Riboregulation in Neisseria meningitidis

Huis in 't Veld, R.A.G.

Citation for published version (APA):
Chapter 3

Deep Sequencing Whole Transcriptome Exploration of the $\sigma_E$ Regulon in Neisseria meningitidis

Robert Huis in 't Veld, Antonius Willemsen, Antonius van Kampen, Edward Bradley, Frank Baas, Yvonne Pannekoek, and Arie van der Ende

PLoS ONE, 2011
Chapter 3

Deep Sequencing Whole Transcriptome Exploration of the $\sigma^E$ Regulon in *Neisseria meningitidis*

Robert Huis in ’t Veld, Antonius Willemsen, Antonius van Kampen, Edward Bradley, Frank Baas, Yvonne Pannekoek, and Arie van der Ende

*PLoS ONE*, 2011
Deep Sequencing Whole Transcriptome Exploration of the σE Regulon in *N. meningitidis*

Contents

Abstract .......................................................................................................................... 34
Introduction ................................................................................................................... 35
Materials & Methods .................................................................................................... 36
  Bacterial strains and growth conditions .................................................................. 36
  Total RNA preparation ............................................................................................. 36
  SOLiD sequencing ..................................................................................................... 37
  qRT-PCR .................................................................................................................... 37
  Data analysis ............................................................................................................ 37
Results .......................................................................................................................... 39
  Output SOLiD ............................................................................................................ 39
  Differentially expressed genes ................................................................................. 40
  Differential transcription in intergenic regions ....................................................... 44
  Antisense transcription ............................................................................................ 46
Discussion ..................................................................................................................... 47
  Positive feedback loop of the σE operon and control of msrA/msrB expression 
  by σE ....................................................................................................................... 47
  Discovery of novel differentially expressed genes ................................................. 48
  Differential expression of non-coding RNA transcripts ....................................... 50
  ρ-independent terminator read-through ............................................................... 51
Conclusion ................................................................................................................... 52
References ................................................................................................................... 53
Supplementary data .................................................................................................... 58
Chapter 3

Abstract
Bacteria live in an ever-changing environment and must alter protein expression promptly to adapt to these changes and survive. Specific response genes that are regulated by a subset of alternative $\sigma^{70}$-like transcription factors have evolved in order to respond to this changing environment. Recently, we have described the existence of a $\sigma^{E}$ regulon including the anti-$\sigma$-factor MseR in the obligate human bacterial pathogen *Neisseria meningitidis*. To unravel the complete $\sigma^{E}$ regulon in *N. meningitidis*, we sequenced total RNA transcriptional content of wild type meningococci and compared it with that of mseR mutant cells ($\Delta$mseR) in which $\sigma^{E}$ is highly expressed. Eleven coding genes and one non-coding gene were found to be differentially expressed between H44/76 wildtype and H44/76$\Delta$mseR cells. Five of the 6 genes of the $\sigma^{E}$ operon, $msrA/msrB$, and the gene encoding a pepSY-associated TM helix family protein showed enhanced transcription, whilst aniA encoding a nitrite reductase and nspA encoding the vaccine candidate Neisserial surface protein A showed decreased transcription. Analysis of differential expression in IGRs showed enhanced transcription of a non-coding RNA molecule, identifying a $\sigma^{E}$ dependent small non-coding RNA. Together this constitutes the first complete exploration of an alternative $\sigma$-factor regulon in *N. meningitidis*. The results direct to a relatively small regulon indicative for a strictly defined response consistent with a relatively stable niche, the human throat, where *N. meningitidis* resides.
Introduction

Bacteria live in an ever-changing environment and must alter protein expression promptly to adapt to these changes and survive. Specific response genes that are regulated by a subset of the alternative σ^{70}-like sigma factors have evolved in order to respond to this changing environment\textsuperscript{1}. In reaction to specific external stimuli, these sigma factors recruit RNA polymerases to the appropriate response genes. The name extracytoplasmic function (ECF) sigma factors or σ^{E} factors refer to the fact that most of the genes under control of these σ-factors encode proteins residing in the outer membrane or periplasmic space\textsuperscript{2}.

Recently, using a proteomic approach, we have described the existence of a σ^{E} regulon in the obligate human bacterial pathogen \textit{Neisseria meningitidis}\textsuperscript{3}. The σ^{E} operon so far identified encompasses an operon consisting of 6 genes (NMB2140-NMB2145) among which the gene encoding σ^{E} itself (NMB2144). In addition \textit{msrA/msrB} (NMB0044) was also found to be subjected to regulation by σ^{E}, it encodes for methionine sulfoxide reductase, an enzyme repairing proteins exposed to reactive oxygen species (ROS)\textsuperscript{4,5}. Furthermore, the anti-σ^{E} factor \textit{MseR} (for \textit{Meningococcal SigmaE Regulator}; NMB2145) was identified. Deletion of \textit{mseR} results in overexpression of the other 5 genes in the σ^{E} operon, thereby increasing the expression of σ^{E} itself as well as \textit{msrA/msrB}\textsuperscript{3}. The proteomic exploration of the σ^{E} regulon of meningococci and a microarray gene expression analysis for the closely related human pathogen \textit{Neisseria gonorrhoeae} performed by others have revealed a surprisingly small regulon for the σ^{E} factor in \textit{Neisseria}\textsuperscript{6}. However, both methods have several inherent and technical limitations. Identification of proteins is limited to those that can be isolated efficiently, are within a certain size range resolvable with SDS-PAGE, and have to be sufficiently expressed to be both visible and able to be confidently identified using mass spectrometry. DNA microarrays are only able to detect products that can sufficiently hybridize with probes that have been created based on available genomic data, usually limited to \textit{in silico} predicted open reading frames (ORFs) of protein coding genes in their sense orientation. Moreover, reliability and reproducibility remains a concern\textsuperscript{7-9}. Both approaches are furthermore unsuitable to detect an important class of RNA regulators called small RNAs (sRNA)\textsuperscript{10,11}. Previous investigations in \textit{Salmonella enterica} and \textit{Escherichia coli} have shown several sRNAs to be regulated by their respective σ^{E} orthologues, making it essential to be able to detect these transcripts in any complete exploration of σ-factor regulation\textsuperscript{12,13}. RNA-seq using next generation sequencing has the potential
to overcome all mentioned limitations by offering an approach to identify all RNA species expressed within a cell without pre-selection\textsuperscript{14,15}. Therefore, in order to unravel the complete $\sigma^E$ regulon in *N. meningitidis*, we sequenced total RNA transcriptional content of wild type meningococci and compared it with that of mseR knock-out cells in which $\sigma^E$ is highly expressed.

Materials & Methods

**Bacterial strains and growth conditions**

*N. meningitidis* strain H44/76, B:P1.7,16:F3-3: ST-32 (cc32), is closely related to serogroup B strain MC58, belonging to the same clonal complex\textsuperscript{16-20}. *N. meningitidis* H44/76 in which mseR was replaced with the erythromycin resistance cassette *ermC* (H44/76ΔmseR) was described previously\textsuperscript{3}. Meningococci were grown on GC plates (Difco), supplemented with 1% (vol/vol) Vitox (Oxoid) at 37 °C in a humidified atmosphere of 5% CO\textsubscript{2}. Broth cultures were incubated in GC medium supplemented with 1% (vol/vol) Vitox on a gyratory shaker (180 rpm) at 37 °C. Where appropriate, plates were supplemented with erythromycin (5 µg/mL). Growth was monitored by measuring optical density of cultures at 530 nm (OD\textsubscript{530}, Pharmacia Biotech Ultraspec 2000) at regular intervals. For RNA isolation, used in both RNA-seq and qRT-PCR experiments, cells were grown to exponential growth phase (OD\textsubscript{530}=0.5-0.6).

**Total RNA preparation**

Immediately after sample removal, RNA was stabilized by adding 1/10\textsuperscript{th} volume of stop solution (95% ethanol/5% phenol pH 4.3) and samples were further processed after rapid cooling on ice\textsuperscript{21}. Isolation of total RNA was performed using an acidic hot phenol method adapted from\textsuperscript{22}. RNA molecules smaller than 200 nucleotides (nts), including 5S and tRNAs, were retained using classic acidic phenol/chloroform extractions for sample clean-up\textsuperscript{23}. The majority of 16S and 23S rRNA molecules were selectively removed using the MICROBExpress kit (Life Technologies, 2008). RNA quality was assessed by resolving samples on 1% agarose gels and measuring RNA Integrity Number using a 2100 Bioanalyzer (Agilent Technologies)\textsuperscript{24}. For RNA-seq experiments RNA with a RIN value ≥ 9.0 was used. Quantification was assessed by UV spectroscopy (Nanodrop 1000, Thermoscientific).
SOLiD sequencing

Transcriptome sequencing was performed according to the manufacturer’s protocol using 1 µg RNA isolated from H44/76 wild type (H44/76 wt) or H44/76ΔmseR cells (SOLiD Whole Transcriptome Analysis Kit, Life Technologies, 2009 Rev. F). Adaptor mix A was used for RNA-adaptor ligation, ensuring strand specificity. In short, this is achieved by ligating the fragmented RNA molecules using double stranded adapters with single stranded random hexamers overhangs either on the 5’ or the 3’ end. Subsequent sequencing is performed using primers that anneal to only one of these two adapters. Bead preparation was performed according to the SOLiD v3 Plus system’s instructions (Life Technologies, 2009). The capacity of a single full slide was used, using barcodes to individually tag two H44/76 wt and two H44/76ΔmseR technical replicates, consisting of separately constructed cDNA libraries.

qRT-PCR

After DNase treatment (Roche, 2008) 2.5 µg RNA was used as input for reverse transcription, performed as per ThermoScript kit instructions (Invitrogen, 2010) with a single adaptation: all procedures were performed in duplo, substituting RT with nuclease-free water in 1 reaction to serve as no-RT controls. The resulting RT+ and RT- cDNA samples were serially diluted and PCR amplification of the constitutively expressed outer membrane protein rmpM was used to titrate relative total cDNA concentrations of the H44/76 wt and H44/76ΔmseR samples. The cDNA samples were adjusted accordingly for use in subsequent real-time PCRs (Lightcycler, Roche) (see Supplementary Table 1 for primers used). Crossing point (Cp) values were calculated manually for all genes using the Fit Points Method in the LightCycler Software package (Roche, version 4.05)25,26. All experiments were carried out in triplo. To normalize target gene Cp values, mean Cp values of the internal control rmpM of both H44/76 wt and H44/76ΔmseR were deducted from the Cp value of the respective target genes. Unpaired two-sample t-tests were performed for assessing the statistical significance of the difference between the means of these normalized Cp values. This was followed by a one-tailed test to derive p-values with 0.05 chosen as the threshold for statistical significance.

Data analysis

The SOLiD Accuracy Enhancer Tool (SAET) was used to correct miscalled bases and subsequently the P2 sequencing adapter was trimmed using a tool provided
by Nicholas Socci (www.pyrodigm.com). The corrected and trimmed reads were mapped to the recently published H44/76 Whole Genome Shotgun (WGS) sequence (GenBank accession AEQZ00000000) using BWA19,27. This mapping was performed in two steps. First the reads were mapped with default BWA settings (version 0.5.7 r1310). Subsequently the unmapped but high quality reads are mapped in a second step with less stringent settings allowing for 2 mismatches in the first 25 nts and 4 mismatches for the total read (up to 50 bp). These mismatches were allowed to compensate for limited genome sequence errors, i.e., the reference H44/76 WGS genome sequence was obtained by 454 pyrosequencing prone to inaccurate sequencing of long homopolymer stretches, and to allow possible mapping of post-transcriptionally altered RNA molecules19. The results of these mappings were merged for further analysis using SamTools28. For every ORF and intergenic region (IGR) we counted the number of aligned reads using the Rsamtools package29,30. Reads that overlapped with two ORFs and/or IGRs both contributed to the read count of these two ORFs/IGRs. Reads that could not be unambiguously mapped to a single location were ignored in subsequent analysis. This concerned an estimated 2% of the genome involving repetitive sequences and duplicated genes like transposase elements (73 annotated), phage related proteins (13 annotated), iron regulated proteins FrpA/B/C/D (includes several pseudogenes), the iron uptake system TpbA/B (duplicated), maf type adhesins (duplicated), Type IV pili, and rRNA genes (one copy annotated) as H44/76 does not contain large duplicated regions19. Library normalization and differential expression analysis for all IGRs larger than 29 nts and all ORFs on both strands was performed using DEseq31. For ORFs this resulted in the separate analysis of both sense and anti-sense transcription. DEseq settings were optimized for the most conservative analysis of technical replicates as described by Anders et al.31. The library sizes of the different samples were adjusted based on the total number of reads per library, without the use of size factor estimation. The p-values resulting from the pairwise comparison between H44/76 wt and H44/76ΔmseR were adjusted by the Benjamini-Hochberg procedure to control the false discovery rate (FDR)32. Visual inspection of mapped reads was facilitated with the Integrative Genomics Viewer33.

In silico prediction of transcriptional start sites (TSSs) was performed using the Neural Network Promoter Prediction program (prokaryote organism setting), ρ-independent terminators were identified manually using Mfold or adapted from those predicted by TransTerm HP in N. meningitidis MCS834-36. All reported nt
positions are according to the published H44/76 whole genome shotgun sequence\(^1\). Potential mRNA targets (in the reference strain MC58) of sRNAs were predicted \textit{in silico} using TargetRNA, its default settings adjusted only to allow for G:U pairs in seed\(^3\).

**Results**

**Output SOLiD**

The total number of reads varied from 50 $\times$ 10\(^6\) to 92 $\times$ 10\(^6\) (\textbf{Table 1}). The proportion of mapped reads ranged from 50\% to 68\%, with the second less stringent mapping accounting for 4-6\% of all mapped reads. Thirty percent to 47\% of the reads mapped to tRNA or rRNA genes. In one technical replicate, H44/76ΔmseR replicate A, the proportion of reads mapped to tRNA/rRNA genes was lower compared to the other 3, while the proportion of unmapped reads was higher. This might be attributable to slightly different gel size fractioning; however this has no observable effect on the total reads mapped to the rest of the genome. Raw read mapping analysis (before library normalization) showed at least 1 read could be mapped to 2593/2598 (100\%) ORFs and 1635/ 2599 (63\%) IGRs, with 43\% of the positive and 48\% of the negative strand covered by 10 or more reads.

\textbf{Table 1. Mapping statistics of SOLiD output.} Two technical replicates were performed for both H44/76 wt and H44/76ΔmseR, resulting in a total of 280 $\times$ 10\(^6\) reads.

<table>
<thead>
<tr>
<th></th>
<th>H44/76 wt</th>
<th></th>
<th>H44/76ΔmseR</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>replicate A (%)</td>
<td>replicate B (%)</td>
<td>replicate A (%)</td>
<td>replicate B (%)</td>
</tr>
<tr>
<td><strong>Mapped to genome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excluding tRNA/rRNA</td>
<td>20 (24%)</td>
<td>12 (21%)</td>
<td>18 (20%)</td>
<td>12 (24%)</td>
</tr>
<tr>
<td>tRNA/rRNA only</td>
<td>37 (44%)</td>
<td>27 (47%)</td>
<td>28 (30%)</td>
<td>22 (44%)</td>
</tr>
<tr>
<td><strong>Unmapped</strong></td>
<td>27 (32%)</td>
<td>19 (33%)</td>
<td>46 (50%)</td>
<td>16 (32%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>84 (100%)</td>
<td>58 (100%)</td>
<td>92 (100%)</td>
<td>50 (100%)</td>
</tr>
</tbody>
</table>
Differentially expressed genes

Eleven coding genes and one non-coding gene were found to be differentially expressed between H44/76 wt and H44/76ΔmseR cells with a corrected p-value \(\leq 0.05\) (Table 2). In contrast to H44/76 wt cells, in H44/76ΔmseR cells no reads were detected that mapped to the mseR locus of the H44/76 reference sequence (Table 2)\(^1\). In addition, RNA from H44/76 wt cells did not contain reads that mapped to ermC, while H44/76ΔmseR cells in which mseR was replaced by ermC yielded reads that mapped to this locus (data not shown). Therefore, differential transcription of artificially introduced and removed genes was confidently identified. Previously, using RT-PCR and proteomics, we identified the σ\(E\) operon and msrA/msrB to be upregulated in H44/76ΔmseR cells\(^3\). In the current RNA-seq analysis the expression of msrA/msrB (Figure 1) and the genes comprising the σ\(E\) operon (Figure 2) was 81-fold and between 6-fold and 36-fold increased, respectively (Table 2). Of interest, four differentially expressed genes were identified which were hitherto not associated with σ\(E\). Of these, one was upregulated: NMH_0518 (analogous to NMB1721; encoding a hypothetical protein of the PepSY-associated TM helix family of proteins). Three genes, NMH_2270 (analogous to NMB0663; nspA), nirK (analogous to NMB1623; aniA), and NMH_1358 (analogous to NMB1053; opcA outer membrane protein), appeared to be downregulated (Table 2). However, the promoter of the latter gene contains a homopolymeric cytidine tract located at the -35 region which can be extended or shortened during DNA replication by a mechanism called slipped strand DNA replication\(^38,39\). As a result, opcA displays variable expression\(^38\). Indeed, classic Sanger re-sequencing of this location revealed the promoter of H44/76 wt opcA contained a tract of 14 cytosines, while that of H44/76ΔmseR contained 15, the latter causing severe reduced transcriptional efficiency\(^38\). Therefore, we considered differential expression of NMH_1358 most likely to be the result of unrelated phase-variation independent of σ\(E\) expression. A complete list of differentially expressed loci with uncorrected p-values \(\leq 0.01\) can be found in Supplementary Table 2. Among the 30 loci with differential sense expression, four clusters of genes and their associated IGRs, either physically arranged in an operon or with a functional relationship, were recognized; cluster 1 comprising three tRNA-Leu genes (NMH_1084, NMH_1085, and NMH_2191); cluster 2 consisting of genes encoding a restriction modification system (NMH_1821-NMH_1823), cluster 3 encoding a leukotoxin type I secretion system (NMH_1851, NMH_0534-NMH_0535) and cluster 4 encoding genes involved in a thiamin biosynthesis operon (NMH_2073, NMH_2074, NMH_2076, thiG). In these
clusters the difference in expression between H44/76 wt and H44/76ΔmseR cells varies between 2-fold and 3.4-fold. Statistical analysis of the total expression levels of these clusters (i.e., using the total read count of all genes in a cluster) did not result in additional significant differential expression considering the corrected p-value. Visual inspection of pairs of 100 randomly picked non-differentially expressed genes using the Integrative Genomics Viewer software revealed consistently highly similar patterns of read distribution between H44/76 wt and H44/76ΔmseR normalized libraries, indicating that variation between strains is limited.
Chapter 3

Table 2. Differentially expressed loci, ranked by locus number, between *N. meningitidis* H44/76 wt and H44/76ΔmseR.

*Reads/locus after library normalization by DEseq. Locus number as annotated in H44/76 WGS (AEQZ00000000). Gene name as annotated in H44/76 WGS (AEQZ00000000). All loci with Benjamini-Hochberg adjusted q-value ≤0.05. Where available *N. meningitidis* strain MC58 orthologues are noted. mseR artificially deleted.*

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>Reads/locus</th>
<th>Fold-change</th>
<th>p-value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMH_0518</td>
<td></td>
<td>1998</td>
<td>5686</td>
<td>2.8</td>
<td>pepSY-associated TM helix family protein</td>
</tr>
<tr>
<td>NMH_1358</td>
<td></td>
<td>3740</td>
<td>1403</td>
<td>0.4</td>
<td>outer membrane OpcA family protein</td>
</tr>
<tr>
<td>NMH_2149</td>
<td></td>
<td>5216</td>
<td>187292</td>
<td>35.9</td>
<td>doxX family protein, σE operon</td>
</tr>
<tr>
<td>NMH_02153</td>
<td></td>
<td>5042</td>
<td>60516</td>
<td>12.0</td>
<td>Hypothetical protein in σE operon</td>
</tr>
<tr>
<td>NMH_2154</td>
<td></td>
<td>4807</td>
<td>63164</td>
<td>13.1</td>
<td>Hypothetical protein in σE operon</td>
</tr>
<tr>
<td>NMH_2155</td>
<td></td>
<td>3071</td>
<td>40218</td>
<td>13.1</td>
<td>Hypothetical protein in σE operon</td>
</tr>
<tr>
<td>NMH_2156</td>
<td>σE</td>
<td>1587</td>
<td>9997</td>
<td>6.3</td>
<td>RNA polymerase σE, σE operon</td>
</tr>
<tr>
<td>NMH_2157</td>
<td>mseR</td>
<td>331</td>
<td>n/a</td>
<td>n/a</td>
<td>Meningococcal σE Regulator, σE operon</td>
</tr>
<tr>
<td>NMH_2253</td>
<td>nirK</td>
<td>638</td>
<td>117</td>
<td>0.2</td>
<td>aniA (MC58)</td>
</tr>
<tr>
<td>NMH_2270</td>
<td>opA50</td>
<td>1098</td>
<td>418</td>
<td>0.4</td>
<td>nspA (MC58)</td>
</tr>
<tr>
<td>NMH_2477</td>
<td>msrA</td>
<td>1726</td>
<td>138968</td>
<td>80.5</td>
<td>methionine sulfoxide reductase msrA/msrB</td>
</tr>
<tr>
<td>σE sRNA</td>
<td></td>
<td>7</td>
<td>114</td>
<td>16.3</td>
<td>σE sRNA</td>
</tr>
</tbody>
</table>
Figure 1. Read coverage visualization of the upregulation of msrA/msrB in H44/76ΔmseR (blue) versus H44/76 wt (red). Transcription on the + strand is visualized on the positive x-axis, - strand transcription on the negative x-axis. Nucleotide position refers to contig9 of H44/76 WGS (AEQZ01000046.1). The relatively high peak of the last 50 nucleotides of msrA/msrB is present in both H44/76 wt and H44/76ΔmseR in both replicates. Expression of msrA/msrB through its ρ-independent terminator can be seen (see text for details), explaining the upregulation of anti-sense transcription of NMH_2475 and NMH_2476 shown in Table 3.

Figure 2. Read coverage visualization of the upregulation of the σE operon in H44/76ΔmseR (blue) versus H44/76 wt (red). Transcription on the + strand is visualized on the positive x-axis, - strand transcription on the negative x-axis. Nucleotide position refers to contig6 of H44/76 WGS (AEQZ01000043.2). The largest absolute transcriptional differences can be found in the first 2 genes of the σE operon, NMH_2149 and NMH_02153.
Differential transcription in intergenic regions

IGRs may contain non-coding RNAs involved in the riboregulation of target mRNAs\textsuperscript{10,11}. Of all IGRs of H44/76, only the IGR between NMH_1566 and NMH_1568 (analogous to NMB1826 and \textit{dnaE} in MC58) showed differential expression with 16 times higher expression in H44/76\textit{ΔmseR} versus H44/76 wt cells (\textbf{Table 1}). A gene encoding a sRNA of 74 nts can be identified in this IGR (\textbf{Figure 3}). Its TSS deduced from read mapping in RNA-seq is at contig4-nt position 59241 (AEQZ01000037.1), which is 11 nts downstream of an \textit{in silico} predicted -10 region of a putative promoter and 3 nts downstream of the predicted TSS (\textbf{Figure 4}). The transcript ends within the T-tract of a \textit{ρ}-independent terminator, with the last reads ending at contig4 nt position 59314 (\textbf{Figure 4}). The gene encoding this sRNA was found to be conserved in all currently sequenced complete and WGS genomes of \textit{N. meningitidis}, \textit{N. gonorrhoeae} and \textit{N. lactamica} strains (query coverage 100\%, identity ≥94\%) and in several but not all commensal \textit{Neisseria} species. Therefore, putative interaction between this sRNA and the 5’ UTR of target mRNAs was assessed using TargetRNA in the related strain MC58\textsuperscript{37}. This way, NMB0205 (\textit{fur}), NMB0792 (\textit{nadC}), NMB0810 (encoding a putative TetR family transcriptional regulator) and NMB1224, NMB1914, NMB2014, and NMB2110, all encoding hypothetical proteins, were identified as putative targets of this novel sRNA.
Figure 3. Read coverage visualization of the upregulation of the σE sRNA in H44/76ΔmseR (blue) versus H44/76 wt (red). Transcription on the + strand is visualized on the positive x-axis, - strand transcription on the negative x-axis. Nucleotide position refers to contig4 of H44/76 WGS (AEQZ01000037.1). The black flag indicates a predicted TSS, the closed circle indicates a predicted ρ-independent terminator. The ORF of hypothetical protein NMH_1566 shows no expression. dnaE (NMH_1568), truncated in this picture with only 150 of 3435 nts shown, is not differentially expressed.

Figure 4. Sequence and predicted secondary structure of the σE sRNA. (A) Mfold prediction of the secondary structure of the σE sRNA. (B) Sequence of the σE sRNA gene including promoter. The boldfaced -35 and -10 regions and transcriptional start site are predicted in silico. The boxed +1 nucleotide is the experimentally validated transcriptional start site of the in vivo present transcripts.
Differential expression between H44/76 wt and H44/76ΔmseR of NMH_0518, NMH_2270, nirK, msrA and NMH_2149 (representing the σE operon) was also assessed by qRT-PCRs on reverse transcribed RNA. The internal control rmpM showed very similar Cp values for both cDNA samples (11.7 ± 1.4 vs 11.9 ± 1.5; p = 0.4573), consistent with RNA-seq results showing less than 6% difference between total normalized reads mapped to rmpM (data not shown). Four genes showed similar direction of differential expression (three significant), confirming the RNA-seq data, although fold-differences show varying order of magnitude in comparison (Supplementary Table 3). Detection of transcription of NMH_2270 and the σE sRNA was insufficient to yield reliable results.

Antisense transcription

Cis-antisense transcripts of genes have been implicated to interact with its mirror image mRNA molecule, regulating transcription, translation or degradation. We investigated differential expression by considering the number of reads per locus in the anti-sense direction of all ORFs and detected 6 genes to have differentially expressed anti-sense transcripts (Table 3). In H44/76 wt cells 771 reads were mapped to the anti-sense strand of msrA/msrB, while in H44/76ΔmseR cells this number was 6.4-fold reduced. In H44/76ΔmseR cells, antisense transcription of both NMH_2475 and NMH_2476 was enhanced 27 and 120-fold, respectively, possibly due to read-through from the strongly transcribed msrA/msrB locus (Figure 1). In addition, an 8.3-fold reduction in antisense transcription of NMH_2156 (part of the σE operon) was observed in H44/76ΔmseR cells. Lastly, antisense transcription of nqrE (NMB0565, Na+-translocating NADH-quinone oxidoreductase subunit E) and σE itself was 4.0 and 5.4-fold increased, respectively.
Deep Sequencing Whole Transcriptome Exploration of the $\sigma^E$ Regulon in *N. meningitidis*

Table 3. Differentially expressed loci in anti-sense direction, ranked by locus number, between *N. meningitidis* H44/76 wild type and *N. meningitidis* H44/76ΔmseR

<table>
<thead>
<tr>
<th>Locus $^\beta$</th>
<th>Gene $^\gamma$</th>
<th>wt</th>
<th>$\Delta$mseR</th>
<th>Fold change</th>
<th>$p$-value $^\delta$</th>
<th>Description $^\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMH_0763</td>
<td>nqrE</td>
<td>30</td>
<td>121</td>
<td>4.0</td>
<td>0.020</td>
<td>NADH:ubiquinone oxidoreductase</td>
</tr>
<tr>
<td>NMH_2154</td>
<td></td>
<td>157</td>
<td>19</td>
<td>0.1</td>
<td>&lt;0.001</td>
<td>Hypothetical protein, $\sigma^E$ operon</td>
</tr>
<tr>
<td>NMH_2156</td>
<td>$\sigma^E$</td>
<td>26</td>
<td>141</td>
<td>5.4</td>
<td>0.001</td>
<td>RNA polymerase $\sigma^E$, $\sigma^E$ operon</td>
</tr>
<tr>
<td>NMH_2475</td>
<td></td>
<td>17</td>
<td>461</td>
<td>27.1</td>
<td>&lt;0.001</td>
<td>Mechanosensitive ion channel family protein</td>
</tr>
<tr>
<td>NMH_2476</td>
<td>msrA/msrB</td>
<td>6</td>
<td>725</td>
<td>120.8</td>
<td>&lt;0.0001</td>
<td>Competence/damage-inducible CinA C-terminal methionine sulfoxide reductase msrA/msrB</td>
</tr>
</tbody>
</table>

$^\alpha$Reads/locus after library normalization by DEseq. $^\beta$Locus number as annotated in H44/76 WGS (AEQZ00000000). $^\gamma$Gene name when annotated in H44/76 WGS (AEQZ00000000). $^\delta$All loci with Benjamini-Hochberg adjusted q-value ≤0.05. $^\epsilon$Where available *N. meningitidis* strain MC58 orthologues are noted.

Discussion
Positive feedback loop of the $\sigma^E$ operon and control of msrA/msrB expression by $\sigma^E$

Previous reports indicated a relatively small $\sigma^E$ operon regulon in pathogenic *Neisseria* when compared to other species with a much wider range of habitats, e.g. *Bacillus subtilis*, where it directs the transcription of many genes that contribute to the formation of mature spores $^{3,6,42}$. Here we used RNA-seq to explore this regulon in *N. meningitidis* in greater depth. We found 4 novel genes of which their expression was regulated by $\sigma^E$. We did not detect any similarities between the promoters of these genes and the promoter consensus of the $\sigma^E$ operon and msrA/msrB described previously, implicating an indirect effect of $\sigma^E$ $^3$. An additional gene, the OpcA superfamily protein NMH_1358, also showed differential expression between H44/76 wt and H44/76ΔmseR cells, but this was
explained by phase variation. Evidently, the $\sigma^E$ operon and $msrA/msrB$ quantitatively show the most dramatic upregulation of sense transcription. The magnitude of differences in expression of the newly identified genes is smaller, which may indicate that the major function of $\sigma^E$ is to provide positive feedback to its own operon and to induce a strong expression of $msrA/msrB$. However, even relatively small transcriptional differences can have a major impact on genome gene expression, e.g. change in transcription factors may propagate among many genes.

Discovery of novel differentially expressed genes

In general, $\sigma^E$ is found to be involved in the regulation of genes encoding proteins residing in the outer membrane or periplasmic space\(^2\). Of the 4 newly identified $\sigma^E$ regulated genes, 3 encode outer membrane proteins or secreted proteins. Of these, only one, NMH_0518 (NMB1721), encoding a protein of the PepSY-associated TM helix family of proteins, was upregulated and its protein structure suggests that it is held at the cell wall or is secreted\(^43\). PepSY domains are present in a diverse family of secreted and cell-wall associated proteins and it has been suggested that they act as protease inhibitors, and as such regulating protease activity in the local environment to the cell, with possible significance for pathogenicity\(^43\). The 2 other $\sigma^E$ regulated protein coding genes are nirK (aniA), and NMH_2270 (nspA); both appeared to be downregulated. NirK is an anaerobically induced outer membrane protein producing nitric oxide through denitrification\(^44\). Its gene is activated by the fumarate and nitrate reduction regulator FNR when iron concentration is low\(^45\). When iron concentration is not limiting than nirK is repressed by the nitrite-sensitive repressor NsrR, replacing FNR at the nirK promoter. NsrR contains a labile Fe-S cluster and when intact it binds to the nirK promoter and represses nirK\(^45,46\). Methionine sulfoxide reductase, ubiquitously present across all organisms and encoded by $msrA/msrB$, might stabilize iron-sulfur clusters as has recently been shown in Saccharomyces cerevisiae\(^4,47\). We suggest that, considering the pervasive similarities of $msrA/msrB$ in both eukaryotes and prokaryotes, high expression of the methionine sulfoxide reductase may favorably affect the stability of the Fe-S complexes of NsrR increasing the relative NsrR-S-Fe concentration shifting the balance towards more repression of nirK.

NMH_2270, encoding the Neisserial surface protein NspA is a vaccine candidate that can bind to human factor H and enhances complement resistance\(^48\). nspA is activated by the Ferric Uptake Regulator protein (Fur) when iron concentration is
not limiting and downregulated when iron concentrations are low. In conditions with sufficient iron, Fur also downregulates itself via a negative feedback loop. The Fur protein contains Fe-S clusters and in the complexed form it binds to its own promoter and represses transcription of itself. The Fur protein contains Fe-S clusters and again with high expression of methionine sulfoxide reductase the Fe-S clusters in Fur may be stabilized. Reduced transcription of fur was detected in H44/76ΔmseR cells albeit not significant, therefore (an) alternative mechanism(s) may explain the downregulation of nspA.

As described by Anders et al., DEseq was designed to improve on previous tools like edgeR, which tends to overestimate differential expression in the lower range of expression, while underestimating differential expression when genes are relatively highly expressed. Nevertheless, examples of the first may still be present, with σE and nqrE, listed in Table 2 as being upregulated in their antisense direction. Both have very low expression in both the H44/76 wt and H44/76ΔmseR cells. Upon closer visual inspection, all reads are scattered within these relatively large loci without any apparent organization. They are therefore either near the detection limit of this experiment or not likely to represent any biological relevance. Examples representing the other side of the quantitative transcriptional spectrum are the genes encoding tRNA-Leu; NMH_1084, NMH_1085 and NMH_2191, which show 2.0 to 3.4-fold upregulation (Supplementary Table 2). Especially, NMH_2191 has a high number of uniquely mapped reads with 224294 reads in H44/76 wt and 771702 reads in H44/76ΔmseR cells. This extremely high number of reads is seen only with tRNAs, 5S RNA and several mRNAs coding for ribosomal proteins. Increased expression of tRNA-Leu genes can be caused by an increased demand for leucine incorporation in leucine-rich proteins. Analysis of the σE operon using Composition Profiler reveals relatively high usage of leucine \(p < 0.001\) compared to the average usage of this amino acid in the total protein coding content of the H44/76 genome. Therefore, artificially created constitutional overexpression of this operon increases the intracellular demand for leucine dramatically, which provides a biological explanation for the observed increase of reads mapped to three tRNA-Leu loci (NMH_2191, NMH_1084 and NMH_1085).

There are three additional clusters of genes which showed differential expression without reaching significance when adjusted according to Benjamini-Hochberg to control for FDR, comprising 1) a restriction modification system, 2) the leukotoxin type I secretion system and 3) the thiamin biosynthesis operon (Supplementary
The increased expression of the restriction modification locus in H44/76ΔmseR cells may be caused by phase variation. The methyltransferase (mod) protein NMH1820 (analogous to NMB1375, pseudogene part 1) contains tetrameric AGCC repeats causing a frameshift that leads to a premature stop codon\(^{55}\). After a stop codon, ribosomes dissociate from mRNAs, making that part of mRNA less protected against degradation and lowering the number of reads of the mRNA posterior of the stop codon in H44/76 wt\(^{56}\). However, Sanger re-sequencing of this repeat region in H44/76 and H44/76ΔmseR using genomic DNA extracted from the samples used for RNA-seq did not show differences in the number of tetrameric repeats. Different mechanisms for differential expression, e.g. transcriptional slippage might be involved, but the lack of genomic differences indicate in favor of the current finding of non-significant differential expression\(^{57}\).

It remains unclear why the leukotoxin type I secretion system is 2-fold downregulated and genes encoding proteins involved in thiamin biosynthesis were 2.2 to 2.8-fold upregulated in H44/76ΔmseR cells. Of note, the thiamin biosynthesis pathway also contains enzymes with iron-sulfur clusters which might suggest a relation with the strong upregulation of msmA/msmB\(^{58}\). Interestingly, oxidative stress leads to auxotrophic requirements for thiamin in *Salmonella enterica* and *Erwinia chrysanthemi*\(^{59,60}\). In *E. chrysanthemi*, thiamin auxothrophy can also results from depletion of the Isc system, which is known to assist Fe-S cluster biogenesis, resulting in drastic alterations in virulence\(^{60}\). In meningococci, increased expression of MsrA/MsrB can prevent Fe-S cluster disintegration, thus supporting increased thiamin biosynthesis.

### Differential expression of non-coding RNA transcripts

Only one IGR showed differential expression between H44/76 wt and H44/76ΔmseR cells. The expression of a sRNA transcript located in this IGR between NMH_1566 and NMH_1568 is upregulated in the H44/76ΔmseR mutant (Figure 3). This \(\sigma^\varepsilon\)sRNA contains no open reading frame and a BLAST search shows no similarity on the nucleotide level with species other than *Neisseria*. Putative mRNA targets of this \(\sigma^\varepsilon\)sRNA are, in addition to several genes coding for hypothetical proteins, Fur, a NadC transporter family protein and a TetR family transcriptional regulator. By regulating the expression of *fur* it may also effect the transcription of *nspA*. Moreover, regulation of *fur* and the TetR family transcriptional regulator would extend the impact of \(\sigma^\varepsilon\) on the global protein expression pattern. Interaction of many sRNAs and their mRNA targets has been
shown to be enhanced by the bacterial sm-like protein Hfq by facilitating base
pairing and stabilizing the double stranded RNA complex. Two proteomic
analyses of the influence of deletion of \textit{hfq} on the protein expression patterns of
\textit{N. meningitidis} strains MC58 and H44/76 have shown an attenuated phenotype
with an extensive profile of differentially expressed proteins. None of the
putative mRNA targets of this $\sigma^E$sRNA have been shown to be differentially
expressed in these studies. This might imply that the targets identified \textit{in silico}
are not relevant \textit{in vivo} or the interaction between $\sigma^E$sRNA and its targets is Hfq
independent. Since the sRNA expression is increased in H44/76ΔmseR cells one
might have anticipated differential expression of its target genes. Nevertheless,
one of the putative target genes of the sRNA show differential expression
between H44/76 wt and H44/76ΔmseR cells. When sRNAs interact with target
mRNAs, the latter are often destabilized due to the activity of the RNase E and
RNase III dependent degradosome. As a result, a target mRNA is degraded to
smaller fragments which can be observed as a less intense or fragmented signal
in northern blots or with a reduced or absent signal in RT-PCRs. RNase III is
also used to process total RNA extracted from the bacterial cultures during SOLiD
library preparation. In addition, with the SOLiD sequencing protocol employed
here, reads between 25 and 50 nts were generated from all RNA molecules with
a native or processed size of 50 nts or larger. Together this makes RNA-seq by
SOLiD sequencing with the currently used protocol less suitable for detection of
sRNA-mRNA interactions. Further study is underway to assess the function of this
interesting non-coding sRNA.

\textit{\rho}-independent terminator read-through

NMH_2475 and NMH_2476 show increased antisense transcription in
H44/76ΔmseR cells. This can be explained by read-through past the \textit{\rho}-
independent terminator of the highly expressed \textit{msrA/msrB} locus located
diametrically oriented just downstream of these genes. This is visualized by small
peaks of the blue H44/76ΔmseR line on the negative x-axis (representing negative
strand expression) at the NMH_2475 and NMH_2476 loci located on the positive
strand, while the red H44/76 wt line shows no antisense expression (\textbf{Figure 1}).
\textit{DEseq} analysis shows that sense expression of NMH_2475 and NMH_2476 is not
significantly altered; therefore, the potential stereometrical obstruction of
opposing RNA polymerases does not seem to influence the expression of these
genes. Similar inefficient termination of transcription has been observed in 454
pyrosequencing RNA-seq experiments in *Helicobacter pylori*, confirmed by northern blot probe hybridization with larger RNA molecules than expected\textsuperscript{68}.

**Conclusion**

This report provides an in-depth analysis of the concise but complete $\sigma^E$ factor regulon in the Beta-proteobacterium *Neisseria meningitidis*. This relatively small regulon implies a strictly defined response suitable for the relatively stable niche, the human throat, where *N. meningitidis* resides.
References


Deep Sequencing Whole Transcriptome Exploration of the $\sigma^E$ Regulon in *N. meningitidis*


Deep Sequencing Whole Transcriptome Exploration of the $\sigma^E$ Regulon in *N. meningitidis*


Supplementary data

**Supplementary Table 1. List of primers used in qRT-PCR experiments.**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>RH_NMH_2477_FW1</td>
<td>agcctatttccaacgcacatcg</td>
</tr>
<tr>
<td>RH_NMH_2477_RV2</td>
<td>gtcgtattctgtggctgaagg</td>
</tr>
<tr>
<td>RH_NMH_2270_FW1</td>
<td>agcacttgccacactgattg</td>
</tr>
<tr>
<td>RH_NMH_2270_RV2</td>
<td>tgacagttgtgactttgcccc</td>
</tr>
<tr>
<td>RH_NMH_2253_FW3</td>
<td>cttctttctacgttcgcacas</td>
</tr>
<tr>
<td>RH_NMH_2253_RV4</td>
<td>ttttccgagatgcagaggctg</td>
</tr>
<tr>
<td>RH_NMH_0518_FW3</td>
<td>gcaacggtagaaacctctgtc</td>
</tr>
<tr>
<td>RH_NMH_0518_RV4</td>
<td>ccgtcattcaataacctcgc</td>
</tr>
<tr>
<td>RH_NMH_2149_FW1</td>
<td>gctggttggctgctgttgc</td>
</tr>
<tr>
<td>RH_NMH_2149_RV2</td>
<td>ttttcagcagctgatccagc</td>
</tr>
<tr>
<td>CT-class4-1</td>
<td>caaacagctgaaattaagcgc</td>
</tr>
<tr>
<td>CT-class4-2</td>
<td>gtgatgattgtgtgccgccc</td>
</tr>
</tbody>
</table>
Supplementary Table 2. List of differentially expressed loci with uncorrected p-values ≤ 0.01. \(^a\) Locus number as annotated in H44/76 WGS (AEQZ00000000), \(^b\) Reads/locus after normalization by DEseq, \(^c\) Unadjusted p-value ≤ 0.01, \(^d\) BH adjusted q-value.

<table>
<thead>
<tr>
<th>Locus</th>
<th>wt (^a)</th>
<th>ΔmseR (^b)</th>
<th>Fold Change</th>
<th>p-value (^c)</th>
<th>p adj-value (^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMH_2477 - msrA/msrB</td>
<td>1726</td>
<td>138968</td>
<td>80.5</td>
<td>2.22E-28</td>
<td>0.000</td>
</tr>
<tr>
<td>NMH_2157 + Meningococcal sigma E regulator, mseR</td>
<td>331</td>
<td>26</td>
<td>0.1</td>
<td>8.23E-15</td>
<td>0.000</td>
</tr>
<tr>
<td>NMH_2149 + doxX family protein</td>
<td>5216</td>
<td>187292</td>
<td>35.9</td>
<td>1.13E-13</td>
<td>0.000</td>
</tr>
<tr>
<td>NMH_2155 + hypothetical protein</td>
<td>3071</td>
<td>40218</td>
<td>13.1</td>
<td>1.40E-13</td>
<td>0.000</td>
</tr>
<tr>
<td>NMH_2154 + conserved hypothetical protein</td>
<td>4807</td>
<td>63164</td>
<td>13.1</td>
<td>6.24E-12</td>
<td>0.000</td>
</tr>
<tr>
<td>NMH_2156 + RNA polymerase sigma factor, sigma-70 family</td>
<td>1587</td>
<td>9997</td>
<td>6.3</td>
<td>3.70E-11</td>
<td>0.000</td>
</tr>
<tr>
<td>NMH_2253 + nirK</td>
<td>638</td>
<td>177</td>
<td>0.3</td>
<td>7.29E-07</td>
<td>0.000</td>
</tr>
<tr>
<td>NMH_0518 - pepSY-associated TM helix family protein</td>
<td>1998</td>
<td>418</td>
<td>0.4</td>
<td>8.88E-05</td>
<td>0.027</td>
</tr>
<tr>
<td>NMH_1358 - outer membrane protein OpcA family protein</td>
<td>3740</td>
<td>1403</td>
<td>0.4</td>
<td>6.45E-05</td>
<td>0.021</td>
</tr>
<tr>
<td>NMH_2270 - opacity protein opA50 (opA30) (V0) (Fragment)</td>
<td>1098</td>
<td>418</td>
<td>0.4</td>
<td>8.88E-05</td>
<td>0.027</td>
</tr>
<tr>
<td>NMH_1821 + DNA methylase family protein</td>
<td>1288</td>
<td>3148</td>
<td>2.4</td>
<td>2.55E-04</td>
<td>0.059</td>
</tr>
<tr>
<td>NMH_1822 + type III restriction enzyme, res subunit</td>
<td>1250</td>
<td>2878</td>
<td>2.3</td>
<td>5.91E-04</td>
<td>0.092</td>
</tr>
<tr>
<td>NMH_1823 + H1054 homolog</td>
<td>211</td>
<td>522</td>
<td>2.5</td>
<td>4.30E-04</td>
<td>0.076</td>
</tr>
<tr>
<td>NMH_1084 - tRNA-Leu</td>
<td>17345</td>
<td>47375</td>
<td>2.7</td>
<td>6.37E-04</td>
<td>0.148</td>
</tr>
<tr>
<td>NMH_2191 - tRNA-Leu</td>
<td>224294</td>
<td>771702</td>
<td>3.4</td>
<td>2.77E-03</td>
<td>0.429</td>
</tr>
<tr>
<td>NMH_1085 - tRNA-Leu</td>
<td>6845</td>
<td>13825</td>
<td>2.0</td>
<td>7.04E-03</td>
<td>0.613</td>
</tr>
<tr>
<td>NMH_2073 + permease for cytosine/purines</td>
<td>882</td>
<td>1916</td>
<td>2.2</td>
<td>1.18E-03</td>
<td>0.156</td>
</tr>
<tr>
<td>NMH_2074 + glycine oxidase ThiO</td>
<td>349</td>
<td>849</td>
<td>2.4</td>
<td>1.03E-04</td>
<td>0.064</td>
</tr>
<tr>
<td>NMH_2076 + hypothetical protein</td>
<td>217</td>
<td>475</td>
<td>2.2</td>
<td>2.35E-03</td>
<td>0.235</td>
</tr>
<tr>
<td>NMH_2078 + Thiazole synthetase ThiG</td>
<td>391</td>
<td>843</td>
<td>2.2</td>
<td>1.67E-03</td>
<td>0.196</td>
</tr>
<tr>
<td>NMH_0534 - outer membrane efflux family protein</td>
<td>1687</td>
<td>895</td>
<td>0.5</td>
<td>8.42E-03</td>
<td>0.659</td>
</tr>
<tr>
<td>NMH_0535 - leukotoxin secretion protein D</td>
<td>3289</td>
<td>1628</td>
<td>0.5</td>
<td>3.78E-03</td>
<td>0.483</td>
</tr>
<tr>
<td>NMH_1851 - leukotoxin translocation ATP-binding protein lktB</td>
<td>5022</td>
<td>2556</td>
<td>0.5</td>
<td>6.12E-03</td>
<td>0.568</td>
</tr>
<tr>
<td>NMH_1441 + hypothetical protein</td>
<td>4014</td>
<td>9306</td>
<td>2.3</td>
<td>1.80E-03</td>
<td>0.204</td>
</tr>
<tr>
<td>NMH_0783 - hypothetical protein</td>
<td>8</td>
<td>42</td>
<td>5.3</td>
<td>1.63E-03</td>
<td>0.309</td>
</tr>
<tr>
<td>NMH_2254 + conserved hypothetical protein</td>
<td>360</td>
<td>139</td>
<td>0.4</td>
<td>4.29E-04</td>
<td>0.076</td>
</tr>
<tr>
<td>NMH_2159 + lipoprotein</td>
<td>125</td>
<td>50</td>
<td>0.4</td>
<td>5.26E-03</td>
<td>0.433</td>
</tr>
<tr>
<td>NMH_1166 - hypothetical protein</td>
<td>1268</td>
<td>2545</td>
<td>2.0</td>
<td>3.92E-03</td>
<td>0.483</td>
</tr>
<tr>
<td>NMH_0258 - putative membrane protein</td>
<td>78</td>
<td>27</td>
<td>0.3</td>
<td>5.40E-03</td>
<td>0.559</td>
</tr>
<tr>
<td>NMH_1696 - hypothetical protein</td>
<td>397</td>
<td>760</td>
<td>1.9</td>
<td>8.52E-03</td>
<td>0.659</td>
</tr>
</tbody>
</table>

Supplementary Table 3. Results of qRT-PCR experiments.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Normalized Cp</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMH_0518</td>
<td>16.86</td>
</tr>
<tr>
<td>NMH_2149</td>
<td>9.82</td>
</tr>
<tr>
<td>NMH_2477</td>
<td>10.99</td>
</tr>
<tr>
<td>NMH_2253</td>
<td>10.67</td>
</tr>
</tbody>
</table>

\(^a\) Locus number as annotated in H44/76 WGS (AEQZ00000000), \(^b\) Reads/locus after normalization by DEseq, \(^c\) Unadjusted p-value ≤ 0.01, \(^d\) BH adjusted q-value.