---|---|---|---|
| *Shigella flexneri*     | 0.4 | > 50 | 0.0046 | 0.7 | 4.8 | > 50 | 10.2 | 0.92 | 44,321 |
| *Shigella sonnei*       | > 50 | > 50 | > 50 | > 50 | > 50 | > 50 | > 50 | > 50 | 29,561 |
| *Salmonella Typhimurium* | > 50 | > 50 | > 50 | > 50 | > 50 | > 50 | > 50 | > 50 | > 50 |
and SSON_nlpB, with SSON_cjrA being the most immunogenic in terms of eliciting antibody-mediated bactericidal responses. Such bactericidal effects against S. sonnei were not observed in vitro, probably due to the protection against complement-mediated killing afforded by its high molecular weight capsule [55]. However, the immunogenicity of these antigens suggests that during S. sonnei infections, the capsule can be modulated to expose these functional proteins in vivo. This highlights the complex pathogenesis of S. sonnei and the difficulty in developing a suitable vaccine candidate.

The biological functions of these six Shigella immunogenic proteins have been characterized previously. FepA and FhuA serve as outer membrane proteins which bind and transport siderophores (ferric enterobactin and ferrichrome, respectively) [56, 57]. CjrA shares substantial homology to Pseudomonas aeruginosa PhuW and potentially acts to sequester iron from heme, the most common iron source in mammals [58]. EmrK and MdtA are subunits of the multidrug efflux pump EmrKY and MdtABC, respectively [59, 60], which contribute to resistance to bile salt and antimicrobials in E. coli. In addition, EmrK has been shown to confer Shigella survival in infected macrophages, facilitating its invasive pathogenesis in the human host [59]. NlpB forms part of the outer membrane protein (OMP) assembly complex, which assembles and inserts beta-barrel proteins into the outer membrane [61, 62]. The immunogenicity of these proteins points to their potentially high expression during Shigella infections, concurring with the survival strategies of pathogenic bacteria. Particularly, within-host iron is key to bacterial replication, and the ability to sequester and transport host iron is pivotal to the pathogenesis of Klebsiella pneumoniae [63] and Staphylococcus aureus [64].

None of the Shigella antigens (i.e., FepA, CjrA, EmrK, FhuA, MdtA, and NlpB) identified as immunogenic in the present study has been previously characterized for immunogenicity either in the context of natural Shigella infections or vaccination. Five of these antigens (i.e., FepA, CjrA, EmrK, FhuA, MdtA, and NlpB) are conserved in pathogenic E. coli, but they have not been tested for immunogenicity following pathogenic E. coli infection either. However, the presence of the genetic cluster cjrABC-senB has been previously linked to uropathogenic E. coli [65]. The immunogenicity of these six proteins may be predictable, since they are either surface exposed outer membrane or periplasmic proteins. Outer membrane particles of Shigella (Generalized Modules of Membrane Antigens—GMMA) have historically been thought to be highly virulent and immunomodulatory and are currently being developed as a vaccine immunogen [66]. GMMA-based Shigella vaccine, 1790GAHB, was shown to be immunogenic in human clinical trials [67]. Immunogenicity of GMMA vesicles has been attributed to the presence of LPS. However, proteomic analysis of the GMMA detected FepA and FhuA [68]. Based on our findings, it is plausible that other Shigella antigens, such as those identified in our current study, are may be partly responsible for the immunogenic properties of GMMA.

Development of a vaccine against Shigella faces many challenges, including the ability to protect against multiple Shigella species and to raise sustained mucosal immunity [10]. Fortunately, all the antigens, with the exception of CjrA, are conserved between S. flexneri and S. sonnei. Therefore, these antigens could be used to create a vaccine that protects against the two globally dominant Shigella species, which accounted for almost 90% of all Shigella cases in the Global Enteric Multicenter study (GEMS) [17]. Furthermore, in addition to strong IgG responses, all six antigens raised significant IgA responses, which is the dominant immunoglobulin at the mucosa. IgA-mediated protection has been explained by both preventing Shigella infection of host cells and downregulating inflammation and intestinal tissue pathology at infected sites [69, 70]. Although the current study measured IgA in serum, it has been previously shown that Shigella-specific serum IgA positively correlates with mucosal IgA in the stool [14]. Additionally, we observed that IgG specific to the six immunogenic Shigella antigens were capable of transplacental transfer, indicating that the antigens could additionally serve as prenatal vaccine candidates to protect neonates.

Conclusions

Shigella infections cause over a quarter of a billion gastrointestinal infections cases globally per annum [2]. Despite the high public health burden, there is currently no licensed vaccine available to prevent Shigella diseases. At present, LPS is a key antigen for the development of a vaccine against Shigella [10, 71]. Here, we identified six novel immunogenic Shigella proteins that could serve as additional vaccine candidates or could be conjugated to O-antigens to provide some cross-protection. Future Shigella challenge studies in animal models or human controlled infection models are needed to test the potency of these identified six antigens as vaccine candidates alone or as new generation glycoconjugates.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13073-020-00824-4.

Additional file 1: Supplementary file 1. In-house script to retrieve complete Shigella and Escherichia coli chromosomal sequences from GenBank (accessed July 2014).

Additional file 2: Table S1. List of Shigella and E. coli strains used for the bioinformatics genome comparison.
Additional file 3: Table S2. Identified protein within subsets of interest during chromosomal genome comparison between S. sonnei, S. flexneri, pathogenic and non-pathogenic E. coli.

Additional file 4: Supplementary file 2. Script used to retrieve annotation and protein sequences from the input Shigella genomes.

Additional file 5: Table S3. List of Shigella proteins and antigens that were successfully tested on the Immunoassay protein microarray.

Additional file 6: Table S4. The recombinant expression of highly reactive Shigella proteins for in vivo immunogen testing.

Additional file 7: Figure S1. No statistically significant seroconversion to Shigella antigens detected in diarrheal patients during Salmonella infections.

Additional file 8: Figure S2. Antibody responses to natural Shigella infections are influenced by severity of disease.

Additional file 9: Figure S3. Convalescent antibody responses to top reactive Shigella antigens by age and sex.

Additional file 10: Table S5. Comparison of SBA titre from Rabbits grown with and without chloramphenicol (50mg/L).

Abbreviations

MSM: Men who have sex with men; LMICs: Low-middle income countries; MDR and XDR: Multi- and extensively drug-resistance; WASH: Water, sanitation, and hygiene; CHIMs: Controlled human infection models; LPS: Lipopolysaccharide; HCMC: Ho Chi Minh City; IVTT: In vitro transcription/translation; FOC: Fold-over-control; BH: Benjamini and Hochberg correction; ELISA: Enzyme-linked immunosorbent assay; SBA: Serum bactericidal activity; OMP: Outer membrane protein; GMMA: Generalized Modules of Membrane Antigens; GEWS: Global Enteric Multicenter study

Acknowledgements

The authors would like to acknowledge and be grateful for all the patient participants, infants, and their families that made this study possible. We also acknowledge the study nurses at Hung Vuong Hospital where the participants were recruited. We thank Corinne N. Thompson whose Shigella O-antigen characterization data of maternal sera were utilized in this study.

Authors’ contributions

Conceptualization: SB. Formal analysis: RdA, LL, EW, HCT, and DHD. Provided samples: VTD. Methodology: RdA, LL, OT, EW, HCT, TNHT, VTD, DDH, and PLF. Writing original draft: RdA and SB. Review and editing: RdA, EW, HCT, DHD, and SB. All authors read and approved the final manuscript.

Funding

This work was supported by a Wellcome senior research fellowship to Stephen Baker to (215151/Z/19/Z). DTP is funded as leadership fellow through the Oak Foundation. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Availability of data and materials

All data are presented in the manuscript (in the supplementary information), and raw data from gene selection from comparative genomics, signals from antigen array, and serum bactericidal data are freely available at https://doi.org/10.5281/zenodo.4320935 [72].

Ethics approval and consent to participate

Human serum samples for the purposes of this investigation were collected from an observational study of children with diarrheal disease and a cohort study of healthy infants, both conducted in Ho Chi Minh City (HCMC), Vietnam. The observational study that collected acute and convalescent blood samples from pediatric diarrheal patients was approved by the institutional review boards (IRB) of three local hospitals in HCMC and the Oxford Tropical Research Ethics Committee (OxTREC No. 1045-13) in the UK. Written informed consent by a legal guardian was obtained prior to enrolment into this observational study. The pediatric cohort study from which the current investigation accessed umbilical cord blood was approved by the local IRB of four local hospitals in HCMC and OxTREC (No. 02-09). Written informed consent was obtained from pregnant mothers prior to enrolment into the cohort. This study conforms to the ethical guidelines of the Helsinki Declaration. All animal work for this study was conducted by GenScript Biotech Corporation under the auditing, monitoring, and approval of GenScript’s Institutional Animal Care and Use Committee (IACUC).

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

1. Program in Emerging Infectious Diseases, Duke-NUS Medical School, Singapore, Singapore. 2. Viral Research and Experimental Medicine Centre, SingHealth Duke-NUS Academic Medical Centre, Singapore, Singapore. 3. Vaccine Research & Development Center, School of Medicine, University of California Irvine, Irvine, CA, USA. Laboratory of Molecular Biology, Cambridge Biomedical Campus, Cambridge, UK. 4. The Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam. 5. Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Level 5, Jeffreys Cheah Biomedical Centre, Cambridge Biomedical Campus, University of Cambridge, Cambridge CB2 0AW, UK.

Received: 8 June 2020 Accepted: 18 December 2020

Published online: 15 January 2021

References

1. Kotloff KL, Riddle MS, Platts-Mills JA, Pavlinac P, Zaidi AKM. Shigellosis. Lancet. 2018;391(10122):801–12.
2. Khalil IA, Troeger C, Blacker BF, Rao PC, Brown A, Atherly DE, et al. Morbidity and mortality due to shigellosis and enterotoxigenic Escherichia coli diarrhoea: the global Burden of Disease Study 1990-2016. Lancet Infect Dis. 2018; 18(11):1229–40.
3. Olson S, Hall A, Riddle MS, Porter CK. Travelers’ diarrhea: an update on the incidence, etiology and risk in military and similar populations - 1990-2005 versus 2005-2015, does a decade make a difference? Tropical Diseases Travel Med Vaccines. 2019;5:1.
4. Anderson JD, Bagamian KH, Muhlb F, Amaya MP, Laytnner LA, Wierzb T, et al. Burden of enterotoxigenic Escherichia coli and shigella non-fatal diarrheal infections in 79 low-income and lower middle-income countries: a modelling analysis. Lancet Glob Health. 2019;7(3):e321–e30.
5. Baker S, Thomson N, Weill FX, Holt KE. Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens. Science. 2018;360(6390):728–8.
6. Mokomane M, Kasvosve I, de Melo E, Pernica JM, Goldfarb DM. The global problem of childhood diarrhoeal diseases: emerging strategies in prevention and management. Ther Adv Infect Dis. 2018;5(1):29–43.
7. KOTHARY MH, BABU US. Infective dose of foodborne pathogens in volunteers: a review. J Food Saf. 2001;21(1):49–68.
8. Lindsay B, Ochieng JB, Ikumapayi UN, Toure A, Ahmed D, Li S, et al. Quantitative PCR for detection of Shigella improves ascertainment of Shigella burden in children with moderate-to-severe diarrhea in low-income countries. J Clin Microbiol. 2013;51(6):1740–6.
9. Platts-Mills JA, Liu J, Rogaski ET, Kabir F, Lertsetthakam P, Sigaus M, et al. Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of shigellosis in children in low-resource settings: a reanalysis of the MAL-ED cohort study. Lancet Glob Health. 2018; 6(12):e1309–e138.
10. Mani S, Wierzb T, Walker RI. Status of vaccine research and development for Shigella. Vaccine. 2016;34(26):2887–94.
11. Cohen D, Green MS, Block C, Rouach T, Ofek I. Serum antibodies to lipopolysaccharide and natural immunity to shigellosis in an Israeli military population. J Infect Dis. 1998;177(5):1068–71.
12. Cohen D, Green MS, Block C, Stepon R, Ofek I. Prospective study of the association between serum antibodies to lipopolysaccharide O antigen and the attack rate of shigellosis. J Clin Microbiol. 1991;29(2):386–9.
13. Cohen D, Meron-Sudal S, Bialik A, Asato V, Goren S, Ariel-Cohen O, et al. Serum IgG antibodies to Shigella lipopolysaccharide antigens - a
correlation of protection against shigellosis. Hum Vaccin Immunother. 2019;15(6):1401–8.

14. Simon JK, Maciel M Jr, Weld ED, Wahid R, Pasetti MF, Picking W, et al. Antigen-specific IgA memory cell responses to Shigella antigens elicited in volunteers immunized with live attenuated Shigella flexneri 2a oral vaccine candidates. Clin Immunol. 2011;139(2):185–92.

15. Ndugo E, Randall A, Hazen TH, Kania DA, Trappil-Kimmons K, Liang X, et al. A novel Shigella proteome microarray discriminates targets of human anti-body reactivity following oral vaccination and experimental challenge. mSphere. 2018;3(4):e00260-18.

16. Levine MM, Kotloff KL, Barry EM, Pasetti MF, Simeen MB. Clinical trials of Shigella vaccines: two steps forward and one step back on a long, hard road. Nat Rev Microbiol. 2007;5(7):540–53.

17. Livio S, Strockbine NA, Panchalingam S, Tannent SM, Barry EM, Marohn ME, et al. Shigella isolates from the global enteric multicenter study inform vaccine development. Clin Infect Dis. 2014;59(7):933–41.

18. Venkatesan MM, Ranallo RT. Live-attenuated Shigella vaccines. Expert Rev Vaccines. 2005;4(5):699–68.

19. Orr N, Katz DE, Atsmon J, Rotu P, Yavzori M, Halperin T, et al. Protective immunity against invasive and noninvasive Shigella sonnei in volunteers vaccinated with a recombinant lipopolysaccharide. J Infect Dis. 2005;176(6):1621–28.

20. Kotloff KL, Liang X, Dumler JS, Barry EM, Pasetti MF, Sztein MB, Wooden SL, Livio S, et al. Safety and immunogenicity of a Shigella sonnei O-specific oligosaccharide-core-protein conjugate vaccine in adults in Israel. Infect Med. 2018;10(456):eaat4615.

21. Dharmasena MN, Hanisch BW, Wai TT, Kopecko DJ. Stable expression of Shigella sonnei form I O-polysaccharide genes recombineered into the chromosome of live Salmonella oral vaccine vector Ty21a. Int J Med Microbiol. 2013;303(4):105–13.

22. Kaminski RW, Wu M, Turbyfill KR, Clarkson K, Tali B, Bourgeois AL, et al. Development and preclinical evaluation of a trivalent, formalin-inactivated Shigella whole-cell vaccine. Clin Vaccine Immunol. 2014;21(3):365–82.

23. McKenzie R, Walker RI, Nabors GS, Van De Verg LL, Carpenter C, Gomes G, et al. Safety and immunogenicity of an oral, inactivated, whole-cell vaccine for Shigella sonnei: preclinical studies and a phase I trial. Vaccine. 2006; 24(18):3735–45.

24. Pozsgay V, Chu C, Pannell L, Wolfe J, Robbins JB, Schneeom R. Protein conjugates of synthetic saccharides elicit higher levels of serum IgG lipopolysaccharide antibodies in mice than do those of the O-specific polysaccharide from Shigella dysenteriae type 1. Proc Natl Acad Sci U S A. 19999;96(9):5194–7.

25. Passwell JH, Ashkenazi S, Banet-Levy Y, Ramon-Saraf R, Farzam N, Lerner-Geva L, et al. Age-related efficacy of Shigella O-specific polysaccharide conjugates in 1-4-year-old Israeli children. Vaccine. 2010;28(10):2231–5.

26. Farzam N, Ramon-Saraf R, Banet-Levy Y, Lerner-Geva L, Ashkenazii K, Kubler-Sonnet et al. Vaccination with Shigella flexneri 2a conjugate induces type 2a polysaccharide from Shigella dysenteriae type 1. Proc Natl Acad Sci U S A. 2005;102(3):747–52.

27. Liveo S, Strockbine NA, Simeen MB. Clinical trials of Shigella sonnei O-polysaccharide genes recombineered into the complete genomes. J Mol Biol. 2001;305(6):857–75.

28. Nielsen H. Predicting secretory proteins with SignalP. Methods Mol Biol. 2017;1611:59–73.

29. Davies DH, Liang X, Hernandez JE, Randall A, Hirt S, Yu M, et al. Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery. Proc Natl Acad Sci U S A. 2005;102(3):547–52.

30. Caboni M, Pedron T, Rossi O, Goulding D, Pickard D, Citiulo F, et al. An O antigen capsule modulates bacterial pathogenesis in Shigella sonnei. PLoS Pathog. 2015;11(3):el004749.

31. de Alwis R, Tu LTP, Quynh NLT, Thompson CN, Anders KL, Van Thuy NT, et al. Immune profiling with a Salmonella Typhi antigen microarray identifies new diagnostic biomarkers of human typhoid. Sci Rep. 2013;3:1045.

32. Fouda GG, Martinez DR, Swamy GK, Permar SR. The impact of IgG transplacental transfer on early life immunity. Immunohorizons. 2018;2(1):14–25.

33. Caboni M, Pedron T, Rossi O, Goulding D, Pickard D, Citiulo F, et al. Antigen capsule modulates bacterial pathogenesis in Shigella sonnei. PLoS Pathog. 2015;11(3):el004749.

34. Thorpe TF, McNeilly DA, Sawyer S, Nayyim NM, Tu le TP, Dung TT, et al. The epidemiology and aetiology of diarrhoeal disease in infancy in southern Vietnam: a birth cohort study. Int J Infect Dis. 2015;35:3–10.

35. Finnt RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, et al. The Pfam protein families database. Nucleic Acids Res. 2008;36(Database issue):D281–8.

36. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305(5):567–80.
59. Pasqua M, Grossi M, Scinicariello S, Ausel L, Barras F, Colonna B, et al. The MFS efflux pump EmrKY contributes to the survival of Shigella within macrophages. Sci Rep. 2019;9(1):2906.

60. Nagakubo S, Nishino K, Hirata T, Yamaguchi A. The putative response regulator BaeR stimulates multidrug resistance of Escherichia coli via a novel multidrug exporter system. MdtABC. J Bacteriol. 2002;184(15):4161–7.

61. Kuntumalla S, Zhang Q, Braisted JC, Fleischmann RD, Peterson SN, Donohue-Rolfe A, et al. In vivo versus in vitro protein abundance analysis of Shigella dysenteriae type 1 reveals changes in the expression of proteins involved in virulence, stress and energy metabolism. BMC Microbiol. 2011;11:147.

62. Wu T, Malinverni J, Ruiz N, Kim S, Silhavy TJ, Kahne D. Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. Cell. 2005;121(2):235–45.

63. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae, an urgent threat to public health. Proc Natl Acad Sci U S A. 2015;112(27):E3574–81.

64. Skaar EP, Humayun M, Bae T, DeBord KI, Schneewind O. Iron-source preference of Staphylococcus aureus infections. Science. 2004;305(5690):1626–8.

65. Mao BH, Chang YF, Scaria J, Chang CC, Chou LW, Tien N, et al. Identification of Escherichia coli genes associated with urinary tract infections. J Clin Microbiol. 2012;50(2):449–56.

66. Ellis TN, Kuehn MJ. Virulence and immunomodulatory roles of bacterial outer membrane vesicles. Microbiol Mol Biol Rev. 2010;74(1):81–94.

67. Launay O, Ndiaye AGW, Conti V, Loulegue P, Scire AS, Landre AM, et al. Booster vaccination with GVG Shigella sonnei 1790GAHB GMMA vaccine compared to single vaccination in unvaccinated healthy European adults: results from a phase 1 clinical trial. Front Immunol. 2019;10:335.

68. Maggiore L, Yu L, Omaits U, Rossi O, Dougan G, Thomson NR, et al. Quantitative proteomic analysis of Shigella flexneri and Shigella sonnei Generalized Modules for Membrane Antigens (GMMA) reveals highly pure preparations. Int J Med Microbiol. 2016;306(2):99–108.

69. Boullier S, Tanguy M, Kadaoui KA, Caubet C, Sansonetti P, Corthesy B, et al. Secretory IgA-mediated neutralization of Shigella flexneri prevents intestinal tissue destruction by down-regulating inflammatory circuits. J Immunol. 2009;183(9):5879–85.

70. Mathias A, Longet S, Corthesy B. Agglutinating secretory IgA preserves intestinal epithelial cell integrity during apical infection by Shigella flexneri. Infect Immun. 2013;81(8):3027–34.

71. Barel LA, Mulard LA. Classical and novel strategies to develop a Shigella glycoconjugate vaccine from concept to efficacy in human. Hum Vaccin Immunother. 2019;15(6):1338–56.

72. de Alwis R, and Baker S. The identification of novel immunogenic antigens as potential Shigella vaccine components (version 1). Zenodo. https://doi.org/10.5281/zenodo.4320935 (2020).

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