Organization of the *Haemophilus ducreyi* 35000 chromosome

Marcia M. Hobbs, Mary Jean Leonardi, Franca R. Zaretzky, Ting-Hsien Wang and Thomas H. Kawula

Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599-7290, USA

Author for correspondence: Thomas H. Kawula. Tel: +1 919 966 2637. Fax: +1 919 962 8103. e-mail: Kawula@med.unc.edu

A physical and rudimentary genetic map of the *Haemophilus ducreyi* strain 35000 genome was constructed. Pulsed-field gel electrophoresis was used to separate restriction fragments of *H. ducreyi* DNA digested with *SfiI*, *I-CeuI*, or *SfiI* plus *I-CeuI*. The sizes of the fragments were determined, and the circular chromosome was estimated to be 1757 kbp. The six *I-CeuI* fragments and four *SfiI* fragments were ordered into macrorestriction maps using Southern blot hybridization with random *H. ducreyi* clones as probes. It was shown that both *H. ducreyi* and the distantly related *Haemophilus influenzae* have six *rRNA* operons marked by the locations of the *I-CeuI* sites. However, the two species displayed distinct *I-CeuI* restriction patterns. A second *H. ducreyi* strain, CIP542, displayed an identical *I-CeuI* pattern to that of *H. ducreyi* 35000, but *SfiI* digests of the two strains were distinct. The orientation of the six *rRNA* operons was determined and thirteen identified *H. ducreyi* genes positioned on the map of strain 35000.

**Keywords:** *Haemophilus ducreyi*, chromosome map, *rRNA* organization

INTRODUCTION

*Haemophilus ducreyi* is the Gram-negative etiological agent of the sexually transmitted disease, chancre, and *H. ducreyi* infection is a cofactor for transmission of the human immunodeficiency virus (Nzila et al., 1991; Plummer et al., 1991; Pepin et al., 1992; Le-BaQ et al., 1993). Considerable attention has been focused on identifying potential *H. ducreyi* virulence determinants (Spinola et al., 1990; Campagnari et al., 1991; Alfa, 1992; Castellazzo et al., 1992; Parsons et al., 1992; Spinola et al., 1992, 1993; Hollyer et al., 1994; Palmer et al., 1994; Elkins et al., 1995) and on developing animal models of *H. ducreyi* infection (Purcell et al., 1991; Spinola et al., 1994; Totten et al., 1994; Hobbs et al., 1995). However, relatively little is known about the genomic structure and organization of this organism.

The development of pulsed-field gel electrophoresis (PFGE) techniques has dramatically influenced the field of physical chromosome mapping and the analysis of bacterial genomic organization. Combined with the identification of restriction enzymes with recognition sequences that occur infrequently in bacterial genomes, PFGE has become a reliable and efficient method for estimating genome sizes and constructing macrorestriction maps of bacterial chromosomes (Fonstein & Haselkorn, 1995).

Here, we describe a macrorestriction map of the chromosome of the virulent *H. ducreyi* strain 35000 created using PFGE and Southern blot analysis. The circular chromosome was estimated to be 1752 kbp based on the sizes of the fragments generated from digesting whole chromosomal DNA with *I-CeuI*, *SfiI*, or *I-CeuI* plus *SfiI*. *I-CeuI*, an intron-encoded endonuclease from the chloroplast large rRNA gene of *Chlamydomonas reinhardtii*, recognizes a 26 bp sequence present in the rRNA genes of a variety of bacterial species (Liu et al., 1993). Digestion of *H. ducreyi* 35000 DNA with *I-CeuI* indicated that this species, like the distantly related *Haemophilus influenzae*, has six *rRNA* operons. We have placed thirteen identified *H. ducreyi* genes on the physical map, resulting in a rudimentary genetic map that may provide a basis for future genetic studies with this organism.

METHODS

**Bacterial strains.** *H. ducreyi* strain 35000 (ATCC 33922) was isolated in Winnipeg, Canada in the 1970s (Hammond et al., 1978). This isolate contained no antibiotic resistance plasmids, and a phenotypically cryptic plasmid reported in some *H. ducreyi* strains (Denecer et al., 1982) was not observed. The virulence of this isolate has been demonstrated in several animal models of infection (Purcell et al., 1991; Spinola et al., 1994; Hobbs et al., 1995). *H. ducreyi* strain CIP542 (ATCC 33940) has been designated the type strain for the species (Skerman et al., 1980);
however, it is less virulent than *H. ducreyi* 35000 in some animal models (Hammond et al., 1978; Hobbs et al., 1995) and in *in vitro* adherence and cytotoxicity assays (Alfa, 1992; Lammel et al., 1993; Holler et al., 1994). *H. influenzae* (ATCC 33391) is the type strain for this species (Skerman et al., 1980). *Haemophilus* species were grown on chocolate agar ([Difco GC agar base, 1\% (w/v) IsoVitaleX (Becton Dickinson, Cockeysville, MD, USA), 1\% haemoglobin] or in BHI broth supplemented with 1\% IsoVitaleX and 50 \( \mu \)g haemin ml\(^{-1}\) at 35 \(^{\circ}\)C in a humidified atmosphere with 5\% (v/v) CO\(_2\). Recombinant *Escherichia coli* strains were grown on LB agar ([Sambrook et al., 1989]) at 37 \(^{\circ}\)C. Longer incubation periods or using more than 1 U of the enzyme has been used to determine the number of operons in many diverse bacterial genomes (Liu et al., 1993).

**RESULTS**

**H. ducreyi** genome size estimate

Like other *Haemophilus* species, *H. ducreyi* has a low G + C content of 39\% (Kilian & Biberstein, 1984). Therefore, we screened restriction enzymes with GC-rich recognition sequences to generate chromosomal digests with relatively few fragments. *Ascl* (GGCGCGCC), *SfiI* (GGCCNNNNNGGCC) and *SgfI* (GCCATCGGC) produced 12, 4 and 14 fragments, respectively, while *NorI* (CGCGCCGC) did not cut *H. ducreyi* 35000 DNA. The intron-encoded endonuclease *I-CeuI* (TAACATAACG-GTCCTAAGGTAGCGA) yielded six fragments in chromosomal digests. The *I-CeuI* recognition sequence occurs in bacterial *rrl* genes, and digestion with this enzyme has been used to determine the number of *rrn* operons in many diverse bacterial genomes (Liu et al., 1993).

The *Ascl* and *SgfI* restriction patterns each contained several unresolved doublet or triplet bands (data not shown) and were not analysed further in this study. The fragments generated from *I-CeuI*, *SfiI* or double digests were resolved by PFGE and used to determine the size and to construct a physical map of the *H. ducreyi* 35000 genome. The double digest yielded 10 fragments, as expected from the total number of sites for the two enzymes (see Fig. 2). Fragment sizes are given in Table 1. The size of the chromosome was estimated to be 1757 ± 54 kbp.

We compared *I-CeuI* and *SfiI* digest patterns of *H. ducreyi* 35000 DNA with *I-CeuI*, *SfiI*, or *I-CeuI* plus *SfiI*.

**Table 1. Fragment sizes after digestion of *H. ducreyi* 35000 DNA with *I-CeuI*, *SfiI*, or *I-CeuI* plus *SfiI***

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Size (kbp) after digestion with:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>I-CeuI</em></td>
</tr>
<tr>
<td>1</td>
<td>956 ± 10</td>
</tr>
<tr>
<td>2</td>
<td>256 ± 10</td>
</tr>
<tr>
<td>3</td>
<td>228 ± 7</td>
</tr>
<tr>
<td>4</td>
<td>129 ± 9</td>
</tr>
<tr>
<td>5</td>
<td>111 ± 7</td>
</tr>
<tr>
<td>6</td>
<td>73 ± 7</td>
</tr>
<tr>
<td>7</td>
<td>92 ± 3</td>
</tr>
<tr>
<td>8</td>
<td>66 ± 4</td>
</tr>
<tr>
<td>9</td>
<td>47 ± 4</td>
</tr>
<tr>
<td>10</td>
<td>20 ± 3</td>
</tr>
<tr>
<td>Total (kbp)</td>
<td>1753 ± 50</td>
</tr>
</tbody>
</table>

Values are means ± SD of two–five independent determinations from different gels.
strain 35000 with strain CIP542 and H. influenzae (Fig. 1). I-CeuI digestion liberated six fragments from each isolate; the patterns of the two H. ducreyi strains were identical to each other and different from that of H. influenzae (Fig. 1a). Thus, the H. ducreyi genome, like that of H. influenzae (Lee et al., 1989; Butler & Moxon, 1990), contains six rrr operons; however the two species have distinct I-CeuI fragment patterns. The two H. ducreyi strains displayed different SfiI restriction patterns (Fig. 1b). Both yielded four SfiI fragments, and the sums of the fragment sizes were similar, but the individual fragment sizes for the two H. ducreyi strains were distinct. SfiI did not cut H. influenzae DNA (data not shown).

Construction of the physical and genetic map

PFGE-separated I-CeuI and SfiI fragments of H. ducreyi 35000 DNA were ordered into macrorestriction maps by probing Southern blots with H. ducreyi DNA probes. Initially, we screened random H. ducreyi 35000 clones from a plasmid library in E. coli (the average insert size was 6–10 kbp) for use as mapping probes. Fewer than 25% of these anonymous clones hybridized to more than one fragment in one or both of the digests (data not shown). Thus, repetitive DNA elements were not sufficiently abundant to confound our analysis. Most of the random probes, and all of those used to construct the map, hybridized to a single fragment in each digest. The random probes used to link the fragments are shown in Fig. 3.

The overlaps of the I-CeuI and SfiI fragments were determined by comparing ethidium bromide-stained pulse field gels and the corresponding Southern blots from individual and double digests (Fig. 2). SfiI digests of H. ducreyi 35000 DNA often contained a faint band of approximately 340 kbp. The size is consistent with a partial digestion product containing fragments S3 and S4. Indeed, the probe used for the Southern blot shown in Fig. 2 hybridized strongly to S3 and weakly to the partial product. The recognition sequence for SfiI is degenerate; the site between S3 and S4 may contain particular bases that are less efficiently cut than those present at the other three sites in the chromosome.

As the physical map evolved, we probed Southern blots with a small collection of cloned H. ducreyi genes obtained from other investigators or constructed in our laboratory. We also used an oligonucleotide probe derived from the gyrB gene of Neisseria gonorrhoeae (Dempsey et al., 1991; Stein et al., 1991). Table 2 lists the cloned gene probes and the fragments to which they hybridized. Each overlap shown in the combined physical and genetic map (Fig. 3) corresponds to the size of a fragment in the double digest and was established by Southern hybridization with at least one probe.

Orientation of rrr loci

The positions of the I-CeuI sites correspond to the locations of the six rrr operons in H. ducreyi. The gene order within rrr loci is rrr(16S)-rrnA-rrl(23S)-rrl(5S)-tRNA (Jinks-Robertson & Nomura, 1987). Honeycutt et al. (1993) described a method for determining the orientation of rRNA operons using sequences found upstream (rrs) and downstream (rrl and rrf) of the I-CeuI site as probes in Southern blot of I-CeuI-digested DNA. We designed oligonucleotide primers to amplify a 707 bp product based on the published H. ducreyi rrs sequence (Dewhirst et al., 1992). PCR using the rrs primers resulted in the amplification of an approximately 700 bp product from H. ducreyi 35000 genomic DNA, but yielded no product with E. coli K-12 chromosomal DNA as a template (data not shown). Another pair of primers, designed to amplify the region between the rrl and rrf genes, were based on the sequence in highly conserved portions of the E. coli tRNA genes (Brosius et al., 1981). The rrl and rrf primers generated two products with both H. ducreyi (340 bp and 610 bp) and E. coli (300 bp and 550 bp) templates (data not shown). The relative locations of the PCR products and the I-CeuI site within a rrr locus are shown in Fig. 4(a). Following the strategy of Honeycutt et al. (1993), we used the fragments amplified from H. ducreyi rRNA genes to probe Southern blots of H. ducreyi.
Fig. 2. PFGE separation of I-Ceul (lane 1), I-Ceul plus SfiI (lane 2), and SfiI (lane 3) fragments of *H. ducreyi* 35000 DNA. (a) Ethidium-bromide-stained gel run with a 20 s pulse time. The six I-Ceul fragments (C1–C6) and ten double digest bands (D1–D10) were separated under these conditions. The largest SfiI fragments, S1 and S2, were compressed on this gel (S1/2). The faint band beneath S1R labelled with an asterisk was probably the result of partial digestion; its size is consistent with the sum of the sizes of S3 and S4. (b) Southern blot probed with random clone IA5. The probe hybridized strongly to fragments C6, D8 and S3, and faintly to the band below S1/2.

**Table 2.** Cloned gene probes and the fragments to which they hybridized

| Probe                  | Genotype and phenotype                                                                 | Reference or source                          | Hybridizing fragments
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>I-Ceul</td>
</tr>
<tr>
<td>pLP14*</td>
<td><em>dnaK</em>, heat shock proteins</td>
<td>Parsons et al. (1995)</td>
<td>C3</td>
</tr>
<tr>
<td>pHD24†</td>
<td><em>fpA</em>, monomer subunit for fine tangled pili</td>
<td>Brentjens et al. (1996)</td>
<td>C3</td>
</tr>
<tr>
<td>pUNCH577*</td>
<td>fur, iron-responsive regulatory protein</td>
<td>Biegel et al. (1995)</td>
<td>C2</td>
</tr>
<tr>
<td>pBS*</td>
<td><em>groES/EL</em>, heat shock proteins</td>
<td>Parsons et al. (1992)</td>
<td>C1</td>
</tr>
<tr>
<td>gyrB oligo‡</td>
<td><em>gyrB</em>, gyrase B subunit</td>
<td>Dempsey et al. (1991)</td>
<td>C3</td>
</tr>
<tr>
<td>pHD28§</td>
<td><em>hbp</em>, <em>H. ducreyi</em>-specific outer membrane lipoprotein</td>
<td>Hilke et al. (1994)</td>
<td>C2</td>
</tr>
<tr>
<td>pUNCH579*</td>
<td><em>hgbA</em>, haemoglobin receptor</td>
<td>Elkins et al. (1995)</td>
<td>C5</td>
</tr>
<tr>
<td>pKLP104*</td>
<td><em>bhdB</em>, haemolysin genes</td>
<td>Palmer &amp; Munson (1995)</td>
<td>C1</td>
</tr>
<tr>
<td>pHD40‡</td>
<td><em>ompP</em>, OmpA homologue</td>
<td>Hilke &amp; Spinola (1995)</td>
<td>C3</td>
</tr>
<tr>
<td>pHD18†</td>
<td><em>pal</em>, peptidoglycan-associated lipoprotein</td>
<td>Fortney et al. (1996)</td>
<td>C1</td>
</tr>
<tr>
<td>pROSE25*</td>
<td><em>asd</em>, aspartate semialdehyde dehydrogenase; <em>soxC</em>, [Cu, Zn]-superoxide dismutase and <em>recF</em>, recombination and DNA repair</td>
<td>San Mateo et al. (1995)</td>
<td>C1</td>
</tr>
</tbody>
</table>

* Cloned from *H. ducreyi* strain 35000.
† Cloned from *H. ducreyi* strain 85-023233.
‡ Sequence determined from *N. gonorrhoeae* strain MUG 116 (Stein et al., 1991).

Ceul digests to determine the orientation of each *rrn* operon (Fig. 4b). The *rrs* (16S) probe hybridized to fragments C2–C6, but not to C1. The two differently sized *rrl* (23S) and *rrf* (5S) probes gave identical results, hybridizing to C1–C4 and C6, but not to C5. Thus, the *rrn* operons flanking C5 are divergently transcribed, with only *rrs* genes on this fragment. The origin of chromosomal replication probably lies within C5. Likewise, the *rrn* operons flanking C1 are convergent with downstream *rrl* and *rrf*, but not *rrs* sequences on this fragment. Replication termination probably occurs within C1. The *rrn* loci have been arbitrarily designated A–F moving clockwise from the putative origin.

**DISCUSSION**

The circular chromosome of *H. ducreyi* strain 35000 was estimated to be 1757 kbp. This is roughly one third of the size of the *E. coli* chromosome (Bachmann, 1990) and slightly smaller than the genomes of other *Haemophilus*
species; *H. influenzae* type b is 2100 kbp (Butler & Moxon, 1990), *H. influenzae* Rd is 1830 kbp (Fleischmann et al., 1995) and *H. parainfluenzae* is 2340 kbp (Kauc & Goodgal, 1989). We constructed a macrorestriction map of the *H. ducreyi* chromosome using Southern blot analysis to identify overlapping I-Ceul and SmI fragments. We positioned thirteen identified *H. ducreyi* genes on the map and determined the orientation of the six *rrn* operons. At the low resolution provided by the current map, we can locate genes within a region of overlap between particular I-Ceul and SmI fragments, but the relative order of markers on separate probes within an overlap cannot be determined. The three genes that map to the C1/S1 overlap (asd, sodC and recF) were identified by nucleotide sequence homology to genes in the GenBank database, and are present in the order indicated within the cloned probe pROSE25.

The locations of the *rrn* loci relative to one another are precise within 10 kbp, which was the largest standard deviation of the mean for the I-Ceul fragment sizes. The rRNA loci in *H. ducreyi*, as in *E. coli* and other eubacteria, are located in approximately one half of the map (Fonstein & Haselkorn, 1995), and their orientation is consistent with transcription of the *rrn* operons away from the origin of replication (*oriC*). The orientation of the *rrn* loci suggested that *oriC* was located on fragment C5. In addition, the *gyrB* probe mapped to the 20 kbp C5/S3 overlap. The *gyrB* gene encodes the B subunit of DNA gyrase and is usually located near the origin in bacterial chromosomes (Ogasawara et al., 1985; Musialowski et al., 1994). The oligonucleotide used in this study was derived from the *gyrB* gene of *N. gonorrhoeae* (Stein et al., 1991) in a region that is highly conserved between *N. gonorrhoeae* and *E. coli*. Its position on the map probably represents the *H. ducreyi* *gyrB* gene and supports the location of *oriC* near the *rrnA* locus.

PCR with the *rrf* and *rrf* primers based on *E. coli* sequences resulted in the amplification of two products, differing in
size by roughly 200 bp, with both E. coli and H. ducreyi genomic DNA templates. One of the seven E. coli tRNA operons (rrnD) has two rrf genes at the 3' end with a tRNA^Thr gene between them (Duester & Holmes, 1980). The appearance of two PCR products with the rrf primer in E. coli probably reflects amplification from each of the rrf genes at this locus. We did not determine the identity of the tRNAs within H. ducreyi tRNA operons. However, based on the amplification of two products using the rrf tRNA primer, it seems likely that at least one rrn locus may have a similar structure to the E. coli rrnD operon.

The locations of several groups of genes deserve mention. The heat shock genes, groS1/1L and dnaK/J are located in different positions in H. ducreyi and E. coli. While these loci are relatively near the origin in both organisms, the H. ducreyi genes flank oriC, whereas the E. coli groS1: operon is between oriC and dnaK/J (Bachmann, 1990). Several outer-membrane components have been identified in molecular studies of H. ducreyi pathogenesis (Table 2). The genes for these potential virulence determinants are scattered over the half of the H. ducreyi chromosome in which the rrn operons are located. Finally, the iron-regulated genes, fum and hgb.A are not physically linked on the chromosome.

During the preparation of this report, the complete nucleotide sequence of the genome of H. influenzae Rd was published (Fleischmann et al., 1995). Despite the shared taxonomic genus of H. ducreyi and H. influenzae, the two species are not closely related (Dewhirst et al., 1992). Using rrn tRNA sequences to compare members of the family Pasteurellaceae, Dewhirst et al. (1992) placed H. ducreyi and H. influenzae in distinct phylogenetic clusters. We compared the chromosomal organization of the H. ducreyi genome described here with sequence-derived data for H. influenzae Rd (Fleischmann et al., 1995). Fig. 5 illustrates the differences between the two species with respect to the location and orientation of the six rrn operons and the locations of several genetic loci that have been identified in both organisms. The H. influenzae Rd genome lacks capsular genes, the sodC gene encoding a [Cu, Zn]-superoxide dismutase and the fimbrial gene cluster characteristic of pathogenic H. influenzae strains. Therefore, some of the potential virulence factors identified for H. ducreyi are absent from the comparison in Fig. 5. It will be some time before the complete nucleotide sequence of the H. ducreyi genome is known; meanwhile, the physical and genetic map will be a valuable tool for further genetic studies on H. ducreyi. As new genetic markers are added to the map and the resolution improves, we will gain more insights into the genomic organization and potential mechanisms of gene regulation for this increasingly important human pathogen.

ACKNOWLEDGEMENTS

This work was supported by NIH NIAID U01 AI31496 (T.H.K.) and NIH Institutional Training Grant T32 AI07151 (M.M.H.). We thank Stanley Spinola, Christopher Elkins, Susan Biegel Carson, Lani San Mateo and Linda Parsons for providing cloned H. ducreyi genes prior to publication. We are grateful to Janine Cannon and Ken Bott for the use of their PFGE equipment, and to Jo Ann Dempsey for helpful discussions during this study.

REFERENCES


Spinola, S. M., Griffiths, G. E., Bogdan, J. & Menegus, M. A.


Received 18 December 1995; revised 9 April 1996; accepted 16 April 1996.