



# Genetic and Structural Variation in the O-Antigen of *Salmonella enterica* Serovar Typhimurium Isolates Causing Bloodstream Infections in the Democratic Republic of the Congo

 Sandra Van Puyvelde,<sup>a,b,c,d</sup> Gianmarco Gasperini,<sup>e</sup>  Michael Biggel,<sup>b</sup> Marie-France Phoba,<sup>f,g</sup>  Maria Michelina Raso,<sup>e</sup> Tessa de Block,<sup>d</sup> Leen N. Vanheer,<sup>d,\*</sup> Stijn Deborggraeve,<sup>d,§</sup> Olivier Vandenberg,<sup>h,i,j</sup> Nicholas Thomson,<sup>c,k</sup> Neil Ravenscroft,<sup>l</sup> Calman A. MacLennan,<sup>m</sup> Barbara Bellich,<sup>n</sup> Paola Cescutti,<sup>n</sup> Gordon Dougan,<sup>a</sup> Jan Jacobs,<sup>d,o</sup> Octavie Lunguya,<sup>f,g</sup> Francesca Micoli<sup>e</sup>

<sup>a</sup>Cambridge Institute of Therapeutic Immunology and Infectious Disease, Department of Medicine, University of Cambridge, Cambridge, United Kingdom

<sup>b</sup>Laboratory of Medical Microbiology, Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, Belgium

<sup>c</sup>Wellcome Trust Sanger Institute, Hinxton, United Kingdom

<sup>d</sup>Institute of Tropical Medicine, Antwerp, Belgium

<sup>e</sup>GSK Vaccines Institute for Global Health, Siena, Italy

<sup>f</sup>Department of Microbiology, Institut National de Recherche Biomédicale, Kinshasa, Democratic Republic of the Congo

<sup>g</sup>Department of Medical Biology, University Teaching Hospital of Kinshasa, Kinshasa, Democratic Republic of the Congo

<sup>h</sup>Clinical Research and Innovation Unit, Laboratoire Hospitalier Universitaire de Bruxelles-Universitaire Laboratorium Brussel, Université Libre de Bruxelles, Brussels, Belgium

<sup>i</sup>Centre for Environmental Health and Occupational Health, School of Public Health, Université Libre de Bruxelles, Brussels, Belgium

<sup>j</sup>Division of Infection and Immunity, Faculty of Medical Sciences, University College London, London, United Kingdom

<sup>k</sup>London School of Hygiene and Tropical Medicine, London, United Kingdom

<sup>l</sup>Department of Chemistry, University of Cape Town, Rondebosch, South Africa

<sup>m</sup>Jenner Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

<sup>n</sup>Department of Life Sciences, University of Trieste, Trieste, Italy

<sup>o</sup>Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

**ABSTRACT** *Salmonella enterica* serovar Typhimurium causes a devastating burden of invasive disease in sub-Saharan Africa with high levels of antimicrobial resistance. No licensed vaccine is available, but O-antigen-based candidates are in development, as the O-antigen moiety of lipopolysaccharides is the principal target of protective immunity. The vaccines under development are designed based on isolates with O-antigen O-acetylated at position C-2 of abequose, giving the O:5 antigen. Serotyping data on recent *Salmonella* Typhimurium clinical isolates from the Democratic Republic of the Congo (DRC), however, indicate increasing levels of isolates without O:5. The importance and distribution of this loss of O:5 antigen in the population as well as the genetic mechanism responsible for the loss and chemical characteristics of the O-antigen are poorly understood. In this study, we Illumina whole-genome sequenced 354 *Salmonella* Typhimurium isolates from the DRC, which were isolated between 2002 and 2017. We used genomics and phylogenetics combined with chemical approaches (<sup>1</sup>H nuclear magnetic resonance [NMR], high-performance anion-exchange chromatography with pulsed amperometric detection [HPAEC-PAD], high-performance liquid chromatography–PAD [HPLC-PAD], and HPLC-size exclusion chromatography [HPLC-SEC]) to characterize the O-antigen features within the bacterial population. We observed convergent evolution toward the loss of the O:5 epitope predominantly caused by recombination events in a single gene, the O-acetyltransferase gene *oafA*. In addition, we observe further O-antigen variations, including O-acetylation of the rhamnose residue, different levels of glucosylation, and the absence of O-antigen repeating units. Large recombination events underlying O-antigen variation were resolved using long-read MinION sequencing. Our study suggests evolutionary pressure toward O-antigen variants in a region where invasive

**Editor** Philippe J. Sansonetti, Pasteur Institute

**Copyright** © 2022 Van Puyvelde et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Sandra Van Puyvelde, Sandra.VanPuyvelde@uantwerpen.be.

\*Present address: Leen N. Vanheer, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom.

§Present address: Stijn Deborggraeve, Médecins Sans Frontières Access Campaign, Geneva, Switzerland.

The authors declare a conflict of interest. GG, MMR and FM are employed by the GSK group of companies.

**Received** 22 February 2022

**Accepted** 29 June 2022

disease by *Salmonella* Typhimurium is highly endemic. This needs to be taken into account when developing O-antigen-based vaccines, as it might impact the breadth of coverage in such regions.

**IMPORTANCE** The bacterium *Salmonella* Typhimurium forms a devastating burden in sub-Saharan Africa by causing invasive bloodstream infections. Additionally, *Salmonella* Typhimurium presents high levels of antimicrobial resistance, jeopardizing treatment. No licensed vaccine is available, but candidates are in development, with lipopolysaccharides being the principal target of protective immunity. The vaccines under development are designed based on the O:5 antigen variant of bacterial lipopolysaccharides. Data on recent *Salmonella* Typhimurium clinical isolates from the Democratic Republic of the Congo (DRC), however, indicate increasing levels of isolates without this O:5 antigen. We studied this loss of O:5 antigen in the population at the genetic and chemical levels. We genome sequenced 354 isolates from the DRC and used advanced bioinformatics and chemical methods to characterize the lipopolysaccharide features within the bacterial population. Our results suggest evolutionary pressure toward O-antigen variants. This needs to be taken into account when developing vaccines, as it might impact vaccine coverage.

**KEYWORDS** Democratic Republic of the Congo, O-antigen, *Salmonella* Typhimurium, genomics, iNTS, surface structures, vaccines, whole-genome sequencing

The bacterium *Salmonella enterica* serovar Typhimurium is globally known as a common cause of gastrointestinal infections. However, *S. Typhimurium* is also responsible for an urgent health burden in sub-Saharan Africa by causing severe bloodstream infections. Two-thirds of invasive nontyphoidal *Salmonella* (iNTS) infections are caused by *S. enterica* serovar Typhimurium, and one-third are caused by *S. enterica* serovar Enteritidis, jointly responsible for over 500,000 bloodstream infections per year in sub-Saharan Africa (1). The fatality rate of iNTS infections can be extremely high and has been estimated to be up to 20% (2). iNTS infections disproportionately affect children under 5 years of age, with malnutrition and coinfection with malaria and HIV as major risk factors (2).

*S. Typhimurium* is identified upon serotyping, presenting a specific combination of O-antigen (OAg) and H-antigen compounds. *S. Typhimurium* further has multiple sequence types (STs), with ST19, -34, and -313 being the most reported. Bloodstream infections observed in sub-Saharan Africa are specifically associated with ST313, which is comprised of lineage I and lineage II that sequentially spread in this region (3). ST313 lineage I was found mostly in East African countries, whereas ST313 lineage II was identified across sub-Saharan Africa and was found to be dominant in recent studies (3–6). A higher diversity of lineages is seen in Kenya, where ST19 isolates were found to cause a substantial proportion of bloodstream infections, similar to ST313 isolates from lineages I and II (7).

Antimicrobial resistance (AMR) is high and further increasing, which threatens effective treatment (8). In the Democratic Republic of the Congo (DRC), an extensively drug-resistant (XDR) sublineage outbreak has recently been identified, which remained susceptible to only one available antibiotic, ciprofloxacin, and two isolates were found to be pan-drug resistant (PDR) (4), i.e., resistant to all available antibiotics.

Alternative intervention strategies for iNTS are urgently needed, especially in regions threatened by a high prevalence and/or high AMR levels. No licensed vaccine is currently available, but different candidates are in development (9). The OAg moiety of bacterial cell surface lipopolysaccharides (LPSs) is the principal target of protective immunity (10–14). OAg-based vaccines targeting iNTS, including glycoconjugate vaccines (LPS-derived OAg linked to carrier proteins) and generalized modules for membrane antigens (GMMA) (outer membrane vesicles consisting of surface polysaccharides and outer membrane proteins), are in development and have been shown to provide protection in *in vitro* and *in vivo* models (11, 15–19).

The *S. Typhimurium* OAg is composed of repeating units containing the monosaccharides mannose (Man), rhamnose (Rha), and galactose (Gal) (O:12 specificity), which can be variably glucosylated at galactose, through (1→6) or (1→4) linkages conferring O:1 and O:12<sub>2</sub> specificities, respectively (20). The immunodominant dideoxyhexose abequose (Abe) linked  $\alpha$ -(1→3) to mannose confers O:4 specificity. The Abe and Rha sugars can be further O-acetylated (21). O:5 specificity is conferred by O-acetylation at position C-2 of Abe (21).

*S. Typhimurium* O:5-negative (O:5<sup>−</sup>) isolates were first discovered by F. Kauffman in 1934 and classified as the “Copenhagen *varietas*.” These O:5<sup>−</sup> variants were reported in pigeons, heifers, cattle, and swine and occasionally in dogs and cats, while cases in humans have been extremely rare (22, 23). However, O:5 specificity is not part of routine *S. Typhimurium* serological identification. Data collected during 10 years of surveillance in the DRC for which the O:5 specificity was studied indicate that there is a large proportion of *S. Typhimurium* clinical isolates lacking O:5, estimated to be 45% of all clinical isolates in 2017 (24). However, the vaccines under development are designed based on O:5-containing (O:5<sup>+</sup>) isolates.

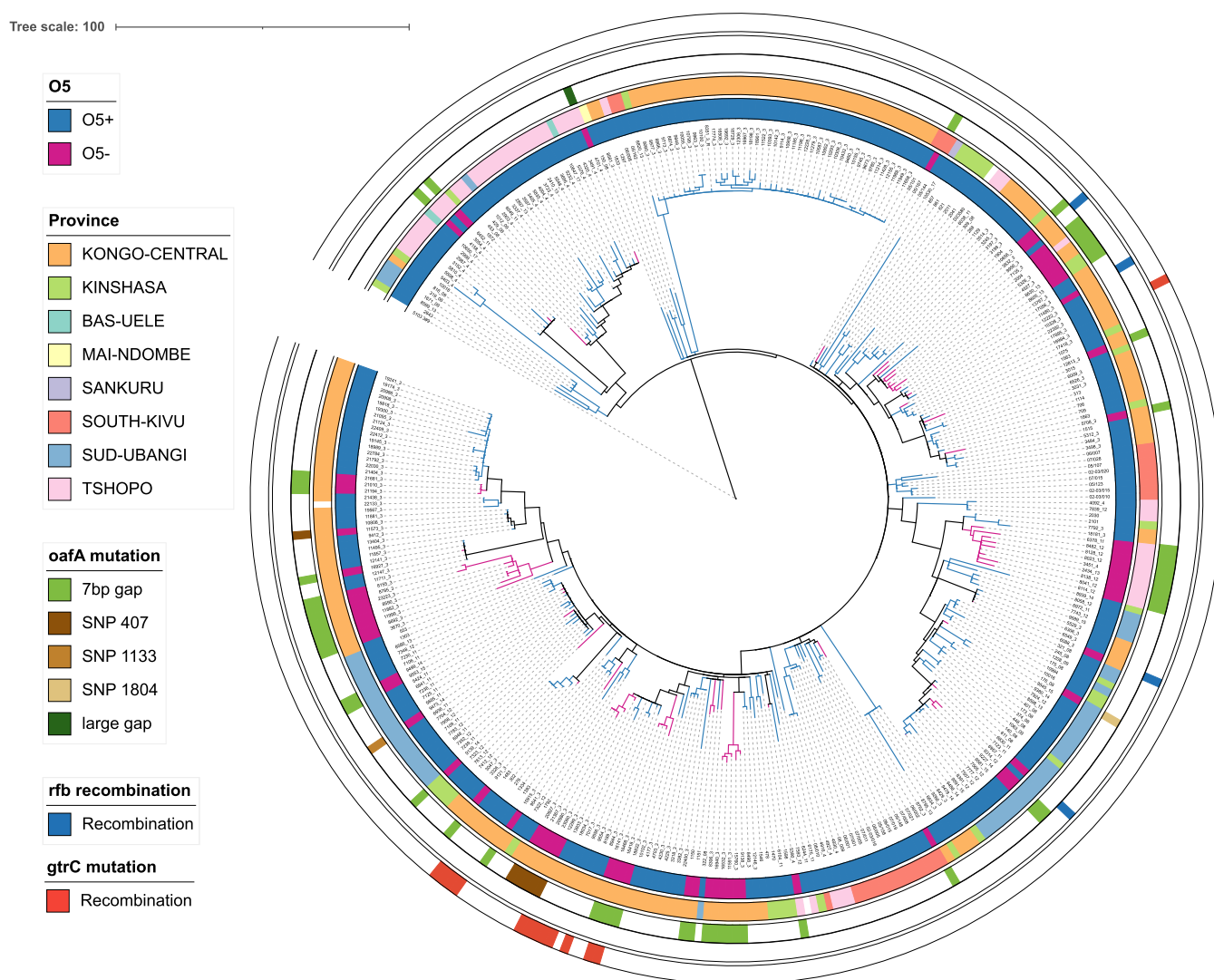
In this work, we used a combined genomics and chemical approach to study this OAg variation in the bacterial population of *S. Typhimurium* causing bloodstream infections in the DRC.

## RESULTS

**Convergent loss of O:5 specificity in the *S. Typhimurium* population is caused by recombination in *oafA*.** O:5<sup>−</sup> invasive *S. Typhimurium* isolates were increasingly observed in the DRC (8). We observed that within a selection of these isolates, the O:5 phenotype was stable after subculturing and over time, which would not be expected if this were a bistable phenotype. This suggests that the loss of the O:5 specificity in the DRC isolates has a genetic cause in the O:5<sup>−</sup> isolates and is not a result of variable expression in the *S. Typhimurium* population.

To identify the genetic cause, a total of 354 *S. Typhimurium* isolates with accompanying metadata and serotype data were whole-genome sequenced. All isolates were identified as closely related ST313 lineage II isolates by building a core-genome phylogenetic tree (Fig. 1). The 71 O:5<sup>−</sup> *S. Typhimurium* isolates were found across ST313 lineage II. When the O:5 phenotype was marked across the phylogeny, it showed at least 30 independent changes from O:5<sup>+</sup> to O:5<sup>−</sup>. Multiple monophyletic branches with many independently arisen O:5<sup>−</sup> isolates were evident. These isolates appear sporadically throughout the tree, implying the frequent local loss and circulation of O:5<sup>−</sup> isolates in the DRC. While the different clusters of O:5<sup>−</sup> isolates predominantly remained within a province, some showed spread across the DRC. Six of these clusters were confined to the Kongo-Central province, one was confined to Sud-Ubangi, another cluster was found in both Tshopo and Sud-Ubangi, and one cluster was found in Kongo-Central and Kinshasa.

A pangenomic analysis combined with a genome-wide association study (GWAS) identified the presence of the O-acetyltransferase gene *oafA* (also named *oatA*, *yrhL*, or STM2232) as being most significantly associated with the loss of the O:5 phenotype (Bonferroni-corrected *P* value of 1.65e−41). The genes neighboring *oafA* are the *umuD* gene and a predicted phage tail fiber assembly gene, which was previously shown to mediate OAg abequose O-acetylation and, hence, O:5 specificity in *S. Typhimurium* (25, 26). By further investigation of single nucleotide polymorphisms (SNPs) and variations in the genome assemblies, the *oafA* gene was found to be intact in all O:5<sup>+</sup> isolates (*n* = 283), while truncations in the *oafA* gene were observed in the O:5<sup>−</sup> isolates. The majority of the O:5<sup>−</sup> isolates showed the same 7-nucleotide (nt) deletion at position 429, resulting in a frameshift mutation that breaks the *oafA* gene into two smaller fragments (*n* = 53). A 7-nt sequence (ATTTTAT) is present as repeating units in the O:5<sup>+</sup> isolate genomes, but one of the repeating units has been lost in the O:5<sup>−</sup> isolates (Fig. 2A). Six isolates forming one monophyletic branch showed a 407G-A SNP in *oafA* resulting in a premature stop codon (TGG [W] to TAG *amber* [stop] codon), one isolate showed an 1133G-A SNP in *oafA* resulting in a premature stop codon (TGG [W] to TAG

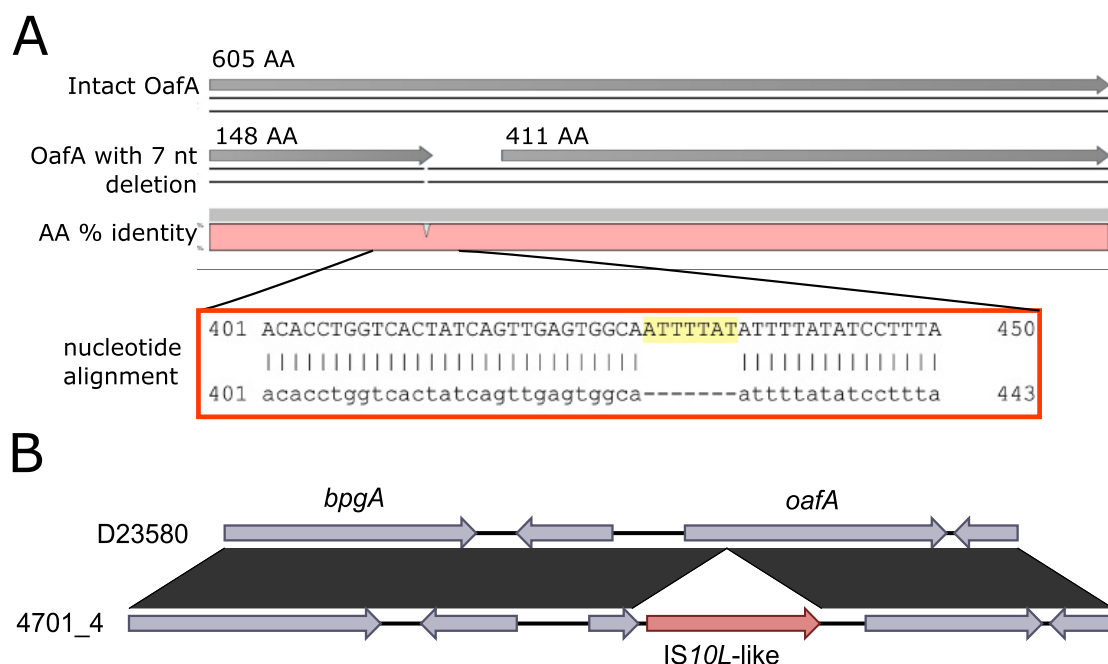


**FIG 1** Phylogenetic tree of invasive *Salmonella* Typhimurium isolates from the Democratic Republic of the Congo (DRC) indicating the presence and absence of the O:5 antigen phenotype. The maximum likelihood phylogenetic tree is based on the 354 genome sequences from this study (summarized in Table S1 in the supplemental material). Sequencing reads were mapped to *S. Typhimurium* ST313 lineage II reference strain D23580 (50). The tree is based on 2,131 chromosomal SNPs. Metadata are visualized on the concentric rings according to the key as follows, from the inside to the outside; O:5 phenotype (ring 1); province of isolation (ring 2); and mutation in the *oafA* gene, *rfb* operon, and *gtrC* gene (rings 3 to 5). Branches are colored by the O:5 phenotype. Isolate DT2 is included as the outgroup to root the tree. Branch lengths represent the number of SNPs, as indicated by the scale bar.

*amber* [stop] codon), and one isolate showed an 1804G-T SNP in *oafA*, again resulting in a premature stop codon (GAA [E] to TAA *ochre* [stop] codon). Finally, the *oafA* gene in isolate 4701/4 was truncated by the insertion of an IS4 transposase (Fig. 2B).

Our data thus show the repeated evolutionary loss of O:5<sup>−</sup> specificity in the *S. Typhimurium* population predominantly driven by an independent deletion of a 7-nt sequence within *oafA* but also with other isolates showing alternative mutational events acting at the level of transcription or that truncate or disrupt the function of the translated OafA protein product (Fig. 1).

**O:5<sup>−</sup> isolates lack O-acetylation of Abe or the full OAg.** Twenty-five invasive *S. Typhimurium* isolates were chosen for full chemical characterization, including 10 O:5<sup>−</sup> isolates presenting *oafA* mutations, 4 O:5<sup>−</sup> isolates with intact *oafA* genes, and 11 O:5<sup>+</sup> isolates with intact *oafA* genes (Table 1; see also Fig. S1 in the supplemental material). These isolates were chosen to reflect (i) the different types of *oafA* mutations observed, (ii) the genetic diversity in the *S. Typhimurium* phylogeny, and (iii) the spatiotemporal diversity in the overall collection. These isolates were analyzed for OAg production



**FIG 2** Mutational events in the *oafA* gene associated with the loss of the O:5 phenotype. (A) Pairwise amino acid (AA) assembly of an intact OafA protein and the OafA protein sequence of an isolate showing a frameshift mutation linked to a 7-nucleotide ATTTAT deletion in the *oafA* gene. (B) Pairwise comparison of the *oafA* gene sequences of D23580 (O:5<sup>+</sup>) and 4701/4 (O:5<sup>-</sup>) presenting an insertion of an IS4 family transposase into *oafA*, as resolved through MinION long-read sequencing.

levels, structural modifications (O-acetylation position and glucosylation level), and the OAg molecular size distribution. The results obtained are summarized in Table 1.

The O:5<sup>-</sup> isolates with *oafA* mutations all showed an absence of 2-O-acetylation at the Abe residue, irrespective of the disruption to *oafA*.

Characterization of the four O:5<sup>-</sup> isolates with intact *oafA* genes by flow cytometry and LPS silver staining suggested that these isolates were OAg negative (Fig. S2). The absence of OAg was confirmed by <sup>1</sup>H nuclear magnetic resonance (NMR) and high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD). These four OAg-negative isolates possessed transposase insertions within the *rfb* locus, which could be fully resolved by MinION long-read sequencing (Fig. 3). *rfb* (or *wb*) genes encode the OAg biosynthesis of Gram-negative bacteria and include genes involved in nucleotide sugar synthesis and O-unit assembly (27).

The OAg-negative isolates containing transposase insertions in the *rfb* locus were found as singletons distributed across the phylogeny, with mutational events that suggest that they have arisen multiple times and independently (Fig. 1).

**Further diversity in the *S. Typhimurium* OAg structure due to O-acetylation and glucosylation.** Our <sup>1</sup>H NMR spectra showed that in addition to the O-acetylation on C-2 of Abe, all isolates, with the exception of 18034/3, displayed OAg with O-acetylation on C-2 and C-3 of Rha, independently from the O:5 specificity (Fig. 4A and Table 1). The presence of 2/3-OAc (indicating that OAc can be in position 2 or 3 of the rhamnose sugar ring) on Rha has previously been observed for the *S. Typhimurium* ST313 reference isolate D23580 and was acquired through the uptake of the BTP1 prophage (21, 28, 29). The genetic determinant of this OAg modification was formerly named *gtrC2* (or *gtrCc* or *gtrC\_BTP1*) by association with GtrABC glycosyltransferase family 2. However, as no functional relationship is present, the gene was recently renamed *oafB*, similar to the name already in use for *oafA* (30). MinION sequencing resolved the full genome of isolate 18034/3 and revealed recombination in which the full region was lost, including *oafB* (Fig. 4B). In the full isolate collection, 13 isolates lacked the *oafB* gene, of which 3 were O:5<sup>-</sup> and 10 were O:5<sup>+</sup>. These 13 isolates are all related strains originating from Kongo-Central (Fig. S1).

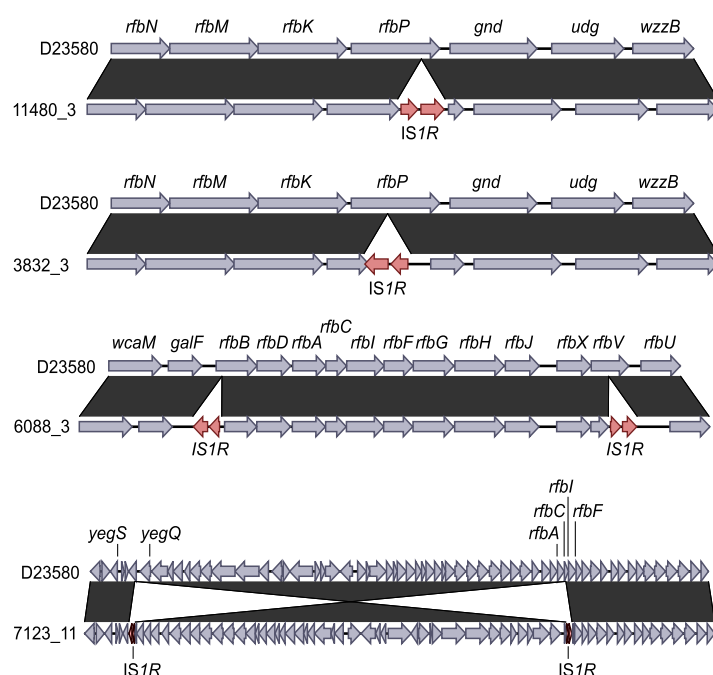
Glucosylation levels varied between 9 and 52% among the different isolates that were chemically characterized (Table 1). All these 25 isolates showed the presence of



**TABLE 1** Overview of genomic and physicochemical characterization of isolates studied<sup>a</sup>

Isolate ID	O:5 Specificity	oafA	μg OAg/OD	OAc Abe	oafB	OAc Rha	gtrC3	% Glc	rfb	OAg length		OAg length	
										% HMM	% MMM	HMM kDa	MMM kDa
O:5 <sup>-</sup> serotype and recombination event in oafA													
8128/12	O:5 <sup>-</sup>	7-nt gap	10.0	-	I	+	I	52	I	13	87	95	28
8692/3	O:5 <sup>-</sup>	7-nt gap	10.6	-	I	+	I	9	I	14	86	85	26
9499/14	O:5 <sup>-</sup>	7-nt gap	11.4	-	I	+	I	28	I	27	73	95	27
11591/3	O:5 <sup>-</sup>	7-nt gap	12.2	-	I	+	I	19	I	13	87	90	27
18832/3	O:5 <sup>-</sup>	7-nt gap	12.4	-	I	+	I	17	I	11	89	91	27
9227/14	O:5 <sup>-</sup>	7-nt gap	10.9	-	I	+	I	29	I	11	89	90	26
20980/3	O:5 <sup>-</sup>	407G-A	13.6	-	I	+	I	9	I	11	89	88	26
9473/14	O:5 <sup>-</sup>	1133G-A	13.1	-	I	+	I	41	I	10	90	87	27
176/08	O:5 <sup>-</sup>	1804G-T	12.2	-	I	+	I	21	I	16	84	93	27
4701/4	O:5 <sup>-</sup>	IS4 insertion	7.4	-	I	+	I	12	I	82	18	85	41
O:5 <sup>-</sup> serotype and no recombination event in oafA													
6088/3	O:5 <sup>-</sup>	I	OAg-neg	n.a.	I	n.a.	I	n.a.	M	n.a.	n.a.	n.a.	n.a.
7123/11	O:5 <sup>-</sup>	I	OAg-neg	n.a.	I	n.a.	I	n.a.	M	n.a.	n.a.	n.a.	n.a.
11480/3	O:5 <sup>-</sup>	I	OAg-neg	n.a.	I	n.a.	I	n.a.	M	n.a.	n.a.	n.a.	n.a.
3832/3	O:5 <sup>-</sup>	I	OAg-neg	n.a.	I	n.a.	I	n.a.	M	n.a.	n.a.	n.a.	n.a.
O:5 <sup>+</sup> serotype and no recombination events in oafA													
18034/3	O:5 <sup>+</sup>	I	14.5	+	M	-	I	38	I	5	95	62	28
4225/3	O:5 <sup>+</sup>	I	6.2	+	I	+	I	43	I	21	79	98	28
4320/4	O:5 <sup>+</sup>	I	7.4	+	I	+	I	25	I	80	20	88	45
7704/12	O:5 <sup>+</sup>	I	4.6	+	I	+	I	16	I	15	85	87	26
8058/12	O:5 <sup>+</sup>	I	6.0	+	I	+	I	18	I	30	70	93	28
9478/14	O:5 <sup>+</sup>	I	7.2	+	I	+	I	51	I	13	87	93	28
13404/3	O:5 <sup>+</sup>	I	7.5	+	I	+	I	9	I	17	83	93	26
10328/3	O:5 <sup>+</sup>	I	7.8	+	I	+	I	47	I	11	89	92	28
173/08	O:5 <sup>+</sup>	I	9.6	+	I	+	I	23	I	47	53	85	26
7565/12	O:5 <sup>+</sup>	I	8.0	+	I	+	I	36	I	12	88	98	29
7323/12	O:5 <sup>+</sup>	I	7.4	+	I	+	I	37	I	8	92	85	26

<sup>a</sup>For each isolate, the isolate identifier is given, along with the O:5 phenotype as determined upon isolation and collection; the conservation of the *oafA*, *gtrC2*, *gtrC3*, and *rfb* genes as determined by Illumina and MinION sequencing analyses; the OAg production levels (micrograms of OAg/OD unit by HPAEC-PAD after OAg isolation); O-acetylation (OAc) of the abequose (Abe) and rhamnose (Rha) residues as determined by <sup>1</sup>H NMR; the percentage of glucosylation (Glc) as determined by HPAEC-PAD; and the OAg length by HPLC-SEC (average size determined by using dextrans as standards and percentage of populations at different relative sizes as determined by dRI areas of the peaks). HMM, high-molecular-mass OAg; MMM, medium-molecular-mass OAg; I, intact; M, mutated; n.a., not applicable.



**FIG 3** Resolved *rfb* recombinations underlying OAg loss using MinION sequencing. The genome sequences of isolates 11480/3, 3832/3, 6088/3, and 7123/11 are pairwise compared to that of isolate D23580 (GenBank accession number [FN424405](#)). Homologous regions between both sequences are colored according to the scale bar. Transposase genes are shown in red. Genes of *S. Typhimurium* D23580 are annotated with the *S. Typhimurium* LT homologs (GenBank accession number [AE006468.2](#)).

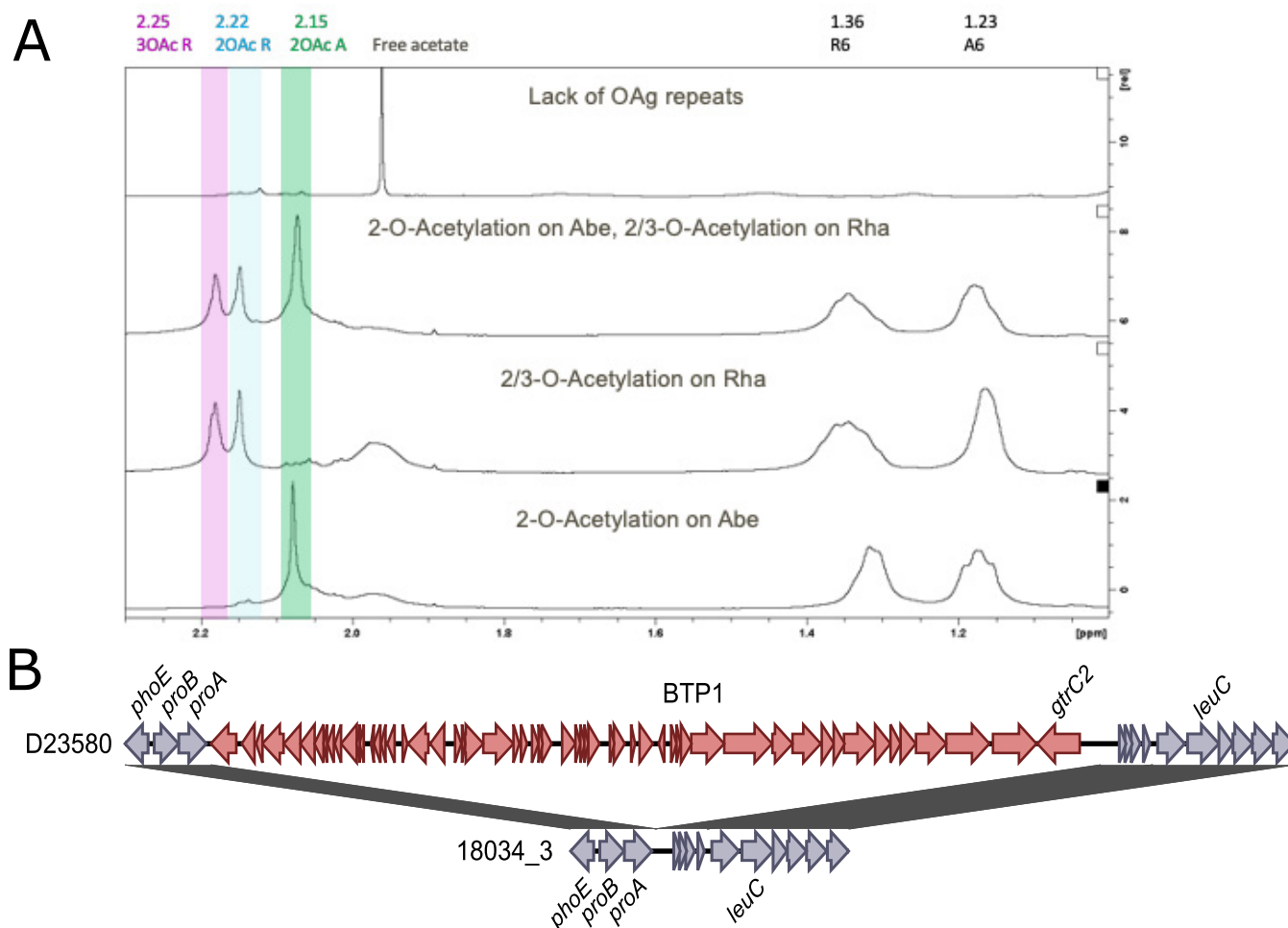
an intact *gtrC3* gene, responsible for the  $\alpha$ 1-4 linkage of glucose on galactose in the OAg repeating unit (O:1<sub>2</sub> specificity). None of these isolates showed the presence of a *gtrC1* gene, responsible for the  $\alpha$ 1-6 linkage of glucose on galactose in the OAg repeating unit (O:1 specificity). A genetic cause underlying the variability in OAg O-acetylation and glucosylation levels in the characterized isolates could not be identified.

All of the OAg-positive isolates showed a bimodal distribution of the OAg chain length, which is associated with the presence of intact *wzzB* and *fepE* genes (31). The relative amounts of high-molecular-mass (HMM) OAg and medium-molecular-mass (MMM) OAg varied among the different isolates (Table 1). HMM OAg varied between 85 and 98 kDa, with the exception of 1 isolate with no O-acetylation on Rha, having an HMM OAg of 62 kDa. MMM OAg ranged between 26 and 29 kDa, but two isolates had MMM OAg of higher molecular masses of around 40 kDa.

Interestingly, it was observed that the OAg expression levels were lower in the strains expressing OAg O-acetylated at both Abe and Rha positions (Table 1).

Five OAg (produced by isolates 18034/3, 8128/12, 8692/3, 13404/3, and 10328/3) differing in the patterns of O-acetylation and levels of glucosylation were selected to be further purified and subjected to more detailed structural characterization by <sup>1</sup>H NMR and gas-liquid chromatography coupled with mass spectrometry (GLC-MS). <sup>1</sup>H NMR confirmed the expected pattern and degree of O-acetylation (Table S2). Composition and linkage analyses by GLC-MS confirmed the identity of the expected monosaccharides together with their glycosidic linkages. In particular, it was confirmed that the Gal residue was glucosylated on C-4, and the measured percentages of glucosylation were in agreement with those found by HPAEC-PAD (Table S3).

**Impact of O:5 specificity on the binding capacity and bactericidal activity of different anti-OAg antibodies.** To investigate whether the observed OAg structural diversity could impact the ability of specific antibodies to bind to the bacterial surface, the panel of 25 selected isolates was analyzed by fluorescence-activated cell sorting (FACS) using commercial anti-OAg antibodies. As a control, *S. Typhimurium* strain D23580 (O:5<sup>+</sup>) was included in the analysis. Interestingly, while a rabbit polyclonal antibody (catalog number 294401; Denka Seiken) showed similar abilities to bind to all tested isolates irrespective of



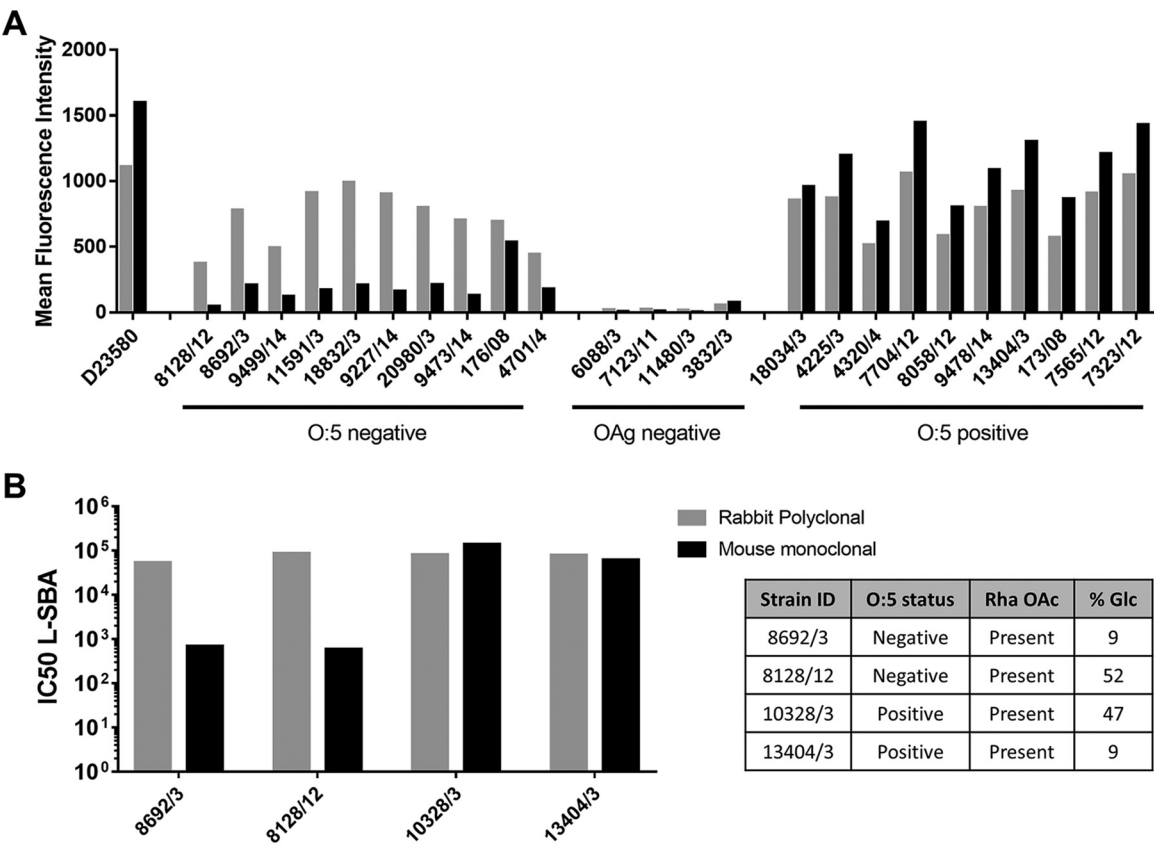
**FIG 4** Diverse O-acetylation patterns. (A) <sup>1</sup>H NMR spectra of OAGs extracted from 4 isolates representative of the variability observed. From top to bottom, the spectra are shown from an isolate that (i) lacks OAg repeats, (ii) is O-acetylated (OAc) on both rhamnose (Rha) and abeque (Abe), (iii) is O-acetylated on Rha only, or (iv) is O-acetylated on Abe only. (B) Resolved *gtrC2* recombination underlying the loss of O-acetylation of rhamnose in isolate 18034/3 compared to isolate D23580. Homologous regions between both sequences are colored according to the scale bar.

the O:5 status, a mouse monoclonal antibody (catalog number ab8274; Abcam) showed higher specificity for O:5<sup>+</sup> isolates (Fig. 5A). The same antibodies were tested in a serum bactericidal assay (SBA) to confirm their ability to trigger complement-mediated killing of isolates 8128/12 (O:5<sup>-</sup>), 8692/3 (O:5<sup>-</sup>), 13404/3 (O:5<sup>+</sup>), and 10328/3 (O:5<sup>+</sup>), also selected based on the different levels of OAg glucosylation (Table 1). Strikingly, while similar killing was observed against all isolates using the rabbit polyclonal serum, significantly less killing was observed against O:5<sup>-</sup> isolates using the mouse monoclonal antibody (Fig. 5B), independently from OAg glucosylation.

## DISCUSSION

We have observed an unexpectedly high level of OAg variation in the *S. Typhimurium* population in the DRC, as summarized in Fig. 6. This variation includes the loss of OAg, Abe or Rha O-acetylation, and differences in glucosylation levels. The absence of O-acetylation of the Abe residue resulted from mutations in *oafA* that explain the O:5<sup>-</sup> phenotype. Disruption of the production of O:5 LPS was highly prevalent in our *S. Typhimurium* data set, with multiple O:5<sup>-</sup> isolates being present across the phylogeny in monophyletic lineages. The distribution of the O:5<sup>-</sup> isolates suggests that these mutations occurred in the DRC itself and that the O:5<sup>-</sup> isolates possessing these genetic lesions have spread further locally. This suggests that there might be a selective advantage to the loss of O:5<sup>+</sup> by mutation of *oafA* among invasive *S. Typhimurium* isolates in the DRC.

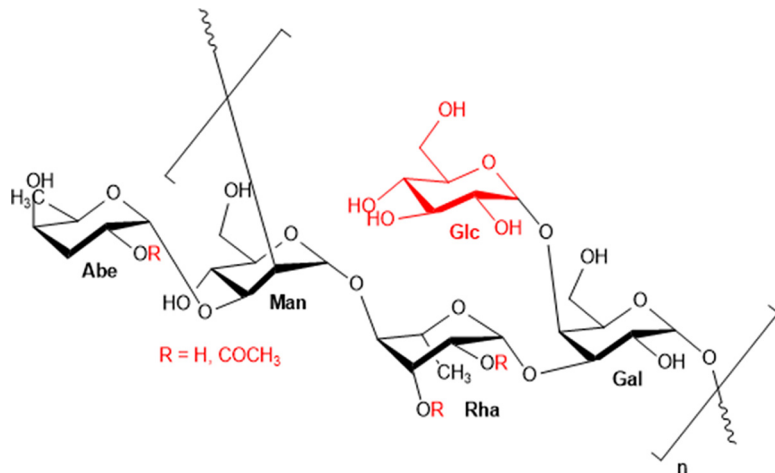




**FIG 5** Binding capacity and bactericidal activity of commercial anti-OAg antibodies. (A) FACS analysis of *S. Typhimurium* isolates using a rabbit polyclonal or a mouse monoclonal antibody. The mean fluorescence intensities (MFIs) of 10,000 acquired events are reported. (B) Serum bactericidal activity of commercial antibodies against the selected *S. Typhimurium* strains. IC<sub>50</sub> values, corresponding to 50% bacterial growth inhibition, are reported.

A lack of O-acetylation of the Rha residue or a loss of the full OAg, however, was observed but at lower rates and is linked to a diverse set of mutational events. As these events were observed in single isolates scattered across the phylogenetic tree, the clinical importance of these variations remains unclear, and these mutations are also likely to be an artifact caused by passage in the laboratory.

Variation in the *S. Typhimurium* OAg can impact infection, as it was previously shown



**FIG 6** *S. Typhimurium* O-antigen repeating unit. The observed modifications are highlighted in red: O-acetylation on Abe or Rha residues and glucosylation.

to play a role in the interaction with the host. The length of the OAg affects colitis and impacts the interaction with macrophages and complement (32, 33). Phase-variable chain length has been shown to be a trade-off between virulence and bacteriophage resistance (34), while phase-variable glucosylation acts as a defense against bacteriophages (35). Unfortunately, no tissue samples or clinical details from patients were available to inform on whether these O:5<sup>-</sup> isolates caused an altered host response or a more severe disease outcome.

Interestingly, a recent study identified *S. Typhimurium* IgA escape mutants in mice that were vaccinated but diseased (36). OAg variants similar to the ones that we identified in the DRC were found among the escape mutants, including a loss of Abe O-acetylation and a gain of  $\alpha$ -(1→4)-linked glucosylation of galactose. These variants also emerged in mice at later stages of chronic infections, suggesting that they also occurred despite the animals having had time to develop an acquired immune response to the bacteria. We hypothesize that the OAg variants in invasive *S. Typhimurium* isolates from DRC might emerge as a result of immune selective pressure following secondary to multiple *S. Typhimurium* infections in a region where iNTS disease is endemic.

O-acetylation of the O:5 antigen impacts the binding of different antibodies, including various monoclonal (37), polyclonal (38), and O:4 antigen (39) antibodies. In mice immunized with an *S. Typhimurium* strain possessing an intact and functional *oafA* gene, the median antibody titer against Abe-O-acetylated LPS was 32-fold higher than the titer against de-O-acetylated LPS. Mice immunized with a strain carrying a disrupted *oafA* gene showed an 8-fold-higher titer against de-O-acetylated LPS. Thus, OAg O-acetylation can alter recognition by some antibodies targeting *S. Typhimurium* OAg (38). This differential antibody recognition of OAg had a statistically significant correlation with protection against subsequent challenge with *S. Typhimurium* *oafA*<sup>+</sup> and *oafA*-deleted strains, carrying an insertion that truncated the O-acetylation (38). It has also been suggested that the three-dimensional (3D) structure of LPS can be altered upon Abe O-acetylation (37). Less is known about the role of the O-acetylation of Rha. In 2015, Lanzilao et al. (40) reported that D23580 OAg, which is O-acetylated on Rha, when conjugated to the carrier protein CRM<sub>197</sub>, induced antibodies that bound the homologous OAg more efficiently than OAg O-acetylated only at Abe but maintained bactericidal activity against *S. Typhimurium* isolates that expressed different OAg forms. *S. Typhimurium* OAg O-acetyls were confirmed to be immunodominant epitopes in more recent work by Baliban et al. (18). Those authors suggest that antibodies to O-acetylated epitopes are induced under conditions where protection is achieved and that immunization with de-O-acetylated OAg glycoconjugates provides significant but reduced protection against *S. Typhimurium* challenge.

O-Acetyl groups are important immune determinants for several bacterial polysaccharide vaccines (41). Konadu et al. (42) reported that O-acetyl groups are required in order to elicit anti-LPS antibodies with bactericidal activity against *S. enterica* serovar Paratyphi A. Sera from mice injected with conjugates of LPS detoxified by hydrazinolysis, which removes O-acetyl groups, had no detectable anti-LPS antibodies and had no bactericidal activity. Also, in *S. enterica* serovar Typhi LPS OAg, glucosylation and O-acetylation impacted serum resistance and antibody recognition and were hypothesized to be immune evasion mechanisms (43).

Our study highlights substantial convergent evolution within the recent *S. Typhimurium* population in the DRC toward O:5 loss, alongside O-acetylation of Rha. Importantly, our results confirmed that O:5<sup>-</sup> isolates show decreased antibody binding and subsequent complement-mediated killing compared to their O:5<sup>+</sup> counterparts when using a mouse monoclonal antibody. Although no information is known about the epitope recognized by this commercial monoclonal antibody, the presence of O-acetylation on Abe significantly impacts antibody recognition and killing of *S. Typhimurium* strains. The availability of other well-characterized monoclonal antibodies will allow further investigation of their ability to recognize different OAg structures. Moreover, there is no information on how a monoclonal antibody can represent the complex polyclonal response after vaccination and how antibodies elicited

by O:5<sup>+</sup> *S. Typhimurium* OAg can protect against circulating O:5<sup>−</sup> isolates. The observed variability in OAg O-acetylation patterns among *S. Typhimurium* clinical isolates may have implications for the optimal design of OAg-based vaccines, and this aspect needs to be further explored.

## MATERIALS AND METHODS

**Ethics statement.** Ethical approval for microbiological surveillance was granted by the Institutional Review Board at the Institute of Tropical Medicine (ITM) in Antwerp, Belgium, and by the Ethics Committees of Antwerp University (Belgium) and the School of Public Health, Kinshasa, DRC.

**Study setting.** *S. Typhimurium* blood isolates from the DRC were available from two collections. The DRC is the largest country in Central Africa, and some surveillance sites are >2,000 km apart without highways connecting them. Isolates from 2002, 2005, 2006, 2007, and 2008 originating from Lwiro (Sud Kivu province, near the Rwandese border in the eastern DRC) were included in this study. These isolates had been stored at the Saint Pierre Hospital in Brussels, Belgium. Additionally, blood culture surveillance in the DRC has been organized by the teams of the National Institute of Biomedical Research (INRB) and the ITM from 2007 to 2017 at the referral hospital of Saint Luc in Kisantu (Kongo-Central province, western DRC); the University Hospital of Kinshasa, the referral Hospital St. Joseph, and Monkole Hospital in Kinshasa (Kinshasa province, western DRC); the referral hospital of Bwamanda (Sud Ubangi province, northwestern DRC); and the referral hospital of Kabondo and the University Hospital of Kisangani (CUKIS) in Kisangani (Tshopo province, northeastern DRC) (24, 44–46). Most health facilities across the country lack the capacity for diagnosing bacterial bloodstream infections, and we therefore have no information about bloodstream infections elsewhere in the DRC except as part of outbreak research (47, 48). The surveillance sites included in this study were not consistently active because of stock shortages, staff movements, and funding and security issues.

**Bacterial isolates.** In this study, 354 *S. Typhimurium* isolates were collected between 2002 and 2017 and are listed in Table S1 in the supplemental material. Twenty-eight isolates were collected between 2002 and 2008 in Lwiro in the South Kivu province of the DRC. A total of 326 isolates were collected between 2007 and 2017 in the provinces of Kongo-Central ( $n = 181$ ), Kinshasa ( $n = 33$ ), Sud-Ubangi ( $n = 58$ ), Tshopo ( $n = 49$ ), Sankuru ( $n = 1$ ), Maniema ( $n = 1$ ), Bas-Uele ( $n = 2$ ), and Mai-Ndombe ( $n = 1$ ).

All isolates were stored in tubes of Trypticase soya agar (Oxoid, Basingstoke, UK) and shipped to the ITM for confirmation and further identification. At the ITM in Belgium, isolates biochemically confirmed as *Salmonella* spp. were serotyped using commercial antisera (Sifin, Berlin, Germany) according to the Kauffmann-White scheme (49), including O:5 reactivity. A representative selection of the isolates (10%) was sent to the National Reference Centre Sciensano in Belgium for confirmation of serotype.

**Stability of O:5 specificity over generations.** The stability of O:5 specificity over time and generations was studied for three random O:5<sup>−</sup> and three O:5<sup>+</sup> isolates. The outcome would then support whether the O:5 specificity is a bistable phenotype or has an underlying genetic cause. Bacteria were repetitively inoculated onto blood agar plates and incubated overnight. Three colonies per plate were subcultured daily 12 times by inoculating them onto fresh blood agar or Mueller-Hinton (MH) agar plates. O:5 serotyping was performed after the 1st, 2nd, 3rd, 4th, 8th, and 12th inoculations.

**Illumina whole-genome sequencing and single nucleotide polymorphism analysis.** Whole-genome sequencing data were obtained as described previously (4). In brief, DNA for Illumina sequencing was extracted using the Gentra PureGene Yeast/Bact kit (Qiagen, Hilden, Germany), according to the manufacturer's guidelines. Genomic DNA was then subjected to indexed whole-genome sequencing on an Illumina HiSeq 2500 platform at the Wellcome Sanger Institute to generate paired-end reads of 100 to 150 bp in length.

Illumina HiSeq reads were mapped to the *S. Typhimurium* reference genome of ST313 lineage II (D23580) (GenBank accession number [FN424405.1](#)) (50) using SMALT v0.7.4 to produce a BAM file. SMALT was used to index the reference using a kmer size of 20 and a step size of 13, and the reads were aligned using default parameters but with the maximum insert size set as 3 times the mean fragment size of the sequencing library. PCR duplicate reads were identified using Picard v1.92 (Broad Institute, Cambridge, MA, USA) and flagged as duplicates in the BAM file.

Variation detection was performed using samtools mpileup v0.1.19 with parameters “-d 1000 -DSugBf” and bcftools v0.1.19 (51) to produce a binary variant call format (BCF) file of all variant sites. The option to call genotypes at variant sites was passed to the bcftools call. All bases were filtered to remove those with uncertainty in the base call. The bcftools variant quality score was required to be greater than 50, and the mapping quality (map\_quality) was required to be greater than 30. If not all reads gave the same base call, the allele frequency, as calculated by bcftools, was required to be either 0 for bases called the same as the reference or 1 for bases called a single nucleotide polymorphism (SNP) ( $af1 < 0.95$ ). The majority base call was required to be present in at least 75% of reads mapping at the base (ratio of <0.75), and the minimum mapping depth required was 4 reads, at least 2 of which had to map to each strand (depth of <4; depth\_strand of <2). Finally, strand\_bias was required to be less than 0.001, map\_bias was required to be less than 0.001, and tail\_bias was required to be less than 0.001. If any of these filters were not met, the base was called uncertain.

**Phylogenetic analysis.** A pseudogenome was constructed by substituting the base call at each site (variant and nonvariant) in the BCF file into the reference genome, and any site called uncertain was replaced with an N. Insertions with respect to the reference genome were ignored, and deletions with

respect to the reference genome were filled with N's in the pseudogenome to keep it aligned and the same length as the reference genome used for read mapping.

Recombinant regions in the chromosome such as prophage regions and the *fljB* coding sequence (CDS) were checked using Gubbins v1.4.10 (52) and removed from the alignment. SNP sites were extracted from the alignment using snp-sites software (53) and used to construct a maximum likelihood phylogeny. RAxML v8.2.8 (54) with the substitution model GTRCAT was used. Support for nodes on the trees was assessed using 1,000 bootstrap replicates. The tree was rooted to *S. Typhimurium* ST313 strain DT2. Trees were visualized using Figtree v1.4.2 and iTOL (55).

**Pangenome analysis.** The Illumina data were *de novo* assembled using VelvetOptimiser v.2.2.5 and Velvet v1.2.10 (56). Assemblies were improved by scaffolding with the best  $N_{50}$  and contigs using SSPACE v2.0 (57), and sequence gaps were filled with GapFiller v1.11 (58). Assemblies were annotated using PROKKA v1.5 (59) and a *Salmonella*-specific database from RefSeq (59).

The pangenome of the full data set was constructed using Roary (60), with a BLASTp percent identity of 95% and a core definition of 99%. Genes significantly associated with the O:5 phenotype as determined using serotyping were identified using Scoary v.1.6.16 (60).

Conservation of the *oafA* gene across all assemblies was analyzed using BLASTN 2.6.0+ using the intact *oafA* sequence of O:5<sup>+</sup> isolate 10808/3 as the query and the assemblies of all isolates as subjects.

**MinION long-read sequencing and genomic analysis.** *S. Typhimurium* isolates 18034/3, 4701/4, 6088/3, 7123/11, 11480/3, 1304, and 3832/3 were grown overnight on MH agar at 37°C. For Oxford Nanopore sequencing, genomic DNA was extracted using the Epicentre MasterPure complete DNA and RNA purification kit (Lucigen, Middleton, WI, USA). A multiplex library was prepared using the 1D ligation sequencing kit (catalog number SQK-LSK109) and the native barcode expansion kit (catalog number EXP-NBD104) (Oxford Nanopore Technologies [ONT], Oxford, UK). Sequencing was performed using the Flo-Min110 flow cell on a MinION Mk1B device. Guppy 4.0.14 (ONT, Oxford, UK) was used for base calling (fast mode) and demultiplexing. Reads were trimmed using Porechop v0.2.4 (61), and high-quality reads were filtered using Filtlong v0.2.0 (62). Hybrid *de novo* assemblies of Illumina and Nanopore reads were generated using Unicycler v0.4.8 (default parameters) (63). Easyfig v2.2.3 (64) and Inkscape v0.92 were used to create figures comparing genomic regions. Insertion sequence (IS) elements were annotated using ISfinder (65).

**Selection of isolates for physicochemical analysis.** A selection of isolates was subjected to further analysis to reveal the physicochemical structure of the OAg. O:5<sup>+</sup> and O:5<sup>-</sup> isolates covering the different types of *oafA* recombination events and encompassing spatiotemporal variation in the collection were selected for further analysis. All O:5<sup>-</sup> isolates with intact *oafA* genes were additionally included in the analysis. In total, 11 O:5<sup>+</sup> and 14 O:5<sup>-</sup> isolates were subjected to physicochemical analysis.

**Physicochemical characterization.** All isolates were grown at 30°C in liquid Luria-Bertani (LB) medium overnight in a rotary shaker at 180 rpm. OAg extraction was performed directly on bacterial cells by acid hydrolysis as previously described (66). In particular, the bacterial pellet was separated by centrifugation, resuspended in 1% acetic acid, and heated at 100°C for 2 h. Supernatants containing OAg were dried, resuspended in water, and purified using a PD10 column prepacked with Sephadex G-25 Superfine (GE Healthcare, Marlborough, MA, USA). Part of the samples was lyophilized and exchanged twice with 99.9% deuterium oxide (D<sub>2</sub>O), dissolved in 600  $\mu$ L of D<sub>2</sub>O, and introduced into a 5-mm nuclear magnetic resonance (NMR) tube for data acquisition. <sup>1</sup>H NMR spectroscopy was used to confirm the OAg identity and determine the O-acetylation pattern (21). Spectra were recorded at 303 K with a Bruker Avance III 400 spectrometer using standard pulse sequences. <sup>1</sup>H NMR spectra were recorded at 400 MHz, chemical shift values are reported in parts per million, and the solvent peak for D<sub>2</sub>O was calibrated at 4.70 ppm. Data acquisition and processing were performed with the TopSpin 3.5 software package (Bruker BioSpin).

High-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) was used to verify the expected OAg composition, evaluate the OAg amount (expressed as micrograms of OAg per bacterial optical density [OD] unit), and quantify the percentage of glucosylation (21).

OAg samples were also characterized by high-performance liquid chromatography–size exclusion chromatography (HPLC-SEC) with differential refractive index (dRI) detection to estimate the molecular size distribution. The OAg samples were run on a TSK gel G3000 PWXL column (30 cm by 7.8 mm, particle size of 7  $\mu$ m, catalog number 808021) with a TSK gel PWXL guard column (4.0 cm by 6.0 mm, particle size of 12  $\mu$ m, catalog number 808033) (Tosoh Bioscience Tokyo, Japan). The mobile phase was composed of 0.1 M NaCl, 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, and 5% CH<sub>3</sub>CN (pH 7.2) at a flow rate of 0.5 mL/min (isocratic method for 35 min). The OAg peak molecular weight was calculated using dextrans as standards in the range of 12 to 150 kDa.

Some samples were selected for composition and linkage analysis by the use of gas-liquid chromatography coupled with mass spectrometry (GLC-MS) and <sup>1</sup>H NMR analysis for quantifying the O-acetylation level. Such samples were further purified by gel filtration chromatography on a HiPrep 16/60 Sephacryl S300 HR column (600 by 16 mm; GE Healthcare, Marlborough, MA, USA), with elution with phosphate-buffered saline (PBS) at 0.5 mL/min.

For the determination of the glucosylation level of Gal, the OAg samples were subjected to linkage analysis: 0.5 mg of each sample was permethylated (67), hydrolyzed with 2 M trifluoroacetic acid at 125°C, and derivatized to alditol acetates (68). The mixtures of partially methylated alditol acetates obtained were subjected to GLC analysis on an Agilent Technologies 6850 gas chromatograph equipped with a flame ionization detector and using helium as the carrier gas. Separation was performed using an HP-1 capillary column (30 m; Agilent Technologies) with the following temperature program: 1 min at 125°C, 125°C to 240°C at 4°C/min, and 2 min at 240°C. GLC-MS analyses were carried out on an Agilent

Technologies 7890A gas chromatograph coupled to an Agilent Technologies 5975C VL MSD, using the same experimental conditions as the ones described above. Integration values of the areas of the partially methylated alditol acetates were corrected by the effective carbon response factors (69).

**Antibody binding to bacteria and bactericidal activity.** Fluorescence-activated cell sorting (FACS) analysis was performed as previously described (70). Bacteria were pelleted, washed with PBS, and then blocked with PBS containing 3% (wt/vol) bovine serum albumin (BSA) for 15 min. Two commercial antibodies targeting *Salmonella* Typhimurium LPS were used: a mouse monoclonal antibody (catalog number ab8274; Abcam) and a rabbit polyclonal serum (anti-O:4, catalog number 294401; Denka Seiken). Bacteria were incubated with antibodies diluted in PBS plus 1% (wt/vol) BSA (1:2,000) for 1 h. After washes, samples were incubated with Alexa Fluor 647 goat anti-mouse IgG or Alexa Fluor 488 goat anti-rabbit IgG (1:500) (Molecular Probes) for 30 min. Finally, bacteria were fixed with 4% (wt/vol) formaldehyde for 20 min, and flow cytometry analysis was performed using a FACSCanto II flow cytometer (BD Biosciences).

The same antibodies were tested against the selected bacterial strains in a serum bactericidal assay (SBA) based on the luminescence readout (71), as previously described (31). The results of the assay were expressed as the 50% inhibitory concentration ( $IC_{50}$ ), the reciprocal serum dilution that resulted in a 50% reduction of luminescence, thus corresponding to 50% growth inhibition of the bacteria present in the assay mixture. GraphPad Prism 7 software was used for curve fitting and  $IC_{50}$  determination.

**Data availability.** Newly generated sequencing data are publicly available at the European Nucleotide Archive (study accession number [PRJEB20135](https://www.ebi.ac.uk/ena/record/PRJEB20135)).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, EPS file, 0.6 MB.

**FIG S2**, TIF file, 0.7 MB.

**TABLE S1**, XLSX file, 0.02 MB.

**TABLE S2**, DOCX file, 0.03 MB.

**TABLE S3**, DOCX file, 0.03 MB.

## ACKNOWLEDGMENTS

This work was supported by the Bacterial Vaccines (BactiVac) Network funded by the GCRF Networks in Vaccines Research and Development, which was cofunded by the MRC and BBSRC with additional support from the Department of Health and Social Care's Global AMR Innovation Fund (GAMRIF). This research was funded by the National Institute for Health Research [Cambridge Biomedical Research Centre at the Cambridge University Hospitals NHS Foundation Trust]. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

G.G., M.M.R., and F.M. are employed by the GSK group of companies.

## REFERENCES

- GBD 2017 Non-Typhoidal Salmonella Invasive Disease Collaborators. 2019. The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infect Dis* 19:1312–1324. [https://doi.org/10.1016/S1473-3099\(19\)30418-9](https://doi.org/10.1016/S1473-3099(19)30418-9).
- Gilchrist JJ, MacLennan CA. 2019. Invasive nontyphoidal *Salmonella* disease in Africa. *EcoSal Plus* 8(2):ESP-0007-2018. <https://doi.org/10.1128/ecosalplus.ESP-0007-2018>.
- Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, Al-Mashhadani MN, Kariuki S, Msefula CL, Gordon MA, de Pinna E, Wain J, Heyderman RS, Obaro S, Alonso PL, Mandomando I, MacLennan CA, Tapia MD, Levine MM, Tennant SM, Parkhill J, Dougan G. 2012. Intracontinental spread of human invasive *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. *Nat Genet* 44:1215–1221. <https://doi.org/10.1038/ng.2423>.
- Van Puyvelde S, Pickard D, Vandelannoote K, Heinz E, Barbe B, de Block T, Clare S, Coomber EL, Harcourt K, Sridhar S, Lees EA, Wheeler NE, Klemm EJ, Kuijpers L, Mbuyi Kalonji L, Phoba MF, Falay D, Ngbondia D, Lunguya O, Jacobs J, Dougan G, Deborggraeve S. 2019. An African *Salmonella* Typhimurium ST313 sublineage with extensive drug-resistance and signatures of host adaptation. *Nat Commun* 10:4280. <https://doi.org/10.1038/s41467-019-11844-z>.
- Post AS, Diallo SN, Guiraud I, Lompo P, Tahita MC, Maltha J, Van Puyvelde S, Mattheus W, Ley B, Thriemer K, Rouamba E, Derra K, Deborggraeve S, Tinto H, Jacobs J. 2019. Supporting evidence for a human reservoir of invasive non-typhoidal *Salmonella* from household samples in Burkina Faso. *PLoS Negl Trop Dis* 13:e0007782. <https://doi.org/10.1371/journal.pntd.0007782>.
- Phoba MF, Barbe B, Ley B, Van Puyvelde S, Post A, Mattheus W, Deborggraeve S, Lunguya O, Jacobs J. 2020. High genetic similarity between non-typhoidal *Salmonella* isolated from paired blood and stool samples of children in the Democratic Republic of the Congo. *PLoS Negl Trop Dis* 14:e0008377. <https://doi.org/10.1371/journal.pntd.0008377>.
- Kariuki S, Mbae C, Van Puyvelde S, Onsare R, Kawai S, Wairimu C, Ngetich R, Clemens J, Dougan G. 2020. High relatedness of invasive multi-drug resistant non-typhoidal *Salmonella* genotypes among patients and asymptomatic carriers in endemic informal settlements in Kenya. *PLoS Negl Trop Dis* 14:e0008440. <https://doi.org/10.1371/journal.pntd.0008440>.
- Tack B, Vanaenrode J, Verbakel JY, Toelen J, Jacobs J. 2020. Invasive non-typhoidal *Salmonella* infections in sub-Saharan Africa: a systematic review on antimicrobial resistance and treatment. *BMC Med* 18:212. <https://doi.org/10.1186/s12916-020-01652-4>.
- Tennant SM, MacLennan CA, Simon R, Martin LB, Khan MI. 2016. Nontyphoidal salmonella disease: current status of vaccine research and development. *Vaccine* 34:2907–2910. <https://doi.org/10.1016/j.vaccine.2016.03.072>.
- Rondini S, Lanzilao L, Necchi F, O'Shaughnessy CM, Micoli F, Saul A, MacLennan CA. 2013. Invasive African *Salmonella* Typhimurium induces bactericidal antibodies against O-antigens. *Microb Pathog* 63:19–23. <https://doi.org/10.1016/j.micpath.2013.05.014>.
- Watson DC, Robbins JB, Szu SC. 1992. Protection of mice against *Salmonella* typhimurium with an O-specific polysaccharide-protein conjugate vaccine. *Infect Immun* 60:4679–4686. <https://doi.org/10.1128/iai.60.11.4679-4686.1992>.



12. Grossman N, Schmetz MA, Foulds J, Klima EN, Jimenez-Lucho VE, Leive LL, Joiner KA, Jimenez V. 1987. Lipopolysaccharide size and distribution determine serum resistance in *Salmonella* montevideo. *J Bacteriol* 169: 856–863. <https://doi.org/10.1128/jb.169.2.856-863.1987>.
13. Svenson SB, Nurminen M, Lindberg AA. 1979. Artificial *Salmonella* vaccines: O-antigenic oligosaccharide-protein conjugates induce protection against infection with *Salmonella* typhimurium. *Infect Immun* 25: 863–872. <https://doi.org/10.1128/iai.25.3.863-872.1979>.
14. Svenson SB, Lindberg AA. 1981. Artificial *Salmonella* vaccines: *Salmonella* typhimurium O-antigen-specific oligosaccharide-protein conjugates elicit protective antibodies in rabbits and mice. *Infect Immun* 32:490–496. <https://doi.org/10.1128/iai.32.2.490-496.1981>.
15. Micoli F, Rondini S, Alfani R, Lanzilao L, Necchi F, Negrea A, Rossi O, Brandt C, Clare S, Mastroeni P, Rappuoli R, Saul A, MacLennan CA. 2018. Comparative immunogenicity and efficacy of equivalent outer membrane vesicle and glycoconjugate vaccines against nontyphoidal *Salmonella*. *Proc Natl Acad Sci U S A* 115:10428–10433. <https://doi.org/10.1073/pnas.1807655115>.
16. De Benedetto G, Alfani R, Cescutti P, Caboni M, Lanzilao L, Necchi F, Saul A, MacLennan CA, Rondini S, Micoli F. 2017. Characterization of O-antigen delivered by generalized modules for membrane antigens (GMMAs) vaccine candidates against nontyphoidal *Salmonella*. *Vaccine* 35:419–426. <https://doi.org/10.1016/j.vaccine.2016.11.089>.
17. Goh YS, Clare S, Micoli F, Saul A, Mastroeni P, MacLennan CA. 2015. Monoclonal antibodies of a diverse isotype induced by an O-antigen glycoconjugate vaccine mediate in vitro and in vivo killing of African invasive nontyphoidal *Salmonella*. *Infect Immun* 83:3722–3731. <https://doi.org/10.1128/IAI.00547-15>.
18. Baliban SM, Yang M, Ramachandran G, Curtis B, Shridhar S, Laufer RS, Wang JY, Van Druff J, Higginson EE, Hegerle N, Varney KM, Galen JE, Tennant SM, Lees A, MacKerell AD, Jr, Levine MM, Simon R. 2017. Development of a glycoconjugate vaccine to prevent invasive *Salmonella* Typhimurium infections in sub-Saharan Africa. *PLoS Negl Trop Dis* 11: e0005493. <https://doi.org/10.1371/journal.pntd.0005493>.
19. Schager AE, Dominguez-Medina CC, Necchi F, Micoli F, Goh YS, Goodall M, Flores-Langarica A, Bobat S, Cook CNL, Arcuri M, Marini A, King LDW, Morris FC, Anderson G, Toellner KM, Henderson IR, Lopez-Macias C, MacLennan CA, Cunningham AF. 2018. IgG responses to porins and lipopolysaccharide within an outer membrane-based vaccine against nontyphoidal *Salmonella* develop at discordant rates. *mBio* 9(2):e02379-17. <https://doi.org/10.1128/mBio.02379-17>.
20. Wang L, Andrianopoulos K, Liu D, Popoff MY, Reeves PR. 2002. Extensive variation in the O-antigen gene cluster within one *Salmonella* enterica serogroup reveals an unexpected complex history. *J Bacteriol* 184: 1669–1677. <https://doi.org/10.1128/JB.184.6.1669-1677.2002>.
21. Micoli F, Ravenscroft N, Cescutti P, Stefanetti G, Londero S, Rondini S, MacLennan CA. 2014. Structural analysis of O-polysaccharide chains extracted from different *Salmonella* Typhimurium strains. *Carbohydr Res* 385:1–8. <https://doi.org/10.1016/j.carres.2013.12.003>.
22. Hegde NV, Cook ML, Wolfgang DR, Love BC, Maddox CC, Jayarao BM. 2005. Dissemination of *Salmonella* enterica subsp. enterica serovar Typhimurium var. Copenhagen clonal types through a contract heifer-raising operation. *J Clin Microbiol* 43:4208–4211. <https://doi.org/10.1128/JCM.43.8.4208-4211.2005>.
23. Frech G, Kehrenberg C, Schwarz S. 2003. Resistance phenotypes and genotypes of multiresistant *Salmonella* enterica subsp. enterica serovar Typhimurium var. Copenhagen isolates from animal sources. *J Antimicrob Chemother* 51:180–182. <https://doi.org/10.1093/jac/dkg058>.
24. Tack B, Phoba MF, Barbe B, Kalonji LM, Hardy L, Van Puyvelde S, Ingelbeen B, Falay D, Ngonda D, van der Sande MAB, Deborggraeve S, Jacobs J, Lunguya O. 2020. Non-typhoidal *Salmonella* bloodstream infections in Kisantu, DR Congo: emergence of O5-negative *Salmonella* Typhimurium and extensive drug resistance. *PLoS Negl Trop Dis* 14:e0008121. <https://doi.org/10.1371/journal.pntd.0008121>.
25. Slauch JM, Lee AA, Mahan MJ, Mekalanos JJ. 1996. Molecular characterization of the oafA locus responsible for acetylation of *Salmonella* typhimurium O-antigen: oafA is a member of a family of integral membrane trans-acylases. *J Bacteriol* 178:5904–5909. <https://doi.org/10.1128/jb.178.20.5904-5909.1996>.
26. Hauser E, Junker E, Helmuth R, Malorny B. 2011. Different mutations in the oafA gene lead to loss of O5-antigen expression in *Salmonella* enterica serovar Typhimurium. *J Appl Microbiol* 110:248–253. <https://doi.org/10.1111/j.1365-2672.2010.04877.x>.
27. Brahmabhatt HN, Wyk P, Quigley NB, Reeves PR. 1988. Complete physical map of the rfb gene cluster encoding biosynthetic enzymes for the O antigen of *Salmonella* typhimurium LT2. *J Bacteriol* 170:98–102. <https://doi.org/10.1128/jb.170.1.98-102.1988>.
28. Kintz E, Davies MR, Hammarlof DL, Canals R, Hinton JC, van der Woude MW. 2015. A BTP1 prophage gene present in invasive non-typhoidal *Salmonella* determines composition and length of the O-antigen of the lipopolysaccharide. *Mol Microbiol* 96:263–275. <https://doi.org/10.1111/mmi.12933>.
29. Owen SV, Wenner N, Canals R, Makumi A, Hammarlof DL, Gordon MA, Aertsen A, Feasey NA, Hinton JCD. 2017. Characterization of the prophage repertoire of African *Salmonella* Typhimurium ST313 reveals high levels of spontaneous induction of novel phage BTP1. *Front Microbiol* 8:235. <https://doi.org/10.3389/fmicb.2017.00235>.
30. Pearson CR, Tindall SN, Herman R, Jenkins HT, Bateman A, Thomas GH, Potts JR, Van der Woude MW. 2020. Acetylation of surface carbohydrates in bacterial pathogens requires coordinated action of a two-domain membrane-bound acyltransferase. *mBio* 11(4):e01364-20. <https://doi.org/10.1128/mBio.01364-20>.
31. Gasperini G, Raso MM, Arato V, Aruta MG, Cescutti P, Necchi F, Micoli F. 2021. Effect of O-antigen chain length regulation on the immunogenicity of *Shigella* and *Salmonella* generalized modules for membrane antigens (GMMAs). *Int J Mol Sci* 22:1309. <https://doi.org/10.3390/ijms22031309>.
32. Crawford RW, Keestra AM, Winter SE, Xavier MN, Tsois RM, Tolstikov V, Baumber AJ. 2012. Very long O-antigen chains enhance fitness during *Salmonella*-induced colitis by increasing bile resistance. *PLoS Pathog* 8: e1002918. <https://doi.org/10.1371/journal.ppat.1002918>.
33. Murray GL, Attridge SR, Morona R. 2006. Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of *Salmonella* enterica serovar Typhimurium with macrophages and complement. *J Bacteriol* 188: 2735–2739. <https://doi.org/10.1128/JB.188.7.2735-2739.2006>.
34. Cota I, Sanchez-Romero MA, Hernandez SB, Pucciarelli MG, Garcia-Del Portillo F, Casades J. 2015. Epigenetic control of *Salmonella* enterica O-antigen chain length: a tradeoff between virulence and bacteriophage resistance. *PLoS Genet* 11:e1005667. <https://doi.org/10.1371/journal.pgen.1005667>.
35. Kim M, Ryu S. 2012. Spontaneous and transient defence against bacteriophage by phase-variable glucosylation of O-antigen in *Salmonella* enterica serovar Typhimurium. *Mol Microbiol* 86:411–425. <https://doi.org/10.1111/j.1365-2958.2012.08202.x>.
36. Diard M, Bakkeren E, Lentsch V, Rocker A, Bekele NA, Hoces D, Aslani S, Arnoldini M, Bohi F, Schumann-Moor K, Adamcik J, Piccoli L, Lanzavecchia A, Stadtmueller BM, Donohue N, van der Woude MW, Hockenberry A, Viollier PH, Falquet L, Wuthrich D, Bonfiglio F, Loverdo C, Egli A, Zandomenighi G, Mezzenga R, Holst O, Meier BH, Hardt WD, Slack E. 2021. A rationally designed oral vaccine induces immunoglobulin A in the murine gut that directs the evolution of attenuated *Salmonella* variants. *Nat Microbiol* 6: 830–841. <https://doi.org/10.1038/s41564-021-00911-1>.
37. Slauch JM, Mahan MJ, Michetti P, Neutra MR, Mekalanos JJ. 1995. Acetylation (O-factor 5) affects the structural and immunological properties of *Salmonella* typhimurium lipopolysaccharide O antigen. *Infect Immun* 63: 437–441. <https://doi.org/10.1128/iai.63.2.437-441.1995>.
38. Kim ML, Slauch JM. 1999. Effect of acetylation (O-factor 5) on the polyclonal antibody response to *Salmonella* typhimurium O-antigen. *FEMS Immunol Med Microbiol* 26:83–92. <https://doi.org/10.1111/j.1574-695X.1999.tb01375.x>.
39. Nakai Y, Ito A, Ogawa Y, Aribam SD, Elsheimer-Matulova M, Shiraiwa K, Kisaka SMB, Hikono H, Nishikawa S, Akiba M, Kawahara K, Shimoji Y, Eguchi M. 2017. Determination of O:4 antigen-antibody affinity level in O:5 antigen positive and negative variants of *Salmonella* enterica serovar Typhimurium. *FEMS Microbiol Lett* 364:fnx062. <https://doi.org/10.1093/femsle/fnx062>.
40. Lanzilao L, Stefanetti G, Saul A, MacLennan CA, Micoli F, Rondini S. 2015. Strain selection for generation of O-antigen-based glycoconjugate vaccines against invasive nontyphoidal *Salmonella* disease. *PLoS One* 10: e0139847. <https://doi.org/10.1371/journal.pone.0139847>.
41. Berti F, De Ricco R, Rappuoli R. 2018. Role of O-acetylation in the immunogenicity of bacterial polysaccharide vaccines. *Molecules* 23:1340. <https://doi.org/10.3390/molecules23061340>.
42. Konadu E, Shiloach J, Bryla DA, Robbins JB, Szu SC. 1996. Synthesis, characterization, and immunological properties in mice of conjugates composed of detoxified lipopolysaccharide of *Salmonella* paratyphi A bound to tetanus toxoid with emphasis on the role of O acetyls. *Infect Immun* 64:2709–2715. <https://doi.org/10.1128/iai.64.7.2709-2715.1996>.
43. Kintz E, Heiss C, Black I, Donohue N, Brown N, Davies MR, Azadi P, Baker S, Kaye PM, van der Woude M. 2017. *Salmonella* enterica serovar Typhi lipopolysaccharide O-antigen modification impact on serum resistance and

- antibody recognition. *Infect Immun* 85:e01021-16. <https://doi.org/10.1128/IAI.01021-16>.
44. Lunguya O, Lejon V, Phoba MF, Bertrand S, Vanhoof R, Glupczynski Y, Verhaegen J, Muyembe-Tamfum JJ, Jacobs J. 2013. Antimicrobial resistance in invasive non-typhoid *Salmonella* from the Democratic Republic of the Congo: emergence of decreased fluoroquinolone susceptibility and extended-spectrum beta lactamases. *PLoS Negl Trop Dis* 7:e2103. <https://doi.org/10.1371/journal.pntd.0002103>.
  45. Ley B, Le Hello S, Lunguya O, Lejon V, Muyembe JJ, Weill FX, Jacobs J. 2014. Invasive *Salmonella enterica* serotype Typhimurium infections, Democratic Republic of the Congo, 2007–2011. *Emerg Infect Dis* 20:701–704. <https://doi.org/10.3201/eid2004.131488>.
  46. Kalonji LM, Post A, Phoba M-F, Falay D, Ngbonda D, Muyembe J-J, Bertrand S, Ceyssens P-J, Mattheus W, Verhaegen J, Barbe B, Kuijpers L, Van Geet C, Lunguya O, Jacobs J. 2015. Invasive *Salmonella* infections at multiple surveillance sites in the Democratic Republic of the Congo, 2011–2014. *Clin Infect Dis* 61(Suppl 4):S346–S353. <https://doi.org/10.1093/cid/civ713>.
  47. Phoba MF, Lunguya O, Mayimon DV, Lewo di Mputu P, Bertrand S, Vanhoof R, Verhaegen J, Van Geet C, Muyembe JJ, Jacobs J. 2012. Multidrug-resistant *Salmonella enterica*, Democratic Republic of the Congo. *Emerg Infect Dis* 18:1692–1694. <https://doi.org/10.3201/eid1810.120525>.
  48. Phoba MF, Barbe B, Lunguya O, Masendu L, Lulengwa D, Dougan G, Wong VK, Bertrand S, Ceyssens PJ, Jacobs J, Van Puyvelde S, Deborggraeve S. 2017. *Salmonella enterica* serovar Typhi producing CTX-M-15 extended spectrum beta-lactamase in the Democratic Republic of the Congo. *Clin Infect Dis* 65:1229–1231. <https://doi.org/10.1093/cid/cix342>.
  49. Grimont PAD, Weill F-X. 2007. Antigenic formulae of the *Salmonellae* serovars. World Health Organization, Geneva, Switzerland.
  50. Kingsley RA, Msefule CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, Harris D, Clarke L, Whitehead S, Sangal V, Marsh K, Achtman M, Molyneux ME, Cormican M, Parkhill J, MacLennan CA, Heyderman RS, Dougan G. 2009. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res* 19:2279–2287. <https://doi.org/10.1101/gr.091017.109>.
  51. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
  52. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 43:e15. <https://doi.org/10.1093/nar/gku1196>.
  53. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, Harris SR. 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom* 2:e000056. <https://doi.org/10.1099/mgen.0.000056>.
  54. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
  55. Letunic I, Bork P. 2016. Interactive Tree of Life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44:W242–W245. <https://doi.org/10.1093/nar/gkw290>.
  56. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
  57. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <https://doi.org/10.1093/bioinformatics/btq683>.
  58. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with Gap-Filler. *Genome Biol* 13:R56. <https://doi.org/10.1186/gb-2012-13-6-r56>.
  59. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
  60. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>.
  61. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial assemblies with multiplex MinION sequencing. *Microb Genom* 3:e000132. <https://doi.org/10.1099/mgen.0.000132>.
  62. Wick RR. 2021. Filtlong v.0.2.0. <https://github.com/rwick/Filtlong>.
  63. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
  64. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* 27:1009–1010. <https://doi.org/10.1093/bioinformatics/btr039>.
  65. Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J Comput Biol* 7:203–214. <https://doi.org/10.1089/10665270050081478>.
  66. Micoli F, Rondini S, Gavini M, Pisoni I, Lanzilao L, Colucci AM, Giannelli C, Pippi F, Sollai L, Pinto V, Berti F, MacLennan CA, Martin LB, Saul A. 2013. A scalable method for O-antigen purification applied to various *Salmonella* serovars. *Anal Biochem* 434:136–145. <https://doi.org/10.1016/j.ab.2012.10.038>.
  67. Harris PJ, Henry RJ, Blakeney AB, Stone BA. 1984. An improved procedure for the methylation analysis of oligosaccharides and polysaccharides. *Carbohydr Res* 127:59–73. [https://doi.org/10.1016/0008-6215\(84\)85106-X](https://doi.org/10.1016/0008-6215(84)85106-X).
  68. Albersheim P, Nevins DJ, English PD, Karr A. 1967. A method for the analysis of sugars in plant cell-wall polysaccharides by gas-liquid chromatography. *Carbohydr Res* 5:340–345. [https://doi.org/10.1016/S0008-6215\(00\)80510-8](https://doi.org/10.1016/S0008-6215(00)80510-8).
  69. Sweet DP, Shapiro RH, Albersheim P. 1975. Quantitative analysis by various g.l.c. response-factor theories for partially methylated and partially ethylated alditol acetates. *Carbohydr Res* 40:217–225. [https://doi.org/10.1016/S0008-6215\(00\)82604-X](https://doi.org/10.1016/S0008-6215(00)82604-X).
  70. Mancini F, Gasperini G, Rossi O, Aruta MG, Raso MM, Alfini R, Biagini M, Necchi F, Micoli F. 2021. Dissecting the contribution of O-Antigen and proteins to the immunogenicity of *Shigella sonnei* generalized modules for membrane antigens (GMMAs). *Sci Rep* 11:906. <https://doi.org/10.1038/s41598-020-80421-y>.
  71. Necchi F, Saul A, Rondini S. 2017. Development of a high-throughput method to evaluate serum bactericidal activity using bacterial ATP measurement as survival readout. *PLoS One* 12:e0172163. <https://doi.org/10.1371/journal.pone.0172163>.