Haemophilus influenzae may be untypable by the multilocus sequence typing scheme due to a complete deletion of the fucose operon

The Gram-negative, non-motile coccobacillus Haemophilus influenzae is a commensal of the upper respiratory tract and is involved in various infections including otitis media, epiglottitis, conjunctivitis and meningitis. H. influenzae is also frequently isolated from patients with chronic infections of the upper and lower respiratory tract (Aubrey & Tang, 2003; Kilian, 2005; Turk, 1984).

There are several methods for subdividing and typing of bacterial species. Serotyping divides encapsulated strains of H. influenzae into six serotypes (a–f) reflecting the capsular polysaccharides produced by the bacterium (Kilian, 2005; Pittman, 1931). Most isolates from invasive infections are capsule (Aubrey & Tang, 2003; Harper & Tilse, 1991), but serotyping does not characterize unencapsulated (nontypable) strains, which comprise the majority of isolates cultured in the clinical microbiological laboratory (Harper & Tilse, 1991; Musser et al., 1986). Multilocus sequence typing (MLST) is a sequence-based approach that allows direct comparison of results between laboratories, and MLST schemes have been developed for several bacterial species including H. influenzae (Meats et al., 2003). Fragments of the housekeeping genes adk (adenylate kinase), atpG (ATP synthase F1 subunit gamma), frdB (fumarate reductase iron–sulfur protein), fucK (fuculokinase), mdh (malate dehydrogenase), gpi (glucose-6-phosphate isomerase) and recA (RecA protein) are sequenced and assigned allele numbers. The combination of allele numbers at the seven loci defines the MLST sequence type (ST) and more than 600 H. influenzae STs are presently included in the database (www.mlst.net).

Fucose is utilized as an energy and carbon source in bacteria and fuculokinase is one of four enzymes involved in the fucose pathway in H. influenzae (Chen et al., 1987; Herman, 1971). The genes encoding fuculokinase (fucK) and the other three enzymes in the pathway, 1-fucose permease (fucP), 1-fuculose-phosphate aldolase (fucA) and 1-fucose isomerase (fucI), constitute a regulatory unit. Also part of this operon is fucR, an 1-fucose operon activator, and fucU (fucose operon protein) (Chen et al., 1987; Herman, 1971). The fucose operon spans the region from nt 642737 to 647375 in H. influenzae KW20 Rd (Fleischmann et al., 1995; Peterson et al., 2001) (Fig. 1). Variant strains closely related to H. influenzae are negative for fucK, and the presence of fucK has been used as a marker of H. influenzae (Nørskov-Lauritsen et al., 2009; Nørskov-Lauritsen, 2009). In the present study, we have analysed the fucose operon in two isolates of H. influenzae, PN127 (CCUG 58365) and CCUG 30048, which failed to hybridize with a fucK-specific probe (Nørskov-Lauritsen et al., 2009; Nørskov-Lauritsen, 2009).

Phenotypic profiles were generated with the Neisseria–Haemophilus card for the VITEK 2 (bioMérieux) and supplemented with selected biochemical tests (Kilian, 2005), and biotypes were defined as described by Kilian (2005). PCR amplification of the entire fucose operon (approx. 10 kb) was performed with the Expand 20 kb PLUS PCR System (Roche). MLST analysis was performed as described by Meats et al. (2003). The H. influenzae database containing 627 STs was downloaded on 17 April 2009 from the MLST website (www.mlst.net). Also included in the phylogenetic analysis were the type strains of H. influenzae (HK 389T), Haemophilus aegyptius (HK 367T) and Haemophilus haemolyticus (HK 386T), seven reference strains of H. influenzae (HK 61, HK 1136, HK 1212, HK 2067, HK 2122, CCUG 35214 and KW20 Rd), and 31 strains previously classified as variant strains (Nørskov-Lauritsen et al., 2009). fucK was excluded from the concatenated sequences used for cluster analysis. Phylogenetic analyses were conducted using MEGA version 4.0 (Tamura et al., 2007). The population structure of H. influenzae was divided into phylogenetic groups I and II (Meats et al., 2003; Musser et al., 1988).

H. influenzae strains PN127, CCUG 30048 and KW20 Rd were subjected to PCR with primers flanking the fucose operon. The forward primer H1_0608.1002fw was positioned in locus H1_0608 and the reverse primer H1_0616.1708rv in locus H1_0616 (hepA) (Fig. 1; primer sequences are given in Supplementary Table S1 available in JMM Online). Visualization of the PCR products by gel electrophoresis revealed a product of approximately 10 kb for H. influenzae KW20 Rd, as expected. In contrast, both PN127 and CCUG 30048 gave rise to amplicons of only approximately 2 kb. Sequencing of the fucose operon region in PN127 and CCUG 30048 showed that nucleotides 641308–648266 (numbers referring to KW20 Rd) were missing. This deleted region includes all the genes in the fucose operon plus 140 and 143 nt of the flanking intergenic regions upstream and downstream of the operon, respectively. Instead, a 54 nucleotide region is present at position 641309, between folD and hepA, in these two strains (Fig. 1; GenBank accession nos GQ859140 and GQ859141). The 54 nucleotide region appears to contain no putative open reading frames and to have no homologous sequences in the NCBI database.

The GenBank/EMBL/DDBJ accession numbers for the DNA sequences generated in this study are GQ859140–50.

A table giving the sequences of primers used in this study is available with the online version of this paper.

The GenBank/EMBL/DDBJ accession numbers for the DNA sequences generated in this study are GQ859140–50.

A table giving the sequences of primers used in this study is available with the online version of this paper.
We also performed PCR targeting the individual genes of the fucose operon, *fucP*, *fucA*, *fucU*, *fucK*, *fucI* and *fucR* (primer sequences are given in Supplementary Table S1). Products of the expected size were obtained with all *H. influenzae* reference strains, but not with CCUG 30048 and PN127. Thus, the fucose operon was not translocated to other parts of the genome.

Fig. 1. Graphical representation of the fucose operon sites in selected *H. influenzae* genomes. (a) *H. influenzae* KW20 Rd (GenBank accession no. NC_000907). The fucose operon (white) contains the loci HI_0610–HI_0615. Numbers refer to positions in *H. influenzae* KW20 Rd. (b) PN127 and CCUG 30048. The entire fucose operon is missing (GenBank accession nos GQ859140 and GQ859141, respectively).

Fig. 2. Neighbour-joining tree of 627 STs from the *H. influenzae* MLST database, 7 *H. influenzae* reference strains (HK 61, HK 1136, HK 1212, HK 2067, HK 2122, CCUG 35214 and KW20 Rd), type strains of *H. influenzae* (HK 389T), *H. aegyptius* (HK 367T) and *H. haemolyticus* (HK 386T), 31 variant strains closely related to *H. influenzae*, and PN127 and CCUG 30048 from this study. Phylogenetic group I is indicated in blue, phylogenetic group II in yellow and variant cluster in red. Strains with equivocal allocation to group II are shown in grey.
of the genome by genomic rearrangement in PN127 and CCUG 30048. PCR with primers HI_0608.1002f and HI_0616.1708r on the H. influenzae reference strains in all cases produced products of approximately 10 kb, indicating that the six fuc genes are located sequentially, similar to the arrangement in H. influenzae KW20 Rd.

An essential question is the identity and allocation to species of the two investigated strains. CCUG 30048 was isolated from the nasopharynx of a child and is consistent by phenotype with H. influenzae biotype IV, except for the detectable emission of hydrogen sulphide (Nørskov-Lauritsen et al., 2009). By 16s rRNA gene sequence comparison (GenBank accession no. EU909661), it is related to H. influenzae serotype e strains, which constitute a distant lineage within the borders of the species (Nørskov-Lauritsen et al., 2009). Phylogenetic analysis based on concatenated housekeeping gene fragments revealed an intermediary position related to both phylogenetic group II strains (including serotype e and serotype f strains; yellow in Fig. 2) and variant strains excluded from the species (‘non-haemolytic Haemophilus haemolyticus’; red in Fig. 2). Nine STs from the MLST database clustered with CCUG 30048, but none of these were from reference strains with validated allocation to a phylogenetic group or species. However, three adjacent STs (ST 461–3) represent a distant lineage within the borders of the species. H. influenzae type b (Nørskov-Lauritsen et al., 2009) is identical to that of the full genome sequenced strain Rd, and multilocus analysis allocated PN127 to phylogenetic group I (Fig. 2). The strain clearly belongs to the core of the species as evidenced by both clinical occurrence and phenotypic and genetic characterization.

Our results show that occasional isolates of H. influenzae lack the entire fuc operon including fucK. The MLST scheme developed for the unambiguous characterization of H. influenzae (Meats et al., 2003) is unable to assign a sequence type to such strains. Furthermore, detection of fucK may not be suitable for specific identification of H. influenzae (Abeldaim et al., 2009; Nørskov-Lauritsen et al., 2009).

Winnie Ridderberg, Mette G. Fenger and Niels Nørskov-Lauritsen
Department of Clinical Microbiology, Aarhus University Hospital Skejby, DK-8200 Aarhus N, Denmark

Correspondence:
Niels Nørskov-Lauritsen (nielnoer@rm.dk)


