Role of inter-species recombination of the \textit{ftsI} gene in the dissemination of altered penicillin-binding-protein-3-mediated resistance in \textit{Haemophilus influenzae} and \textit{Haemophilus haemolyticus}

Elizabeth A. Witherden\textsuperscript{*}, Maria Paula Bajanca-Lavado\textsuperscript{2}, Stephen G. Tristram\textsuperscript{1} and Alexandra Nunes\textsuperscript{2}

\textsuperscript{1}School of Human Life Sciences, University of Tasmania, Launceston, Tasmania, Australia; \textsuperscript{2}Department of Infectious Disease, National Institute of Health, Av. Padre Cruz, 1649-016 Lisbon, Portugal

*Corresponding author. Tel: +61-3-63-245400; Fax: +61-3-63-243658; E-mail: elizabeth.witherden@utas.edu.au

Received 23 September 2013; returned 10 December 2013; revised 15 December 2013; accepted 18 January 2014

**Objectives:** To screen the \textit{ftsI} gene sequences obtained from clinical isolates of non-typeable \textit{Haemophilus influenzae} (NTHi) and \textit{Haemophilus haemolyticus} for the presence of mosaic \textit{ftsI} gene structures, and to evaluate the role of inter-species recombination of the \textit{ftsI} gene in the formation and distribution of resistant \textit{ftsI} genes.

**Methods:** The \textit{ftsI} genes of 100 \textit{Haemophilus} isolates comprising genetically defined \textit{β}-lactamase-negative ampicillin-susceptible (gBLNAS), \textit{β}-lactamase-positive ampicillin-resistant (gBLPAR), \textit{β}-lactamase-negative ampicillin-resistant (gBLNAR) and \textit{β}-lactamase-positive amoxicillin/clavulanate-resistant (gBLPACR) isolates of NTHi (\textit{n}=50) and \textit{H}. \textit{haemolyticus} (\textit{n}=50) were analysed in this study. Both the flanking regions and the full-length \textit{ftsI} gene sequences of all study isolates were screened for mosaic structures using \textit{H}. \textit{influenzae} Rd and \textit{H}. \textit{haemolyticus} ATCC 33390 as reference parental sequences, and bioinformatics methods were used for recombination analysis using SimPlot.

**Results:** Of the 100 clinical isolates analysed 34\% (34/100) harboured mosaic \textit{ftsI} gene structures containing distinct \textit{ftsI} gene fragments similar to both reference parental sequences. The inter-species recombination events were exclusively encountered in the \textit{ftsI} gene of gBLNAR/gBLPACR isolates of both NTHi and \textit{H}. \textit{haemolyticus}, and were always associated with the formation of a mosaic fragment at the 3′ end of the \textit{ftsI} gene. There was no evidence supporting horizontal gene transfer (HGT) involving the entire \textit{ftsI} gene among the clinical isolates in vivo.

**Conclusions:** We provide evidence for the HGT and inter-species recombination of the \textit{ftsI} gene among gBLNAR/gBLPACR isolates of NTHi and \textit{H}. \textit{haemolyticus} in a clinical setting, highlighting the importance of recombination of the \textit{ftsI} gene in the emergence of altered penicillin-binding protein 3 and BLNAR-mediated resistance.

**Keywords:** BLNAR, \textit{β}-lactam resistance, mosaic genes, horizontal gene transfer

**Introduction**

Resistance mediated by altered penicillin-binding proteins (PBPs) has emerged as an important mechanism of \textit{β}-lactam resistance in many bacterial pathogens, including non-typeable \textit{Haemophilus influenzae} (NTHi).\textsuperscript{1} In \textit{β}-lactamase-negative ampicillin-resistant (BLNAR) isolates of NTHi, low-level resistance and reduced susceptibility to \textit{β}-lactam antibiotics is specifically mediated by a reduction in the binding affinity of PBP3 for \textit{β}-lactams.\textsuperscript{2,3}

Until the late 1990s little was known about the molecular evolution and genetic mechanism of altered PBP3-mediated resistance in BLNAR isolates of NTHi, as they remained relatively uncommon in most countries.\textsuperscript{4–7} In 2001, the genetic basis of BLNAR was clarified and attributed to the presence of either an N526K or an R517H substitution in PBP3, encoded by \textit{de novo} point mutations in the \textit{ftsI} gene.\textsuperscript{3} In early epidemiological typing studies, high levels of genetic diversity were found among BLNAR isolates, supporting the notion of their independent evolution.\textsuperscript{5,6–10}

Over the past decade many countries have documented a rapid increase in the prevalence of BLNAR NTHi isolates, particularly among respiratory tract isolates.\textsuperscript{2,11} As a result, genotypically defined BLNAR (gBLNAR) isolates account for 15\%–30\% of all NTHi isolates in Australia, the USA and Europe and an alarming 50\% in Japan.\textsuperscript{6,9,12–15} Recent molecular characterization of gBLNAR isolates has identified the presence of identical \textit{ftsI} gene sequences among genetically diverse NTHi isolates, whilst similar PBP3 substitution profiles among geographically distinct
gBLNAR isolates are frequently reported.\textsuperscript{10,15–17} As a result, it is now postulated that the presence of identical \textit{ftsI} genes in diverse \textit{NTHi} backgrounds may have arisen from recombination events involving the horizontal transfer and homologous recombination of the \textit{ftsI} gene either among \textit{NTHi} isolates or between \textit{NTHi} and other closely related \textit{Haemophilus} species.\textsuperscript{15,17–21}

The ability of \textit{Haemophilus} species to transfer the \textit{ftsI} gene horizontally was first investigated in 2007 by Takahata et al.\textsuperscript{22} In that study they described the presence of mosaic \textit{ftsI} gene structures in clinical isolates of \textit{H. influenzae} and these were shown to arise through inter-species recombination of the \textit{ftsI} gene from the closely related \textit{Haemophilus haemolyticus}.\textsuperscript{22} In addition, intra-species recombination events were demonstrated by \textit{in vitro} transformation experiments in which susceptible \textit{H. influenzae} isolates were successfully transformed to a BLNAR genotype by co-incubation with a clinical \textit{NTHi} gBLNAR isolate.\textsuperscript{22}

Although Takahata et al.\textsuperscript{22} described genetic recombination events involving the \textit{ftsI} genes of both \textit{H. influenzae} and \textit{H. haemolyticus} in 2007, little was known about the molecular mechanisms of \textbeta-lactam resistance in \textit{H. haemolyticus}. Recently, Witherden and Tristram\textsuperscript{14} showed that \textit{H. haemolyticus} isolates harboured the same mechanisms of \textbeta-lactam resistance as \textit{NTHi}, and described mutated \textit{ftsI} genes in gBLNAR isolates of \textit{H. haemolyticus} that were similar to those currently recognized in gBLNAR isolates of \textit{NTHi}.\textsuperscript{14} Furthermore, the molecular characterization of \textit{ftsI} gene sequences among resistant and susceptible \textit{H. haemolyticus} isolates demonstrated that some of the commonly reported BLNAR-associated substitutions in \textit{NTHi} (at positions 350, 377, 547 and 569 of PBP3) appear to form part of the baseline amino acid sequence in ampicillin-susceptible isolates of \textit{H. haemolyticus}.\textsuperscript{14} This suggests that some of the BLNAR-associated amino acid substitutions in \textit{NTHi} might have originated from recombination with \textit{H. haemolyticus}.\textsuperscript{14}

Since the genetic recombination of the \textit{ftsI} gene among \textit{H. influenzae} and \textit{H. haemolyticus} isolates was first described in 2007, little work has been done to clarify the contribution of this mechanism to the dissemination of gBLNAR-mediated resistance in \textit{NTHi} or other closely related \textit{Haemophilus} species. In the present study we aimed to molecularly characterize the \textit{ftsI} genes of clinical \textit{NTHi} and \textit{H. haemolyticus} isolates and to investigate the existence of mosaic \textit{ftsI} gene structures through inter-species recombination events. Furthermore, we aimed to clarify whether these structures contribute to the dissemination of BLNAR-mediated resistance in isolates of \textit{NTHi} and \textit{H. haemolyticus}.

### Materials and methods

#### Bacterial isolates

A study collection was established and comprised a total of 100 partially characterized \textit{Haemophilus} isolates. Isolates were collected as part of other studies, and included \textit{NTHi} isolates from paediatric patients suffering from recurrent acute otitis media and nasopharyngeal \textit{NTHi} and \textit{H. haemolyticus} isolates from age-matched healthy controls in Perth, Western Australia,\textsuperscript{23} whilst normal flora isolates of \textit{H. haemolyticus}, representing nasopharyngeal colonization, were collected from healthy adults in Tasmania, Australia.\textsuperscript{14} All isolates were previously characterized with respect to ampicillin susceptibility (determined by microbroth dilution), \textbeta-lactamase production (identified by PCR) and \textit{ftsI} genotype and identified as either \textit{H. haemolyticus} or \textit{NTHi} by PCR for marker genes \textit{sodC}, \textit{fucK} and \textit{hpd} (which encode a [Cu,Zn]-superoxide dismutase, a fuculose kinase and protein D, respectively), as validated and described in a previous study.\textsuperscript{15} Basically, isolates were identified as \textit{NTHi} if they were negative for \textit{sodC} but positive for \textit{fucK} and/or \textit{hpd}, while \textit{H. haemolyticus} isolates were identified if they were negative for both \textit{fucK} and \textit{hpd} but positive for \textit{sodC}. The 100 isolates in the study collection were randomly selected to represent four populations with respect to the amino acid residue at position 526 of PBP3: \textbeta-lactamase-negative ampicillin-susceptible (gBLNAS) or \textbeta-lactamase-positive ampicillin-resistant (gBLPAR) \textit{NTHi} isolates (\textit{n} = 25), all with the normal amino acid at position 526 of PBP3 (N526), gBLNAR or \textbeta-lactamase-positive amoxicillin-clavulanic-resistant (gBLPACR) \textit{NTHi} isolates, all with the N526K-positive altered PBP3 (\textit{n} = 25), \textit{H. haemolyticus} gBLNAS or gBLPAR isolates with the normal N526 in PBP3 (\textit{n} = 30) and \textit{H. haemolyticus} gBLNAR or gBLPACR isolates with the N526K-positive altered PBP3 (\textit{n} = 20). The characteristics, origin and respective GenBank accession numbers of the \textit{ftsI} gene sequences of the study isolates are given in Table S1 (available as Supplementary data at JAC Online).

#### Amplification and sequencing of genomic regions

In order to evaluate the existence of putative recombination events involving \textit{ftsI}, the full-length \textit{ftsI} gene (1833 bp) and its upstream and downstream flanking regions were amplified using the primers and conditions of Takahata et al.\textsuperscript{22} Purified PCR products were then subjected to three sequencing reactions using the following primers: \textit{TKA-F}, CTCTGTTA TCCGTTCAGCAG (\textit{\textbullet}317 to \textit{\textbullet}298 bp); \textit{ftsI-FL}, ACGGTATTTGCAGGACTTGG (620 to 639 bp); and \textit{ftsI-ORF}, TCCGTTCAGCAG (1157 to 1175 bp) [bp numbering relative to the \textit{ftsI} open reading frame (ORF)].

Overall, contigs of \textit{\textbullet}2418 bp, comprising the last \textit{\textbullet}210 bp of the upstream \textit{ftsI} gene, the entire \textit{ftsI} ORF (1833 bp) and the first \textit{\textbullet}374 bp of the downstream murE gene, were assembled using CLCbio Main Workbench version 6.6 (JPR Technologies, Queensland, Australia). Each sequence was manually edited for quality rather than using only automated quality editing.

#### Phylogenetic analysis

Analyses of genetic diversity and phylogeny were conducted using Molecular Evolutionary Genetics Analysis software, version 5.2.2 (MEGA; Institute of Molecular Evolutionary Genetics, Pennsylvania State University; http://www.megasoftware.net), as previously described.\textsuperscript{14} Briefly, the contigs for all study isolates were aligned against the corresponding reference sequences of \textit{H. influenzae} Rd and \textit{H. haemolyticus} ATCC 33390 (GenBank accession numbers L42023 and AB267855, respectively) using the CLUSTALW algorithm. Variability was inferred by computing the number of variable sites (sites that contain at least two types of nucleotides or amino acids), the number of parsimony-informative sites (variable sites occurring in at least two sequences) and matrices of pairwise comparisons. Phylogenies were generated using the neighbour-joining method with bootstrapping,\textsuperscript{25} and the evolutionary nucleotide or protein distances were estimated using both nucleotide differences and Kimura two-parameter\textsuperscript{26} or gamma models,\textsuperscript{27} respectively. For all these analyses, the pairwise deletion option was selected to remove all sites containing missing data or alignment gaps from all distance estimations only when necessary in the distance estimation and not prior to analyses.

Considering that \textit{Haemophilus} species are naturally competent for transformation and DNA uptake occurs in a sequence-specific manner,\textsuperscript{28} we screened our contigs for the presence of the \textit{Haemophilus}-specific 9 bp DNA uptake signal sequence (USS) (5′-AAGTGCGGT-3′) using MEGA software.

#### Recombination analysis

A search for potential mosaic structures involving the \textit{ftsI} gene was performed using SimPlot 3.5.1 (http://sray.med.som.jhmi.edu/SCRFoftware/),\textsuperscript{29} as previously described.\textsuperscript{14} In brief, SimPlot was used to study the similarity
of each study isolate sequence (Query) to the putative parental reference sequences *H. influenzae* Rd and *H. haemolyticus* ATCC 33390, using a sliding window size of 200 bp moved across the alignment in a step size of 20 bp. The suitability of *H. influenzae* Rd and *H. haemolyticus* ATCC 33390 as *ftsI* reference controls was assessed through a BLAST search for the ~2418 bp *ftsI* + flanking region sequence (used in this study) from clinical NTHi and *H. haemolyticus* isolates available in GenBank. Analysis of the 13 available NTHi *ftsI* + flanking region sequences showed ~98% similarity to that of *H. influenzae* Rd, whilst the 5 available *H. haemolyticus* *ftsI* + flanking region sequences shared 100% sequence identity with that of type strain *H. haemolyticus* ATCC 33390. Nucleotide pairwise distances were calculated using the neighbour-joining method with the Kimura two-parameter model (gaps excluded; ts/tv of 2.0). A BootScan sliding window analysis was also performed for all study (Query) isolates, where each was compared with sequences from the probable reference parental isolates *H. influenzae* Rd and *H. haemolyticus* ATCC 33390, and an outgroup *ftsI* gene sequence from *Haemophilus parahaemolyticus* ATCC 10014 (GenBank accession number AB267860). For each nucleotide window range, a phylogenetic analysis was performed using the neighbour-joining topology on the basis of pairwise genetic distances (Kimura two-parameter method; gaps strip off; ts/tv of 2.0). Bootstrap confidence levels were determined by 1000 replicates. Significant changes in phylogenetic relationships among the four taxa from window to window resulted in changes in BootScan values that were indicative of putative recombination events. The likelihood that the observed distributions of informative sites favouring specific phylogenetic groupings might occur randomly was assessed using the maximum χ² test. The most likely crossover region occurred where the observed distribution was least likely to occur randomly (maximum χ² value). The null hypothesis for recombination was that informative sites are drawn from a single distribution. P values for the specified crossover sites were determined by Fisher’s exact test using 2 × 2 contingency tables. Bonferroni multiple correction testing was applied to discard false positives. Results with P values of <0.05 were considered statistically significant at the 5% level.

To support the determined recombination regions and to accurately infer the phylogenetic relationships of the putative recombinant isolates, the defined breakpoint (crossover) regions were used to divide the alignments into delimited genomic regions, and phylogenetic trees were inferred by neighbour-joining (Kimura two-parameter) analysis.

Comparative sequence representation figures were generated (Figures 3 and 4) for the entire ORF of the *ftsI* gene by comparing the *ftsI* gene sequences of NTHi isolates with that of *H. influenzae* Rd reference parental sequence, and *H. haemolyticus* isolates with the *H. haemolyticus* ATCC 33390 reference parental sequence, using the variable-function tool within MEGA 5. Figures were drawn such that solid black lines represent the positions of the nucleotides that differ from the corresponding nucleotides in the respective parental sequence.

**Retrospective analysis**

A retrospective recombination analysis was performed on the three isolates described to have mosaic *ftsI* genes by Takahata et al. In brief, full-length *ftsI* gene sequences for *H. influenzae* isolates MSC07169 and MSC07771 and *H. haemolyticus* isolate MSC07286 were obtained from GenBank (accession numbers AB267865, AB267866 and AB267867, respectively) and analysed as described above.

**Results and discussion**

**Recombination summary**

In the present study, we attempted to evaluate the impact of homologous inter-species recombination on BLNAR-mediated resistance in both NTHi and *H. haemolyticus* isolates. Results from the recombination analysis performed on the 100 NTHi and *H. haemolyticus* isolates are summarized in Table 1 and show an overall incidence of 34% (34/100) for *ftsI* gene recombination. More specifically, 36% (18/50) of NTHi and 32% (16/50) of *H. haemolyticus* isolates displayed evidence of recombination and contained mosaic *ftsI* genes with discrete DNA fragments similar to both reference parental sequences, respectively. All of the isolates identified as inter-species recombinants were of the gBLNAR or gBLPACR N526K genotype, whilst no gBLNAS or gBLPAR isolates showed evidence of inter-species recombination events.

Similar to previous findings of Takahata et al., all our study isolates exhibit two copies of the *Haemophilus*-specific USS, both located in proximity to the 3’ end of the *ftsI* gene. One copy was located in the middle of the *ftsI* gene (1006–1014 bp) whilst the other was located 22 bp downstream of the *ftsI* ORF. Considering the role of the USSs in efficient uptake of species-specific DNA in *Haemophilus* species, our findings reinforce the potential for homologous recombination of the *ftsI* gene between *Haemophilus* species. Comparison of *ftsI* gene sequences from type strains of *H. influenzae* Rd and *H. haemolyticus* ATCC 33390 (used as putative parental sequences) revealed a high degree of homology between the *ftsI* gene and PBP3 protein sequences of the two species. The overall divergence in *ftsI* gene sequence of the two species was 12.7% (233/1833 bp), whilst the *ftsI* sequences of the study isolates diverged by 0.33%–8.35% over the gene’s 1833 bp length when compared with their respective reference isolates. According to PBP3 genotype classifications, the gBLNAS/gBLPAR isolates showed lower levels of *ftsI* gene sequence diversity than their gBLNAR/gBLPACR counterparts, in both species populations, respectively (NTHi, gBLNAS/gBLPAR 0.33%–3.82%; gBLNAR/gBLPACR 0.49%–6.6%; *H. haemolyticus* gBLNAS/gBLPAR 2.40%–3.93%; gBLNAR/gBLPACR 3.11%–8.35%). The highest level of sequence diversity was evident among the recombination isolates, with the majority of sequence heterogeneity encountered towards the 3’ end of the *ftsI* gene in these isolates, and this correlates with the location of the USSs in this region. Similarly, other authors have reported an increase in sequence heterogeneity at the 3’ end of the *ftsI* gene among some clinical gBLNAS isolates.

On the basis of SimPlot similarity plots and informative site analysis, the exact size and position of the mosaic *ftsI* fragment could be calculated. As a result, isolates harbouring recombinant *ftsI* genes could be classified into one of eight distinct recombination profiles based on the size and position of the respective mosaic gene fragment (Table 1). Figures 1 and 2 show representative examples of the SimPlot and BootScan analyses performed on all isolates, and show the exact positions of the breakpoint regions (with P values) supporting the recombination events for an NTHi recombinant isolate (Figure 1) and *H. haemolyticus* recombinant isolate (Figure 2). The number of informative sites shared by the recombinant sequences and the putative parental isolates that support the position of the recombination crossover regions are also shown. It is important to note that the choice of reference control sequences poses some methodological constraints. The choice of reference sequences influences the ability to identify recombinant sequence patterns, and this is most evident in cases where the query sequences are not highly similar to either of the reference sequences. As a result, we acknowledge that the re-analysis of this dataset using alternative reference sequences may alter the position and size of the recombinant fragments identified. Furthermore, this study takes the view that
regions identified within the ORF of the ftsI gene of 100 clinical isolates of NTHi and H. haemolyticus

<table>
<thead>
<tr>
<th>Species/recombinant groupa</th>
<th>Number of isolates</th>
<th>Crossover regions within ftsIb</th>
<th>Recombinant fragmentd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>position (bp)c</td>
<td>P value</td>
</tr>
<tr>
<td>H. influenzae (n = 50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-recombinant</td>
<td>32</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>recombinant group 1</td>
<td>7</td>
<td>2</td>
<td>1189–1265 (L); 1594–1619 (R)</td>
</tr>
<tr>
<td>recombinant group 2</td>
<td>3</td>
<td>1</td>
<td>1465–1505</td>
</tr>
<tr>
<td>recombinant group 3</td>
<td>7</td>
<td>1</td>
<td>1067–1103</td>
</tr>
<tr>
<td>recombinant group 4</td>
<td>1</td>
<td>2</td>
<td>519–557 (L); 1321–1346 (R)</td>
</tr>
<tr>
<td>H. haemolyticus (n = 50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-recombinant</td>
<td>34</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>recombinant group 5</td>
<td>4</td>
<td>2</td>
<td>1321–1346 (L); 1465–1538 (R)</td>
</tr>
<tr>
<td>recombinant group 6</td>
<td>9</td>
<td>1</td>
<td>1321–1346</td>
</tr>
<tr>
<td>recombinant group 7</td>
<td>2</td>
<td>2</td>
<td>712–746 (L); 1067–1103 (R)</td>
</tr>
<tr>
<td>recombinant group 8</td>
<td>1</td>
<td>1</td>
<td>559–587</td>
</tr>
</tbody>
</table>

Isolates were assigned to a recombinant group on the basis of the position and size of the recombinant fragment identified on SimPlot/BootScan analysis.

Data for the crossover regions are calculated from the SimPlot and BootScan analysis using the ftsI+flanking regions contig; n, number of crossover regions identified within the ORF of ftsI; position of the crossover region is relative to the bp numbering of the ftsI ORF.

(L) denotes the first crossover site and (R) denotes the second crossover site.

Size, size of recombinant fragment (bp); position in ftsI, position of the recombinant fragment relative to the ORF of the ftsI gene in bp; position in PBP3, deduced position of the recombinant fragment in the encoded PBP3 protein; percentage divergence, the degree by which the nucleotide sequence of the recombinant fragment diverges from the reference sequence and is calculated by dividing the total number of variable nucleotide sites in the recombinant fragment by the length of the recombinant fragment.

Isolated examples used as representative examples of each recombination profile group. These example isolates were used for the recombination fragment calculations.

as the parental origin of the clinical isolates is unknown, use of the ftsI gene sequences from type strains of both species (H. influenzae Rd and H. haemolyticus ATCC 33390) are the most appropriate for inferring putative parental origins, because these isolates are routinely used to phylogenetically separate the two species.

For all the recombinant isolates identified in this study, the recombinant fragment was most frequently located towards the 3′ end of the ftsI gene, and always (with the exception of group 4 recombinants) surrounded the region of the ftsI gene that encodes the most common marker of BLNAR-mediated resistance, N526K (1576–1578 bp of ftsI). There was no evidence to support the horizontal transfer of the entire ftsI gene between the two species.

Schematic summaries representing the nucleotide divergence among ftsI gene sequences of recombinant and non-recombinant NTHi and H. haemolyticus study isolates are shown in Figures 3 and 4, respectively.

Recombination of the ftsI gene in H. influenzae isolates

Both group 1 and group 4 recombinants appear to have undergone two putative crossover events within the ftsI gene that resulted in a small internal ftsI gene fragment originating from H. haemolyticus, although the exact position and size of the fragment differed between the two groups (Table 1 and Figure 3). For group 1 recombinants, this fragment correlated with amino acids 422–531 of PBP3 and contained the key substitutions N526K, G490E and A530S. Interestingly, the D350N substitution was also present in these group 1 isolates, but it did not appear to be carried within the recombinant fragment. Similarly, the group 4 recombinants possessed an internal 762 bp fragment, consistent with an origin in H. haemolyticus that encoded BLNAR-associated substitutions D350N, M377I and I449V but not the key BLNAR-defining substitution N526K. In both recombinant groups 2 and 3, distinctly different lengths of the 3′ end of the ftsI gene appeared to originate from H. haemolyticus, and carried the important section of the ftsI gene that encodes the BLNAR-defining substitution (Figure 3). This suggests that in these isolates, and those from group 1, the BLNAR-defining N526K substitution may have originally arisen in an H. haemolyticus isolate and was then transferred via horizontal gene transfer (HGT) into a previously susceptible NTHi isolate, inducing BLNAR-mediated resistance in NTHi.
Figure 1. *H. influenzae* group 1 recombinant (using L27 as a model). (a) Number of informative sites shared by the recombinant sequences (light blue) and the putative parental *H. influenzae* and *H. haemolyticus* strains (red and blue, respectively). The outgroup sequence of *H. parahaemolyticus* is shown in grey. Four-member trees consistent with these sites are also shown for each region adjacent to the crossover region. (b) Similarity plot between the recombinant sequence and the putative donor, with a sliding window size of 200 bp and a step size of 20 bp showing the phylogenetic relatedness (% of permuted trees) between these sequences. For (b) and (c), the crossover regions are located between each pair of green vertical lines, and nucleotides at the bottom of each plot correspond to alignment positions (ftsI+flanking regions) (not to chromosomal locations) of the continuous genomic regions analysed. The P values supporting each crossover are also shown. For group 1 isolates, the two crossovers are located within the ftsI gene. (d) Phylogenetic reconstructions for each specific region bounded by the recombination breakpoint region supporting each crossover (1000 bootstrapped trees).

Figure 2. *H. haemolyticus* group 8 recombinant (using L274 as a model). (a) Number of informative sites shared by the recombinant sequences (light blue) and the putative parental *H. influenzae* and *H. haemolyticus* strains (red and blue, respectively). The outgroup sequence of *H. parahaemolyticus* is shown in grey. Four-member trees consistent with these sites are also shown for each region adjacent to the crossover region. (b) Similarity plot between the recombinant sequence and the putative donor, with a sliding window size of 200 bp and a step size of 20 bp showing the phylogenetic relatedness (% of permuted trees) between these sequences. For (b) and (c), the crossover region is located between the pair of green vertical lines, and nucleotides at the bottom of each plot correspond to alignment positions (ftsI+flanking regions) (not to chromosomal locations) of the continuous genomic regions analysed. The P value supporting the crossover is also shown. (d) Phylogenetic reconstructions for each specific region bounded by the recombination breakpoint region supporting each crossover (1000 bootstrapped trees).
<table>
<thead>
<tr>
<th>Reference</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Non-recombinant</th>
</tr>
</thead>
<tbody>
<tr>
<td>H. haemolyticus</td>
<td>L27</td>
<td>L28</td>
<td>L35</td>
<td>L45</td>
<td>L46</td>
</tr>
<tr>
<td>H. influenzae</td>
<td>L30</td>
<td>L31</td>
<td>L32</td>
<td>L33</td>
<td>L34</td>
</tr>
</tbody>
</table>

**Figure 3.** Mosaic ftsI gene structures identified in clinical NTHi isolates. Schematic representation of the divergence in the ftsI gene of NTHi study isolates compared with the ftsI reference sequence of H. influenzae Rd. Solid black lines denote the positions of the nucleotides that differ from the corresponding nucleotides in the reference sequence. Solid grey shading, most similar to H. influenzae Rd sequence; diagonal striped shading, most similar to H. haemolyticus ATCC 33390; solid black shading, crossover location.
Figure 4. MosaicftsIgene structures identified in clinicalH. haemolyticusisolates. Schematic representation of the divergence in theftsIgene ofH. haemolyticusstudy isolates compared with theftsIreference sequence ofH. haemolyticusATCC 33390. Solid black lines denote the positions of the nucleotides that differ from the corresponding nucleotides in the referenceH. haemolyticussequence. Crossover schematics show the proportion of theftsIgene most similar to reference sequences. Solid grey shading, most similar toH. influenzaeRd sequence; diagonal striped shading, most similar toH. haemolyticusATCC 33390; solid black shading, crossover location.
Rd by up to 3.8% over its 1833 bp length. In these isolates the range of PBP3 substitutions varied greatly and, in addition to NS26K \((n = 7)\), included A239E \((n = 5)\), D350N \((n = 2)\), A368T \((n = 3)\), A437S \((n = 1)\), I449V \((n = 1)\), A502T \((n = 5)\), V547I \((n = 3)\), N569S \((n = 4)\) and N586S \((n = 4)\). In these isolates there was no evidence of \(ftsI\) recombination events and this might suggest that the acquisition of the BLNAR-defining NS26K substitution and the additional BLNAR-associated substitutions was through the accumulation of independent point mutations rather than the HGT of resistant \(ftsI\) genes.

**Recombination of the \(ftsI\) gene in \(H. haemolyticus\) isolates**

The group 5 recombinants \((n = 4)\) possessed a small internal 117 bp fragment most similar to \(H. influenzae\), at position 1346–1465 bp of the \(ftsI\) gene (Figure 4). Interestingly, in these isolates the portion of the \(ftsI\) gene that encodes BLNAR-mediated resistance (1576–1578) did not occur within the recombinant fragment originating from \(H. influenzae\). Instead, it appeared within the second putative crossover region, where the exact origin of the sequence cannot be statistically confirmed but appears to be returning to an \(H. haemolyticus\) origin. Most of the \(H. haemolyticus\) recombinants \((n = 9)\) were categorized as having the group 6 recombination profile, which was characterized by the presence of a 486 bp fragment most similar to \(H. influenzae\), at the 3' end of the \(ftsI\) gene (Table 1 and Figure 4). Only two isolates were classified as group 7 recombinants, and both isolates contained an internalized DNA fragment most similar to \(H. influenzae\); however, the exact position of the first putative crossover event differed by ~100 bp between the two isolates. In both isolates, neither recombination fragment contained the region of the \(ftsI\) gene that encodes NS26K. Instead, as in the group 5 recombinants, the NS26K substitution appeared to occur within a crossover region, where the \(ftsI\) gene is probably from the \(H. haemolyticus\) isolate itself. This suggests that recombination in these isolates might not confer BLNAR-mediated resistance, and supports the notion that the NS26K substitution originated in situ in \(H. haemolyticus\).

**Retrospective analysis of the Takahata et al.\(^{22}\) study**

In this retrospective analysis, \(H. influenzae\) strains MSC07169 and MSC07771 were compared with \(H. influenzae\) Rd and \(H. haemolyticus\) strain MSC07286 was compared with \(H. haemolyticus\) ATCC 33390, instead of \(H. influenzae\) Rd as in the original study. The additional bioinformatics analysis performed in this study showed that isolate MSC07169 harboured a 172 bp fragment at the 3' end of the \(ftsI\) gene derived from \(H. haemolyticus\), whilst the presence of two putative crossover sites within the gene of \(H. influenzae\) strain MSC07771 resulted in an internal DNA fragment (size 432 bp; position 419–849) derived from \(H. haemolyticus\) (shown in Figure S1, available as Supplementary data at JAC Online). As a result, the \(H. influenzae\) strain MSC07169 shared a similar recombination profile to that of the group 2 recombinants described in this study, whilst strain MSC07771 shared a similar recombination profile to the group 4 recombinant.

In the \(H. haemolyticus\) strain MSC07286 a large 835 bp fragment most similar to \(H. influenzae\) was identified at the 3' end of the \(ftsI\) gene (position 998–1833 bp), making it most similar to the group 8 recombinants described in this study.

**Conclusions**

In summary, we provide evidence for the HGT and homologous recombination of the \(ftsI\) gene between \(H. influenzae\) and \(H. haemolyticus\) isolates, and show that mosaic \(ftsI\) gene structures form frequently in gBLNAR isolates in vivo. In every case, only partial \(ftsI\) gene fragments were exchanged between \(H. influenzae\) and \(H. haemolyticus\) and these gene fragments frequently encoded important BLNAR-defining and BLNAR-associated amino acid substitutions in PBP3. This highlights the potential role of recombination of the \(ftsI\) gene in the dissemination of gBLNAR in \(Haemophilus\) species. The subsequent analysis of the mosaic \(ftsI\) gene fragments showed that the majority of NTHi recombinant isolates appeared to acquire the BLNAR-defining NS26K substitution from a resistant \(H. haemolyticus\) isolate. This led us to hypothesize that in many gBLNAR NTHi isolates the group 5 recombinant substitution might originate in \(H. haemolyticus\) isolates and be subsequently transferred to NTHi, whilst both isolates are co-localized within the respiratory tract. This is similar to what is currently recognized in \(Streptococcus pneumoniae\), where the resistant PBP2x genes frequently originate in the respiratory commensals \(Streptococcus oralis\) and \(Streptococcus mitis\) before undergoing HGT to the pathogen \(S. pneumoniae\).

As no attempt was made to characterize the clonal nature of the isolates studied here, we cannot exclude the possibility that identical or similar mosaic profiles might represent clonally related isolates. However, NTHi isolates are generally considered heterogeneous, such that clonal spread is very uncommon and gBLNAR isolates are infrequently reported.\(^{5,10}\) Furthermore, as we identified identical mosaic \(ftsI\) gene profiles with identical recombination breakpoints among NTHi isolates from two geographically distinct cohorts (evident among group 6 recombinants), our data strongly suggest that the genetic transfer of resistant \(ftsI\) genes occurs relatively frequently among \(Haemophilus\) isolates in vivo, and that there is some pattern to these recombination events. This conclusion is supported by the fact that the retrospective analysis of the three original mosaic \(ftsI\) gene structures described by Takahata et al.\(^{22}\) which were isolated from clinical samples in Japan, harboured mosaic fragments similar to those of the Australian isolates studied here.

In conclusion, the data of this study confirm that inter-species recombination of the \(ftsI\) gene in \(Haemophilus\) species has a role in NTHi gBLNAR/gBLPACR dissemination, particularly in NTHi. Our data also provide supporting evidence that resistant \(ftsI\) genes associated with the N526K-altered PBP3 genotype might originate in the respiratory commensal \(H. haemolyticus\), and be transferred to NTHi forming mosaic \(ftsI\) genes whilst the species are co-localized within the respiratory tract of an individual.

**Funding**

This project was funded by a grant from the Clifford Craig Medical Research Trust, Launceston, Tasmania. A. N. is a recipient of a postdoctoral fellowship (SFRH/BPD/75295/2010) from the Portuguese Fundação para a Ciência e a Tecnologia.

**Transparency declarations**

None to declare.
### Supplementary data

Table S1 and Figure S1 are available as Supplementary data at JAC Online (http://jac.oxfordjournals.org/).

### References


