1a →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
| 3/4(*oacB*) ↑ 4(*gtrI*)
OAc α-D-Glc*p*

1b →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
| 3/4(*oacB*) | 2(*oac*) ↑ 4(*gtrI*)
OAc OAc α-D-Glc*p*

1c(7a) →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑ 4(*gtrI*)
(*gtr1c*)
α-D-Glc*p*(1→ 2)α-D-Glc*p*

1d →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1
↑ 3(*gtrX*) ↑ 4(*gtrI*)
α-D-Glc*p* α-D-Glc*p*

2a →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
| 3/4(*oacB*) ↑ 4(*gtrII*) | 6(*oacD*)
OAc α-D-Glc*p* OAc

2b →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑ 3(*gtrX*) ↑ 4(*gtrII*) | 6(*oacD*)
α-D-Glc*p* α-D-Glc*p* OAc

3a →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑ 3(*gtrX*) | 2(*oac*) | 6(*oacD*)
α-D-Glc*p* OAc OAc

3b →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
| 2(*oac*)
OAc

4a →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑ 6(*gtrIV*)
α-D-Glc*p*

4av →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑3(*opt*) ↑ 6(*gtrIV*)
PEtN α-D-Glc*p*

4b →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
| 2(*oac*) ↑ 6(*gtrIV*)
OAc α-D-Glc*p*

5a →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
| 3/4(*oacB*) ↑ 3(*gtrV*)
OAc α-D-Glc*p*

5b →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑ 3(*gtrX*) ↑ 3(*gtrV*)
α-D-Glc*p* α-D-Glc*p*

X →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑ 3(*gtrX*) | 6(*oacD*)
α-D-Glc*p* OAc

Xv →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑ 3(*gtrX*) ↑3(*opt*)
α-D-Glc*p* PEtN

Y →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
| 3/4(*oacB*) | 6(*oacD*)
OAc OAc

Yv →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑3(*opt*) ↑3(*opt*) | 6(*oacD*)
PEtN PEtN OAc

6 →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-β-D-Gal*p*A-(1→3)-β-D-Glc*p*NAc-(1→
| 3/4(*oacC*)
OAc
OAc
| 2(*oac*)

7b →2)-α-L-Rha*p*$^{III}$-(1→2)-α-L-Rha*p*$^{II}$-(1→3)-α-L-Rha*p*$^{I}$-(1→3)-β-D-Glc*p*NAc-(1→
↑ 4(*gtrI*)
α-D-Glc*p*(1→ 2)α-D-Glc*p*
(*gtr1c*)

Figure S1:Chemical composition and specific O-antigen modification genes of different serotypes of *Shigella flexneri*. All serotypes except 6 share a common O-antigen consists of repeating units of tetrasaccharide. Serotypes differ by the addition of either glucosyl or O-acetyl groups (2-O-acetyl, 3/4-O-acetyl and 6-O-acetyl), or PEtN to different sugars within the tetrasaccharide repeat unit via the linkages indicated. Specific O-antige modification genes for different serotypes were indicated in parentheses.