doi:10.1111/j.1558-5646.2009.00658.x

EXTENSIVE VARIATION IN NATURAL COMPETENCE IN *HAEMOPHILUS INFLUENZAE*

Heather Maughan[1,2,3] and Rosemary J. Redfield[1]

[1]*Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada*

[2]*E-mail: heather.maughan@utoronto.ca*

Received May 13, 2008
Accepted January 24, 2009

The ability of some bacteria to take up and recombine DNA from the environment is an important evolutionary problem because its function is controversial; although populations may benefit in the long-term from the introduction of new alleles, cells also reap immediate benefits from the contribution of DNA to metabolism. To clarify how selection has acted, we have characterized competence in natural isolates of *H. influenzae* by measuring DNA uptake and transformation. Most of the 34 strains we tested became competent, but the amounts of DNA they took up and recombined varied more than 1000-fold. Differences in recombination were not due to sequence divergence and were only partly explained by differences in the amounts of DNA taken up. One strain was highly competent during log phase growth, unlike the reference strain Rd, but several strains did not develop competence under any of the tested conditions. Analysis of competence genes identified genetic defects in two poorly transformable strains. These results show that strains can differ considerably in the amount of DNA they take up and recombine, indicating that the benefit associated with competence is likely to vary in space and/or time.

KEY WORDS: Competence, DNA uptake, polymorphism, selection—natural, transformation, variation.

Natural competence is a regulated physiological state in which bacterial cells actively transport DNA from their external environment into the cytoplasm. (This differs from artificial competence that uses electrical or chemical treatments to passively permeabilize cellular membranes.) The evolutionary function of this genetically programmed ability remains controversial because DNA taken up by cells has three potential benefits, as a food source, as a template for DNA repair, and as a source of new alleles. First, the cell's pool of nucleotides is enriched by degradation of one DNA strand in the cytoplasm, by uptake of nucleotides that have been released by degradation of the other DNA strand at the cell surface, and by degradation of any chromosomal DNA strands displaced by homologous recombination. Second, the internalized DNA strand may be used as a template for recombinational repair of a homologous strand containing otherwise-irreparable DNA damage. Third, the genotype of the cell may change if recombination introduces a different allele ("transformation").

[3]Present address: Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada.

The controversy does not concern the existence of these benefits but their relative importance in the evolution of natural competence. We have argued that selection arising from enrichment of nucleotide pools is likely to be the most significant because the DNA abundant in many microbial environments is a concentrated source of nucleotides and because all cells need nucleotides—the ubiquity of uptake systems is evidence of the need to defray the cost of nucleotide synthesis (Nyhan 2005; Vlassov et al. 2007). However, some argue that the ability of recombination to generate adaptive new genotypes is the only significant selective force (e.g., see Kramer et al. 2007), whereas others emphasize DNA repair because of the prevalence of environmental stressors that can create damage requiring recombinational repair (Michod et al. 2008).

Competence can also be costly. Immediate costs include expression and assembly of the DNA uptake machinery and weakening of the cell envelope containing this machinery (Seto and Tomasz 1975). In *H. influenzae*, DNA uptake can cause lysis because incoming DNA induces the SOS DNA repair response which reactivates dormant prophages resident in the chromosome

(Setlow et al. 1973). Although benefits like those proposed to explain the evolution of eukaryotic sexual reproduction are likely to apply to transformation (Kondrashov 1993; Otto and Gerstein 2006), these will be offset by an increased frequency of deleterious mutations in incoming DNA originating from dead cells (Redfield 1988; Redfield et al. 1997). Competence gene expression is under complex regulation in most species, indicating that the balance between these costs and benefits fluctuates on a physiological time scale.

The variation in transformability reported for different isolates of competent species suggests that costs and benefits of competence may also fluctuate on an evolutionary time scale. Seventy-seven of 120 *Pseudomonas stutzeri* soil isolates were not transformable in the laboratory, and the transformation frequencies of the 43 isolates that were ranged widely, from 10^{-9} to 10^{-4} (Sikorski et al. 2002). Similarly, only 18 of 31 serotype b isolates of *H. influenzae* were transformable, at frequencies that ranged from 10^{-6} to 10^{-3} (Rowji et al. 1989). Comparable variation has also been found in *H. parainfluenzae*, *Actinobacillus actinomycetemcomitans*, *Bacillus subtilis*, *Ralstonia solanacearum*, and *Streptococcus pneumoniae* (Zawadzki et al. 1995; Ramirez et al. 1997; Gromkova et al. 1998; Fujise et al. 2004; Coupat et al. 2008).

The above studies measured only the genetic consequences of DNA uptake (transformation), and so may have only limited value for addressing the evolution of competence. Transformation is the usual assay for competence because acquisition of selectable markers is easy to detect in laboratory cultures. However, cells may be capable of competence and yet fail to transform for many different reasons. First, the tested conditions may not have induced competence. Competence is regulated in very different ways in different species, and strain-specific variation has also been reported (Pozzi et al. 1996; Solomon and Grossman 1996; Martin et al. 2006). Second, the DNA provided may have been inappropriate; some bacteria do not efficiently take up circular plasmid DNA or DNA lacking specific uptake sequences (Scocca et al. 1974; Stuy and Walter 1986). Third, DNA may have been taken up but not recombined into the chromosome because of cytoplasmic degradation or insufficient sequence similarity (Albritton et al. 1984; Majewski and Cohan 1999) (Palchevskiy and Finkel 2006). The impacts of the above factors can only be determined by direct assays of DNA uptake, which we have now completed for 34 strains of *H. influenzae*.

H. influenzae, the model system for studies of competence in the gamma-proteobacteria, is an obligate commensal of the human upper respiratory tract, where it routinely experiences DNA concentrations of ~300 μg/mL of pulmonary secretion (Matthews et al. 1963). Carriage of *H. influenzae* is usually asymptomatic, but in infants, the elderly, or people with stressed immune systems, *H. influenzae* can infect the middle ear (otitis media), lungs (bronchitis), blood (sepsis), and brain (meningitis). *H. influenzae* strains are grouped by the serotype of their capsule, a polysaccharide that surrounds the cell and often assists in evasion of the host immune response. In addition to the six capsule types (a through f), many nonencapsulated ("nontypeable") strains are known.

Competence in *H. influenzae* has been extensively studied only in laboratory cultures of strain Rd, a nonencapsulated derivative of a serotype d strain. In this strain, the genes whose products are required for DNA uptake are coregulated by two nutritional signals reflecting availability of preferred sugars and nucleotides (MacFadyen et al. 2001; Redfield et al. 2005; Cameron et al. 2008). As a result of this regulation, cells in exponential growth do not express these genes and thus do not detectably take up DNA or transform (Redfield 1991; Redfield et al. 2005). Moderate levels of competence develop in rich medium when cultures become dense and growth slows, but competence is most strongly induced by transferring exponentially growing cells to a starvation medium. Thus, depending on growth conditions, transformation frequencies in cultures of strain Rd range from lower than the background due to mutation (~10^{-9}) to higher than 10^{-2}. To determine how extensively other strains of *H. influenzae* differ in their competence phenotype, we measured the ability of 34 strains to take up DNA and integrate it into the chromosome under a range of culture conditions.

Materials and Methods

H. *INFLUENZAE* STRAINS AND GROWTH CONDITIONS

The strains used are listed in Table 1; they were obtained from R. Moxon (Oxford University), A.L. Smith (University of Washington), R. Munson (Ohio State University), and G. Ehrlich (Center for Genomic Sciences, Pittsburgh). All strains were cultivated in Bacto brain–heart infusion broth (BHI), 37 g/L, supplemented with a final concentration of 2 μg/mL NAD and 10 μg/mL hemin (sBHI). Strains were incubated at 37°C either as 5 mL liquid sBHI cultures in tubes on a roller-wheel, as larger cultures in flasks on a rotary shaker, or on sBHI agar plates (12 g agar per liter).

COMPETENCE INDUCTION AND TRANSFORMATION

Competence induction with starvation medium

A 5-mL overnight culture, started from a frozen stock, was diluted 1:100 into fresh sBHI (25 mL in a 250-mL flask). Once cells were growing exponentially (OD_{600} 0.1–0.4), they were collected by filtration, washed in MIV competence medium (Herriott et al. 1970; Poje and Redfield 2003) and then transferred to an equal volume of MIV and returned to the shaker for 100 min. Cells were then either immediately incubated with DNA or mixed with glycerol (16% final concentration) and frozen at −80°C for assay at a later date. (Although frozen cells had slightly reduced viability

Table 1. Strains used in this study.

Strain	Serotype	Environment of isolation	Doubling time (min)	Optical density	Colony forming units (per mL)	Source
Rd[1]	d	N	45	2.3	2×10^8	Redfield
R2846[1]	–	OM	50	2.9	7×10^9	A.L. Smith
R2866[1]	–	M	38	2.7	7×10^9	A.L. Smith
86-028NP[1]	–	OM	58	0.1	3×10^8	R. Munson
22.1-21[1]	–	HN	51	2.6	4×10^8	G. Ehrlich
22.4-21[1]	–	HN	43	3.1	1×10^8	G. Ehrlich
R3021[1]	–	HN	44	1.8	3×10^9	G. Ehrlich
PittAA[1]	–	OM	50	2.1	8×10^9	G. Ehrlich
PittBB	–	OM	48	2.0	2×10^9	G. Ehrlich
PittDD	–	OM	42	2.0	6×10^9	G. Ehrlich
PittEE[1]	–	OM	38	0.4	3×10^7	G. Ehrlich
PittGG[1]	–	O	40	1.8	3×10^9	G. Ehrlich
PittHH[1]	–	OM	40	0.9	2×10^9	G. Ehrlich
PittII[1]	–	OM	48	1.7	6×10^7	G. Ehrlich
RM7033	d	P	40	2.3	3×10^8	R. Moxon
RM7271	d	RI	43	1.7	3×10^9	R. Moxon
RM7429	d	unk.	38	2.7	5×10^8	R. Moxon
RM6181	e	B	30	2.9	5×10^9	R. Moxon
RM6169	e	A	35	1.8	3×10^9	R. Moxon
RM6158	e	unk.	35	1.5	5×10^9	R. Moxon
1008	–	OM	50	1.7	4×10^9	R. Moxon
1209	–	OM	35	1.6	5×10^9	R. Moxon
432	–	OM	53	1.8	7×10^9	R. Moxon
477	–	OM	45	2.3	3×10^9	R. Moxon
375	–	OM	43	1.6	6×10^9	R. Moxon
1124	–	OM	45	1.9	7×10^9	R. Moxon
1181	–	OM	68	1.9	2×10^9	R. Moxon
1247	–	OM	50	2.1	6×10^9	R. Moxon
1207	–	OM	53	1.6	5×10^9	R. Moxon
1233	–	OM	45	2.5	4×10^9	R. Moxon
1241	–	unk.	50	1.7	1×10^8	R. Moxon
176	–	OM	65	3.1	1×10^9	R. Moxon
Eagan	b	M	53	3.4	6×10^9	R. Moxon
1158	–	OM	40	2.0	3×10^9	R. Moxon
1159	–	OM	60	2.6	6×10^9	R. Moxon

[1]Strains with genome sequence data available.

M, meningitis; N, nasopharynx; OM, otitis media; HN, healthy nasopharynx; O, otorrhea; P, pneumonia; RI, respiratory infection; B, bronchitis; A, asthma; Unk., no information available.

Table 2. Growth properties and their relationships with DNA uptake and transformation.

	DNA uptake	Transformation
Doubling time	r^2=0.001, P=0.86	r^2=0.005, P=0.69
Optical density (600 nm)	r^2=0.03, P=0.33	r^2=0.02, P=0.43
Colony forming units per milliliter (CFU/mL)	r^2=0.01, P=0.58	r^2=0.01, P=0.51
Cell clumping	r^2=0.005, P=0.69	r^2=0.008, P=0.61
Cell freezing	r^2=0.01, P=0.13	r^2=0.02, P=0.07

they did not differ significantly from fresh cells in the amount of DNA taken up and recombined [Table 2].) Cells frozen in MIV were quick-thawed at 37°C, pelleted by centrifugation at 13,200 rpm for 60–90 sec, and resuspended in freshly made MIV at room temperature. Each strain was cultured in MIV at least twice (biological replicates), with each culture used for at least two transformation assays on different days (technical replicates).

Transformation assays

The same transformation assay was used for MIV-treated cells and for cells in sBHI cultures (time course assays). Half (0.5) milliliter of cells was incubated with 10–50 ng of linearized plasmid DNA containing an *H. influenzae* Rd chromosomal *gyrA* allele that confers resistance to novobiocin (Redfield 1991). After 15 min uptake was terminated by addition of 10 μg DNase I. In some assays 1 mL of sBHI was added and cells were incubated for an additional 45–50 min; this did not alter transformation frequencies. Appropriately diluted cells were plated on sBHI (to obtain the total number of viable cells) and on sBHI containing 25 μg/mL novobiocin (to obtain the number of viable transformants). Transformation frequencies were calculated by dividing the number of transformants/mL by the number of viable cells/mL. For assays that gave no transformants, the level of detection replaced the transformation frequency.

A single no-DNA control assay was done for MIV prepared cells from each strain, plating approximately 10^6 to 10^8 cells on sBHI containing 25 μg/mL novobiocin; these consistently produced no colonies. When larger numbers of cells from overnight sBHI cultures were plated on sBHI with novobiocin, the only colonies came from strains PittGG (1 colony from 3.3×10^9 CFU) and RM6158 (3 colonies from 5.4×10^9 CFU); the reported transformation frequencies have been corrected for these background mutation frequencies.

Competence and transformation in overnight cultures

For each assay a 5-mL overnight culture, started from a frozen stock, was diluted 1:100 into 5 mL fresh sBHI containing 1 μg chromosomal DNA from strain MAP7 (Poje and Redfield 2003). After ∼24 h of incubation, cells were plated on sBHI agar plates with and without 250 μg/mL streptomycin.

Time course experiments

To follow the timing of competence development during growth in broth and in MIV, transformability was measured at 10 time points in five strains. Cells from overnight cultures were diluted 1:100 into fresh sBHI and the optical density of the culture was followed. Once the cells were in log phase (approximately OD_{600} 0.2) an aliquot was transferred to MIV starvation medium. Samples (0.5 mL) were removed for transformation assays at the time points indicated in Figure 3A.

DNA UPTAKE MEASUREMENTS

A 222 bp polymerase chain reaction (PCR) fragment containing a synthetic uptake signal sequence (Smith et al. 1995) was labeled by one round of DNA synthesis using Klenow polymerase as follows. The 20 μl reaction mixture contained NEB buffer #2, 1.25 nmol each of dCTP, dGTP, and dTTP, 160 pmol each of the forward and reverse primers and 250 μCi (0.083 nmol) α-^{33}P dATP. After 3 min of denaturation at 94°C and cooling on ice, 2 U Klenow were added and the reaction was incubated for 20 min at 37°C. To ensure complete copying of the DNA, 1.2 nmol of unlabeled dATP was then added and the reaction was incubated for an additional hour. The labeled DNA fragment was purified with a Sigma GenElute PCR clean-up column and eluted into 50 μl 10 mM Tris:HCl (pH 8.0) to a final concentration of 5–10 ng/μl and a specific activity of $2–6 \times 10^5$ cpm/ng.

To assay DNA uptake, 8–12 ng of ^{33}P-labeled DNA fragment was added to 0.5 mL competent cells in MIV and incubated, with gentle mixing, at 37°C for 15 min. Cells were washed three times to remove radioactivity that had not been taken up. For each wash, 0.5 mL of cells in MIV were pelleted by centrifugation at 13,200 rpm for 3 min before removal of the supernatant and subsequent resuspension in 0.5 mL MIV. The radioactivity in the final cell pellet and in each of the three supernatants was counted. The fraction of DNA taken up by cells was calculated as the amount of radioactivity taken up (i.e., in the final pellet) divided by the total amount of radioactivity in all tubes (the final pellet plus three supernatants). Log phase cells of Rd, known not to take up DNA, were used as a negative control.

ANALYSIS OF COMPETENCE GENES

Genes whose products take up DNA or are involved with transformation have been previously described for strain Rd (Maughan et al. 2008). We used BLAST and Rd sequences to query the genome sequences of strains PittAA, PittEE, PittGG, PittII, PittHH, R2846, R2866, R3021, 22.1-21, 22.4-21, and 86028NP in GenBank. Homologs were aligned either by eye or using CLUSTALW. Tajima's D and Fu and Li's *D*-statistics and π_A/π_S values were calculated in DnaSP (Rozas et al. 2003).

SEQUENCING COMPETENCE GENES

Polymerase chain reactions were mixed using Qiagen Hotstart kit (Qiagen, Valencia, CA) and contained 25 pmol of each primer, 0.2 mM dNTPs, 1X buffer, 1X Q, 1 U Taq, and ∼100 ng of template DNA. Thermal cycler run parameters were: 94° for 2 min; three cycles of 94° for 1 min, 45 or 48° for 30 sec, and 72° for 1 min; and 41 cycles of 94° for 35 sec, 45 or 48° for 30 sec, and 72° for 45 sec. Amplicons were sequenced by the Nucleic Acid and Protein Service Unit at the University of British Columbia.

Results

H. INFLUENZAE STRAINS AND THEIR GROWTH PROPERTIES

Strains were chosen to include some close relatives and to encompass the genetic diversity in the *H. influenzae* phylogeny, using the multilocus sequence typing (MLST) analysis of Meats et al. as a guide (Meats et al. 2003). Although this analysis did not resolve the deeper phylogenetic relationships, pairwise distances identified close relatives (such as the serotype d strains) and some more divergent strains. The final set of 34 strains was dominated by nontypeable strains isolated from patients with otitis media but also included a variety of capsulated strains isolated from other sites (Table 1). The standard laboratory strain Rd was also used as a control in all assays.

Initial culturing revealed that some strains differed considerably from Rd in their growth properties. Because the physiological state of Rd cultures determines whether they become competent, the first step was to characterize these differences. We measured doubling times in exponential growth and the optical densities and viable counts of overnight cultures, and examined cultures for clumping using phase contrast microscopy. In addition to the between-strain differences described below, many of the clinical strains exhibited substantial culture-to-culture variation in doubling times and overnight densities, even when grown from the same freezer stock, suggesting that these strains were more sensitive than Rd to subtle differences in media or culture conditions. Growth of strains PittEE, 1181, and R3021 was especially variable.

Doubling times were measured for strains growing exponentially in sBHI (log phase); most strains doubled every 30 to 70 min (Table 1). Strains RM6181, RM6169, RM6158, and 1209 grew fastest, averaging less than 40 min, whereas strains 1159, 176, and 1181 averaged greater than 60 min. Rd reproducibly doubled every 45 min. Cell density can be measured two ways: optical density (OD_{600}) measures biomass (as light scattering, primarily by intact living and dead cells), whereas viable counts detect only cells capable of growing to form colonies on solid medium (colony forming units per milliliter of culture [CFU/mL]). For exponentially growing cultures these measures were consistent, indicating that few dead cells were present.

These two measures often gave discordant results with overnight cultures (Table 1). OD_{600} in the overnight cultures ranged from 0.08 to 3.6 (equivalent to about 2×10^8 to 10^{10} intact cells), suggesting strains differed in maximum biomass produced during overnight growth. CFU/mL in the overnight cultures ranged between 10^7 and 10^9; if this 100-fold variation were due to differences in growth, we would expect OD_{600} to strongly predict CFU/mL. However, most of the variation in CFU/mL was not explained by differences in OD_{600} ($r^2 = 0.188$, $P = 0.007$) so differences in CFU/mL may be at least partly due to differences in survival after maximum density was reached. Some cells could also have been viable but not able to form colonies, as has been reported for some otitis media strains of *H. influenzae* (Rayner et al. 1998).

A lower than expected CFU/mL could also be due to clumping. Although cultures had been gently agitated continuously during overnight growth, microscopic examination showed that cells from 20 strains formed clumps containing up to hundreds or thousands of cells. Clumping may be caused by hemagglutinins on the cell surface, an association made in *H. influenzae*'s close relative, *H. parainfluenzae* (Kahn and Gromkova 1981); such surface properties are known to be highly strain-specific. However, this clumping did not explain any variation in CFU/mL ($r^2 = 0.03$, $P = 0.30$), suggesting that the variation in CFU/mL is primarily due to variation in cell survival.

Differences in strains' growth properties are likely to reflect differences in colonization and survival in the human host. However, they did not explain any of the differences in competence described below (Table 2).

DIFFERENCES IN DNA UPTAKE

The ability to bring DNA into the cell is the defining feature of competence. However, because DNA uptake assays are much less sensitive than transformation assays, uptake is best measured when competence has been fully induced. In strain Rd this is accomplished by transferring exponentially growing cells to MIV, a starvation medium containing amino acids but no nucleotide precursors or essential cofactors (Herriott et al. 1970). As this method has been successfully used to induce competence in other strains (e.g., see Swords et al. 2000; Bakaletz et al. 2005), we used it to prepare all strains for DNA uptake assays.

DNA uptake was measured by incubating cells with a radiolabeled DNA fragment containing a strong consensus *H. influenzae* uptake signal sequence (Smith et al. 1995). Because both OD_{600} and CFU/mL differed between strains, we examined the effects of normalizing the amount of uptake by each of these measures. Both normalizations gave quantitatively similar results; normalization by optical density is presented in Figure 1A because this takes into account any cells that were able to take up DNA but not form colonies.

As shown in Figure 1A, the median percentages of DNA taken up by different strains ranged over three orders of magnitude, with some strains taking up almost as much DNA as competent Rd and others having uptake values at or below the detection threshold. This threshold was measured as "uptake" by noncompetent Rd cells; its range and median are indicated by the horizontal dotted lines in Figure 1A. These results show that strains differ extensively in the amount of DNA taken up.

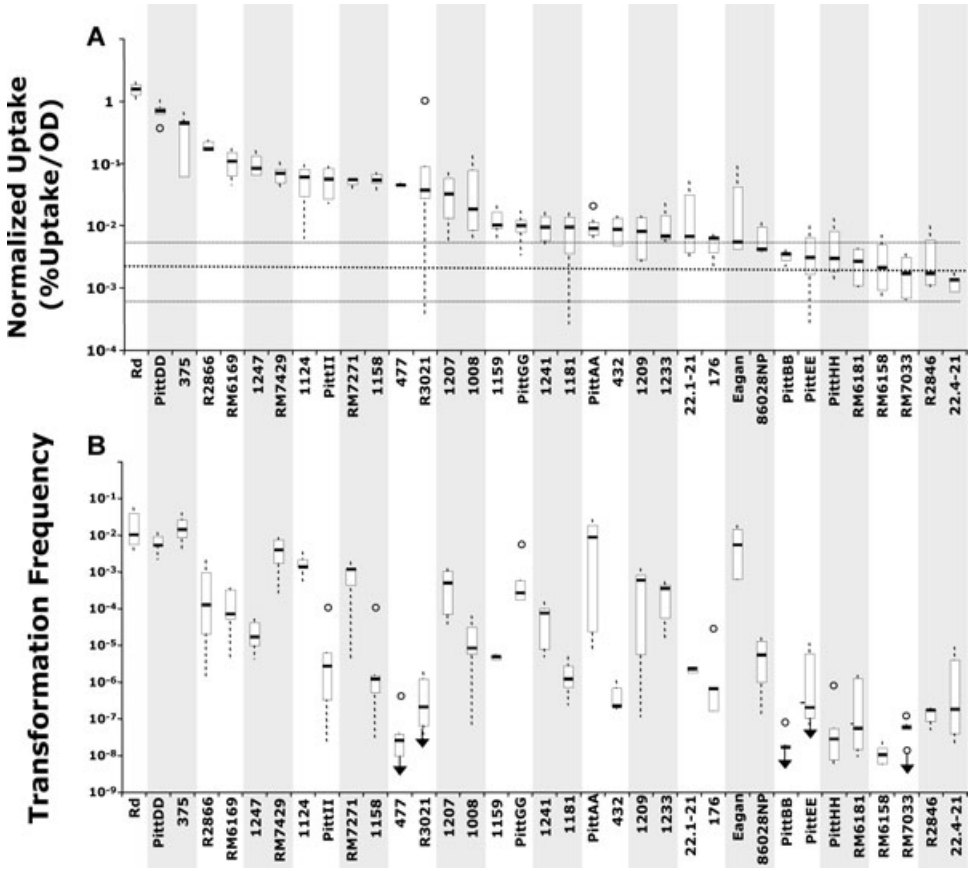

Figure 1. DNA uptake and transformation frequencies in MIV starvation medium. (A) Box plot of normalized DNA uptake for each strain tested. Experiments for each strain were replicated four to 12 times. The dark black line in each box indicates the median, the upper and lower limit of the box represents the upper and lower quartile, respectively, and the range of the data is indicated by dotted lines. Circles outside of the boxes represent outliers, which are defined as points further away from the box than 1.5 times the length of the box. The horizontal dotted lines indicate the median (large dots) and range (small dots) of DNA taken up by noncompetent Rd. (B) Box plot of transformation frequencies in MIV. Experiments for each strain were replicated four to eight times. The structures of the box plots are the same as described for (A). Downward-pointing arrows begin at the detection limit for strains in which transformants were not observed.

DIFFERENCES IN TRANSFORMATION

Although the nutritional benefits of competence are expected to depend only on DNA uptake, any genetic benefits will require that the DNA recombine with the chromosome. To measure transformability, MIV-treated cells were incubated for 15 min with a homologous *gyrA* DNA fragment carrying a mutation conferring novobiocin resistance. The ratio of novobiocin-resistant transformants to total cells is the transformation frequency; for strain Rd typically 1 cell in 100 was transformed.

As seen in Figure 1B, median transformation frequencies ranged much more widely than DNA uptake, from 10^{-8} to 10^{-2}. Given that the amount of DNA taken up imposes an absolute upper limit on the amount of DNA that can be recombined, the correlation between transformation frequency and DNA uptake was surprisingly low (Figure 2; Spearman's $\rho = 0.64$; $P = 0.0001$). Moreover, the correlation was largely due to strains with the highest ($n = 3$) and lowest ($n = 5$) DNA uptake; removal of these from the analysis gave a correlation coefficient of only 0.08. There was no obvious relationship between transformation frequency and serotype or site of isolation, but sample sizes were too small for statistical analysis.

Five strains produced no transformants (477, R3021, PittBB, PittEE, and RM7033), but several analyses indicated that this was unlikely to be an artifact of the assay method. To rule out a marker effect specific to the *gyrA* gene, these strains were retested using Rd-derived chromosomal DNA containing a str^R point mutation marker in the *rps12* gene. No transformants were observed in these experiments (data not shown). Because efficient homologous recombination requires at least 95% identity between donor and recipient DNA (Shen and Huang 1986), divergence between the *gyrA* alleles of the Rd-derived donor sequence (nov^R) and recipients (nov^S) might have prevented recombination. However, the *gyrA* sequences from the 11 sequenced strains all had 97–98% sequence identity to the Rd allele, although the strains had widely

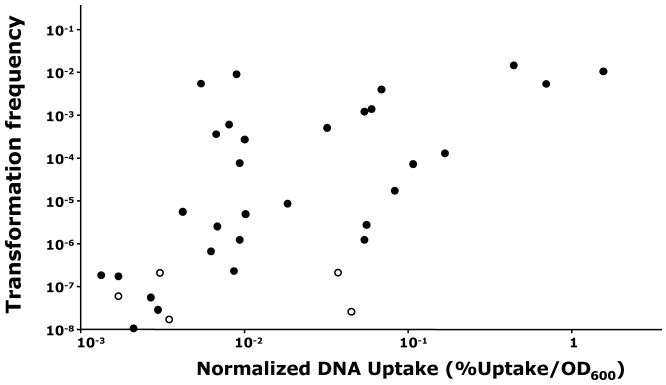

Figure 2. Transformation frequency plotted as a function of DNA uptake. Median transformation frequencies in MIV were plotted as a function of median percent DNA taken up (normalized by OD_{600}). Each point represents one strain. Open circles represent strains in which transformants were never detected in MIV-prepared cells. These strains were excluded from the statistical analysis presented in the text.

differing transformation frequencies. Thus our measures of transformability are unlikely to have been confounded by divergence between donor and recipient alleles. As a further check, we were able to isolate spontaneous rifampicin resistant mutants from three of these strains (477, R3021, and PittEE) and a spontaneous novobiocin resistant mutant from PittBB, allowing us to test whether these strains could be transformed with their own DNA. Strains R3021, 477, and PittBB were still not transformable, but PittEE transformed at a frequency of 6.4×10^{-6}. The different results with self and Rd DNAs may not be significant; the low CFU/mL of this strain limited the sensitivity of transformation assays, and it also gave a small number of transformants with a Rd-derived rifampicin resistance allele.

Some strains produced significant numbers of transformants despite having uptake values at or below the detection limit (Eagan, 86028NP, PittHH, RM6181, RM6158, R2846, and 22.4-21) confirming that the uptake assay lacks the sensitivity needed to detect uptake of low but biologically relevant amounts of DNA. On the other hand, strains R3021 and 477 took up more DNA than did noncompetent Rd but never yielded detectable transformants, confirming the importance of assaying for uptake as well as transformation.

DIFFERENCES IN THE REGULATION OF COMPETENCE

Regulation of competence is known to be much more evolutionarily labile than the mechanisms of DNA uptake and recombination (Solomon and Grossman 1996), so the between-strain differences in DNA uptake and transformation seen under the standard condition of MIV treatment could be due to differences in how competence is regulated rather than to mechanistic differences in DNA uptake and transformation. Because this is the first study of competence in other strains, we used two culture strategies that encompassed a broad range of growth conditions. First, the time course of competence development was followed for six strains growing in sBHI and after transfer to MIV. Second, the transformation frequency was measured in all strains after they had been cultured overnight in rich medium supplemented with chromosomal DNA; this assay revealed the cumulative effects of differences in the strains' responses to environmental or physiological signals, integrated over the whole growth period.

Time course of competence induction

For this analysis, we chose two strains whose MIV-induced transformation frequencies were like that of Rd (375 and PittAA), one strain with a very low and variable transformation frequency (PittHH), and two strains that gave no transformants (R3021, PittBB). Two time courses were done for each strain, and Rd was used as a control. We tested whether the timing of induction differed between the strains by following competence development over several hours of growth in sBHI and after transfer to MIV. Figure 3A shows the approximate cell densities at which transformation frequencies were measured, superimposed on a typical growth curve of Rd. Table 3 summarizes the time course results, with the results of the MIV assays described above included for comparison. (In this table the terms "log" and "late log" refer to actively growing cultures with OD_{600} less than 0.2 and greater than 0.6, respectively).

As expected from previous studies, Rd in log phase in sBHI produced no transformants (Redfield 1991); transformation became detectable when the OD_{600} reached ~0.2 and reached a frequency of $\sim 1 \times 10^{-4}$ once OD_{600} reached 1.0 (Fig. 3B). Strains PittBB and R3021 were consistently nontransformable at all time points (thus they are not shown in Fig. 3B). This suggests that their failure to transform in the MIV experiments shown in Figure 1 was due to a genuine inability to transform and not to regulatory differences. The two PittHH replicates gave different results, with one giving no significant transformation and the other low but significant transformation frequencies of 10^{-7} in log phase and 1.5×10^{-8} in late log (Table 3 and Fig. 3B). Similar differences between replicate cultures were seen in the MIV assays shown in Figure 1B. Despite the variation, the failure to induce competence in late log clearly distinguishes PittHH from strain Rd.

The two other strains tested resembled Rd in being much more competent in late log than in log phase (Table 3 and Fig. 3B), although neither exhibited the kinetics typical of Rd. Like Rd, strain 375 failed to transform in log phase but the onset of competence was more abrupt and also declined sooner and more rapidly. Strain PittAA differed from Rd by reproducibly transforming in log phase; however, its late log induction paralleled that of Rd.

In MIV, Rd cells develop competence slowly; expression of competence genes peaks after about 30 min, but maximal

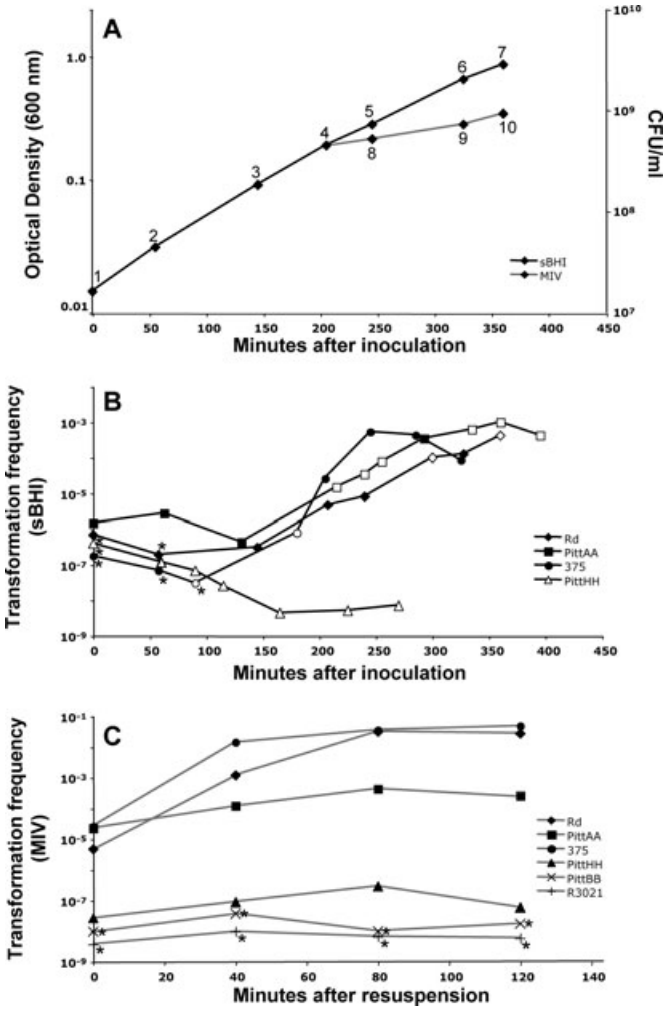

Figure 3. **Time course of competence induction. (A) A schematic growth curve for Rd. Numbered points indicate the approximate times at which cultures were sampled and transformation assays performed. (B and C) Transformation frequencies. Symbols with asterisks indicate the detection limit for time points where no transformants were observed. (B) Transformation frequencies in sBHI (time points 1–7 in A) as a function of time after inoculation. Because strains grew at different rates and to different final densities, samples for transformation assays were not always taken at the same times. Furthermore, samples from replicate cultures of the same strain were not always taken at the same time because of differences in growth characteristics between experiments. Thus, the solid symbols in Figure 3B represent means of replicate (same time) samples and open symbols represent single samples from differing times in the replicates. (C) Transformation frequencies in MIV (time points 8–10 in A) as a function of time after resuspension in MIV. The value for time zero is the transformation frequency immediately prior to resuspension in MIV.**

transformation frequencies are not reached until 60–100 min after transfer (Redfield et al. 2005). To detect any kinetic differences between strains, several time course samples were taken after cells were resuspended in MIV. Figure 3C shows the transformation frequencies for all six strains. The data are also summarized in the second to last column in Table 3, with the data from Figure 1B reproduced in the last column for comparison. As expected, the transformation frequency of Rd increased 10,000-fold after transfer to MIV, stabilizing at 80 min after transfer. The peak transformation frequencies of the other tested strains were consistent with those seen in the experiments summarized in Figure 1B. Strain 375 reached the same high transformation frequency as Rd, but took less time to do so. In contrast, the already-high transformation frequency of PittAA only increased 10-fold upon transfer to MIV. In both strains, transformation frequencies remained stable over the next 80 min. The three strains that transformed poorly or not at all in sBHI behaved similarly in MIV. Strain PittHH showed slight induction in MIV, reaching a peak transformation frequency of 3×10^{-7} after 80 min. The other strains PittBB and R3021 gave no significant transformants.

Overall the time courses showed two things. First, strains that had very low or undetectable transformation frequencies after 100 min in the original MIV assays (Fig. 1B) did not become transformable under other growth conditions. Thus, their failure to transform is not due to differences in the dependence of competence on growth. Second, one of the strains, PittAA, clearly regulates competence very differently from Rd, with substantial transformation during log phase and much less induction by MIV.

Transformation in rich medium

This technically simple assay allows transformation to occur in any of a wide range of physiological states, from the onset of growth after transfer to rich medium, through exponential growth, to the cessation of growth and overnight adaptation to postgrowth conditions. Each of these conditions is known to stimulate competence development in at least one other bacterial species (Solomon and Grossman 1996). Using chromosomal DNA (carrying a str^R marker) more closely mimicked the DNA available in mucosal environments.

These assays were expected to be less sensitive than the MIV assays due to two factors. First, strains that regulate competence like Rd are expected to have transformation frequencies 10- to 100-fold lower in this assay than in MIV. Second, assays using chromosomal DNA are less sensitive; although chromosomal DNA is taken up very efficiently, the effective concentration of the selective marker is much lower than in the MIV assays, which used a cloned nov^R DNA fragment. Consistent with this, Rd transformed at a frequency of 4×10^{-6}, 2500-fold lower than its median transformation frequency in MIV. Furthermore, 17 strains that produced transformants in MIV were not detectably transformable in this assay (Table S1).

Most of the other 17 strains had, as expected, lower transformation frequencies in this assay than in MIV. However, the assay

Table 3. Transformation frequencies in a subset of *H. influenzae* strains. Means and ranges are shown. Bold indicates the mean value.

	Time course			MIV
	Log TF (sBHI)	Late log TF (sBHI)	TF (MIV)	
Rd	**$< 1.4\times10^{-7}$**	**1.6×10^{-4}**	**0.010**	**0.013**
	7.5×10^{-8}–4.1×10^{-7}	9.0×10^{-5}–1.0×10^{-4}	3.7×10^{-4}–4.0×10^{-2}	3.7×10^{-3}–4.5×10^{-2}
PittBB	**$< 2.5\times10^{-7}$**	**$< 6.1\times10^{-9}$**	**$< 1.9\times10^{-8}$**	**$< 1.9\times10^{-8}$**
	4.8×10^{-8}–7.7×10^{-7}	5.7×10^{-10}–4.4×10^{-8}	4.7×10^{-9}–2.2×10^{-7}	1.5×10^{-8}–2.9×10^{-8}
R3021	**$< 1.9\times10^{-8}$**	**$< 1.3\times10^{-9}$**	**$< 7.2\times10^{-9}$**	**$< 2.4\times10^{-7}$**
	6.0×10^{-9}–2.3×10^{-7}	4.1×10^{-10}–3.3×10^{-9}	4.4×10^{-9}–2.1×10^{-8}	7.7×10^{-8}–1.8×10^{-6}
PittHH	**9.9×10^{-8}**	**1.3×10^{-8}**	**1.2×10^{-7}**	**3.0×10^{-8}**
	6.9×10^{-8}–1.3×10^{-7}	5.3×10^{-9}–3.1×10^{-8}	1.1×10^{-8}–3.9×10^{-7}	5.9×10^{-9}–5.5×10^{-8}
375	**1.0×10^{-7}**	**2.0×10^{-4}**	**0.030**	**0.014**
	3.1×10^{-8}–8.1×10^{-7}	1.1×10^{-5}–1.0×10^{-3}	6.0×10^{-3}–8.5×10^{-2}	4.0×10^{-3}–4.0×10^{-2}
PittAA	**4.4×10^{-7}**	**4.7×10^{-4}**	**2.3×10^{-4}**	**1.1×10^{-3}**
	1.3×10^{-7}–1.5×10^{-6}	1.8×10^{-4}–1×10^{-3}	1.3×10^{-4}–5.9×10^{-4}	7.2×10^{-6}–2.6×10^{-2}

identified two strains that appeared to regulate competence differently than Rd: they were more transformable than Rd in this assay although less transformable than Rd in MIV. Strains R2866 and 86028NP transformed at 8.2×10^{-5} and 8.7×10^{-6} in this assay but at 1.3×10^{-4} and 5.5×10^{-6} in MIV, respectively. Together the two culture strategies identified three strains with regulatory differences from Rd. But the transformation phenotypes of most strains were consistent with the hypothesis that they regulate their competence in the same way as Rd.

VARIATION IN COMPETENCE GENES

To identify candidate genetic changes underlying the observed variation in competence, we examined the sequences of competence genes in those 12 strains with available genome sequences (see the first footnote in Table 1). Products of 32 Rd genes are associated with DNA uptake and transformation, and alleles at each locus differed from each other by 2.7% ($\pm$ 0.5%) of nucleotides and 3.6% ($\pm$ 0.4%) of amino acids (Table 4). Below we describe differences between strains in competence gene content and sequence (summarized in Fig. 4).

All but six competence genes were present and apparently functional in all 12 strains. We first discuss two genes that were missing from some strains. Most strains lacked a homolog of the Rd gene HI1631, which is also absent from other Pasteurellaceae species (Maughan et al. 2008). Most strains also had incomplete homologs of the Rd gene HI0660, either lacking the first 177 nucleotides or having only the first 127 nucleotides. Although nothing is known about these genes' functions, transcription of both is strongly induced upon transfer from sBHI to MIV and is dependent on the competence regulatory proteins Sxy and Crp (Redfield et al. 2005). They are unlikely to be essential for competence, because their presence or absence did not correlate with competence development.

Frameshifts or premature stop codons were found in 33 out of 360 sequences obtained from the databases (Table S2). However, resequencing revealed that 29 of the apparent frameshifts were due to errors in the original sequences; almost all of these were from incomplete projects. Although the three frameshifts and one nonsense mutation described below create truncated proteins, each allele contains only the single inactivating mutation, and alignment of the gene sequences with apparently functional alleles from other strains shows little divergence. This suggests that the inactivating mutations occurred sufficiently recently that additional mutations have not accumulated.

An early frameshift truncates *comM* in strain R2866 from 509-amino acids (aa) to only 47 aa. ComM protects incoming DNA from cytoplasmic nucleases; its primary function is unlikely to be in transformation. Although a knockout in Rd reduces transformation 300-fold (Gwinn et al. 1998), *comM* homologs are found in many bacteria not known to be competent (reviewed in Maughan et al. [2008]). Despite this defect R2866 takes up a lot of DNA and its transformation frequency is quite high ($\sim10^{-4}$; Fig. 1).

pilD has a frameshift in strain 22.1-21, truncating the protein from 230 to 181 aa. *pilD* is a member of the *pilABCD* operon, and encodes a protease required for assembly of pili, hair-like structures that extend from the outer membrane and whose retraction is thought to generate the force driving DNA uptake in almost all competent bacteria (Chen and Dubnau 2004). The predicted amino acid sequence of the 22.1-21 PilA is also very divergent from the other strains. Although pili are usually essential for DNA uptake, this strain does take up DNA and produces transformants at a frequency of $\sim10^{-6}$ (Fig. 1).

comD has a frameshift in strain R3021, truncating the protein from 137 to 37 aa. *comD* is a member of the *comABCDE* operon that is induced over 100-fold in Rd during competence

Table 4. Competence genes used in the analysis.

Gene name	Length (bp)	Nucleotide diversity (π)	Tajima's D	π_A/π_S	Function
comA	798	0.02	−1.097	0.19	Uptake
comB	507	0.018	−1.22	0.12	Uptake
comC	522	0.037	0.274	0.19	Uptake
comD[1]	414	0.026	0.908	0.26	Surface exposed
comE	1338	0.019	0.605	0.18	Surface exposed
comE1	339	0.065	1.45	0.47	Surface exposed
comF	687	0.02	0.45	0.24	Uptake
comM[2]	1530	0.031	−0.21	0.03	Transformation
comN	513	0.022	−0.84	0.08	Surface exposed
comO	717	0.018	−0.915	0.08	Surface exposed
comP	684	0.031	−0.44	0.26	Surface exposed
comQ	306	0.013	−1.55	0.42	Surface exposed
crp	675	0.0037	−0.098	0	Regulation
cyaA	2532	0.018	−1.06	0.09	Regulation
dprA	1122	0.036	1.41	0.03	Transformation
icc	825	0.036	0.266	0.07	Regulation
ligA[3]	558	0.032	0.329	0.14	Unknown
murE	1467	0.012	−0.43	0.13	Unknown
pilA	450	0.07	0.07	0.34	Surface exposed
pilB	1395	0.048	−0.1	0.13	Uptake
pilC	1221	0.041	−0.22	0.09	Uptake
pilD[4]	693	0.052	0.45	0.25	Uptake
pilF2	540	0.035	0.43	0.19	Surface exposed
radC	705	0.02	−0.86	0.16	Transformation
rec-1	1065	0.02	−0.22	0	Transformation
rec-2	2367	0.022	0.03	0.24	Uptake
ssb	507	0.037	0.09	0.16	Transformation
sxy	654	0.008	1.12	0.27	Regulation
0365	1173	0.032	0.456	0.04	Unknown
0659	297	0.015	−0.376	0	Unknown

[1]frameshift in R3021; [2]frameshift in strain R2866; [3]nonsense mutation in PittAA; [4]frameshift in strain 22.1-21.

development (Redfield et al. 2005). A *comD* frameshift is expected to be polar on the essential DNA uptake gene *comE*, and this defect may explain its 50-fold lower uptake than Rd. However, the function of ComD is not known, and in principle its *comD* defect might be responsible for R3021's inability to transform.

A point mutation in *ligA* of strain PittAA creates a premature stop codon that eliminates the last 18 of its 268 aa. LigA is a ligase localized to the periplasm; although its expression is induced 100-fold in Rd upon transfer to MIV, a deletion in strain Rd reduced transformation only fivefold (VanWagoner et al. 2004; Redfield et al. 2005). Consistent with this, strain PittAA takes up moderate amounts of DNA and transforms efficiently.

In addition to examining the sequences of competence genes in each strain individually, we used population genetic analyses to investigate how selection has acted on competence. Because molecular population genetics relies on measures of sequence polymorphism, we were concerned that sequencing errors could affect our results. However, our resequencing suggests an error rate of ~0.08% (29 mistakes in 38 kb), too low to significantly affect estimates of polymorphism based on genes with 3% nucleotide diversity (Clark and Whittam 1992). To determine whether most single nucleotide variants are selectively neutral, we examined polymorphism in competence genes using Tajima's D and Fu and Li's D statistics (Tajima 1989; Fu and Li 1993). Significantly negative values of these statistics would indicate that a bottleneck has occurred or an advantageous allele has swept through the population, whereas significantly positive values would indicate that selection has favored diversity, for example, alleles have been beneficial because they are rare. The actual values for these statistics were not significantly different from zero, consistent with most sequence differences between strains being selectively neutral (Table 4 shows Tajima's *D*-values). This suggests that competence gene alleles are functionally equivalent with respect

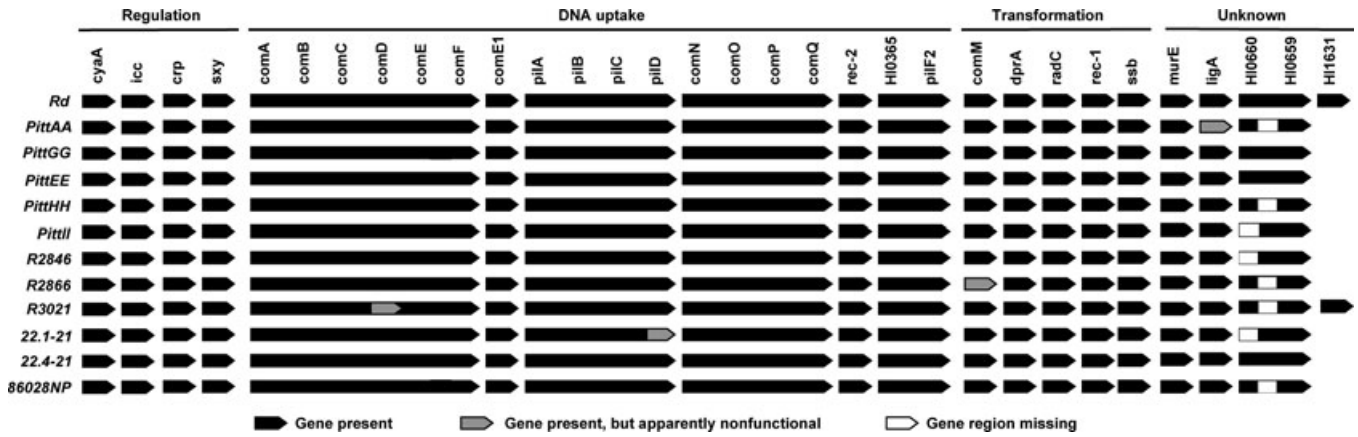

Figure 4. **Competence genes in 12 strains with genome sequences available. Strain names are listed on the far left and competence genes, divided by functional category, are listed on top. Black arrows indicate transcriptional units that are present and predicted to be functional. Gray arrows indicate genes that are present but inactivated by mutation. White regions indicate areas missing from some strains. Arrows do not scale with the real sizes of genes.**

to fitness and most mutations altering function were deleterious and removed from the population by selection.

To confirm that selection has removed most deleterious mutations from the population, we calculated the ratio of nonsynonymous to synonymous polymorphisms (π_A/π_S). High values of π_A/π_S indicate that selection favors mutations that change the encoded amino acid or is not effective in removing them, whereas low values indicate that most amino acid changes are deleterious and efficiently purged from the population. π_A/π_S values were all well below 1, with an average of 0.16 and a range of 0 (*crp*) to 0.47 (*comE1*) (Table 4), within the range of values described for many bacterial species (Hughes et al. 2008). Strong purifying selection thus is likely to be the dominant force acting on competence genes, a result consistent with the *D*-statistics described above.

To determine whether selection has acted similarly on all competence genes, we pooled genes into functional categories and calculated whether values of π_A/π_S differed significantly between functions. Because some DNA uptake proteins are located on the cell surface, and pathogenic and commensal bacteria are subject to diversifying selection imposed by the host immune system, we further divided these into surface-exposed and internal groups. This resulted in five groups: regulation, uptake (internal), transformation, surface-exposed, and unknown. Regulation, transformation, uptake, and unknown genes had statistically indistinguishable values of π_A/π_S. Genes in the cell-surface associated category had π_A/π_S values significantly higher than regulation, transformation and unknown genes but did not differ significantly from uptake genes (Fig. 5). This analysis shows that all competence genes evolve at similar rates unless their products are localized to the cell surface.

Discussion

The summary in Figure 2 emphasizes that measuring transformation gives only a partial picture of the extent of variation. The variation is particularly striking for strains with transformation frequencies between 10^{-3} and 10^{-2}, whose DNA uptake values varied over more than 100-fold. Similarly, strains whose normalized DNA uptake was around 10^{-2} had transformation frequencies that ranged over nearly five orders of magnitude.

The standard *H. influenzae* strain Rd provides a reference point for thinking about this variation. Rd took up more DNA than any other strain; its transformation frequency was similar to some strains but 10- to 100-fold higher than many others. Although Rd has been used to study competence in the laboratory for decades, this high competence is likely to result from evolution in its natural environment rather than to laboratory selection. This strain was originally chosen for study because of its naturally high transformation frequency (Alexander and Leidy 1951), and the original frozen stock has only been subcultured a few times (although many derivative strains have been created by transformation and given new names). Its high transformation frequency may also not be due to selection to maximize the recombination benefits of a limited supply of DNA, as strains PittAA and Eagan produced similar numbers of transformants while taking up 100-fold less DNA. On the other hand, some strains were less efficient, taking up substantial amounts of DNA but recombining it much less efficiently than Rd, and some produced no transformants at all. These discrepancies suggest that the benefits of DNA uptake are partially uncoupled from those of transformation.

The adaptive significance of any phenotypic variation must be assessed in the context of its proximate (genetic) and ultimate (evolutionary) causes. In addition to the specific proximate and

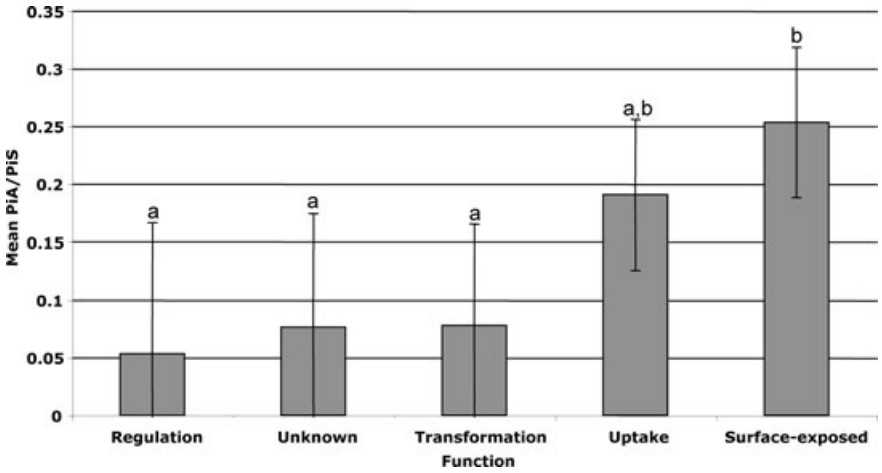

Figure 5. Mean π_A/π_S values for each functional category. Gray bars indicate the mean and error bars represent 95% confidence intervals. Letters above gray bars represent significantly different means, as determined by Tukey–Kramer HSD and Wilcoxon/Kruskal–Wallis rank sum tests.

ultimate factors affecting evolution of DNA uptake and transformation considered below, several aspects of *H. influenzae* population biology should be mentioned. First, the amount of genetic variation in MLST (Meats et al. 2003) and competence genes suggests a large effective population size, so deleterious mutations are likely to be removed efficiently. However, weakly deleterious mutations are expected to persist for extended periods. Second, there is no evidence that strains are locally adapted to specific niches in the host. Strains in the respiratory tract are known to turn over frequently, and there is no evidence that different strains have become adapted to specific subniches (Trottier et al. 1989; Faden et al. 1995; Moller et al. 1995; van Alphen et al. 1997; Dhooge et al. 2000; Farjo et al. 2004; Mukundan et al. 2007; Erwin et al. 2008). Selection is unlikely to act at other sites of infection—sepsis and meningitis are the most serious *H. influenzae* infections, but adaptation to the corresponding bloodstream and meninges niches is very unlikely as both are evolutionary dead ends, environments from which cells cannot easily pass to new hosts. Third, analysis of the evolutionary causes is necessarily speculative, because so little is known about *H. influenzae*'s competence and transformability in its natural environment. There has only been one in vivo study of transformation, with frequencies of 10^{-4} observed in the lining of rat abdominal cavities (Dargis et al. 1992). In its natural human environment *H. influenzae* cells live in DNA-rich biofilms on human respiratory mucosa, where cells are known to express competence-regulated pilus genes (Jurcisek et al. 2007).

The proximate cause of several strains' complete inability to take up DNA could be loss-of-function mutations in the competence-regulating genes *crp* and *sxy* or in any of the genes whose products are essential components of the DNA uptake machinery (Maughan et al. 2008). Less severe uptake defects in other strains could of course be due to less severe defects in these genes. Our analysis of competence genes showed that strains were identical in their predicted sequence of Crp, but we found several amino acid differences in Sxy and in some genes required for DNA uptake, and identified two genes expressed in Rd during competence development but absent from the majority of other strains examined. Differences in DNA uptake could also be due to differences in genes whose role in competence has not been discovered; some strains of *H. influenzae* differ in their gene content by over 500 genes (Hogg et al. 2007).

Two additional factors could cause variation in DNA uptake. First, the proportion of cells in the culture that become competent may differ between strains; the fraction of Rd cells that become competent is known to be very dependent on the inducing condition (Goodgal and Herriott 1961). Second, DNA uptake activates a Mu-like prophage in the Rd genome, with the extent of killing depending on the divergence of the donor DNA (Setlow et al. 1973; Albritton et al. 1984). Because different strains commonly have different prophages (Hogg et al. 2007), the level of killing of competent cells could also differ between strains. Selection on these factors would not have affected variation in the competence genes we examined.

Selection for DNA uptake is likely to fluctuate on multiple time scales. On a physiological time-scale (one or a few cell generations) the tight regulation of competence genes must be due to selection for optimal expression under circumstances in which the costs and benefits of DNA uptake fluctuate. Longer-term differences in the ratio of costs to benefits are also likely, and might be expected to expose whole populations to differing selection pressures for extended periods. For example, an ongoing arms race between a phage and its *H. influenzae* host might make DNA uptake consistently deleterious for many generations because of persistently high levels of phage DNA in the environment. This would select for strains that took up little or no DNA under any

conditions, like strains R2846 and 22.4-21. On the other hand, a mutation reducing the efficiency of de novo nucleotide synthesis or a consistent scarcity of nucleotides and their precursors in the host mucosa might produce strong selection favoring uptake of large amounts of DNA, resulting in strains Rd and PittDD.

Strain-to-strain variation in transformation was even more dramatic than the variation in uptake. The proximate causes of this variation must be differences in the activities of proteins that affect DNA degradation or the physical recombination of DNA strands. Survival of incoming DNA requires the products of the competence genes *dprA* and *comM*, which appear to protect DNA from cellular nucleases (Karudapuram et al. 1995; Gwinn et al. 1998; Berge et al. 2003; Mortier-Barriere et al. 2007). Recombination of incoming DNA in *H. influenzae* is known to depend on the actions of the recombinational repair protein Rec-1 (RecA homolog); the genome also contains homologs of many other proteins known to affect recombination (Kuzminov 1999). Mutations affecting these activities could thus be responsible for discrepancies between DNA uptake and transformation. Differences in transformation could also be due to strain-specific genes affecting DNA degradation or recombination, or to differing activities of as yet unidentified factors such as cytoplasmic nucleases. Differences in restriction modification systems are unlikely to contribute, as the single-stranded incoming DNA is not a substrate for most restriction enzymes.

Finally, we consider two hypotheses for the ultimate causes of variation in transformation: selection for DNA repair and selection for genetic exchange. If a cell's chromosome contains physical damage, recombination with incoming DNA can provide a template for repair, but the benefit is limited by the need to take up an undamaged strand that can base pair with the damaged chromosomal segment. Although we did not directly assay for recombinational repair, our measurements of transformation also measured the capacity for this repair, as the cellular enzymes required are the same. Causes of DNA damage are ubiquitous (reactive byproducts of metabolism, errors by DNA polymerase, UV radiation, and environmental chemicals), and damage that affects both strands of a chromosome is lethal unless a template for repair is available. If transformation were very important for repair, strains that are completely nontransformable should be very rare. Strains of many bacterial species have been shown to differ in their mismatch repair (Matic et al. 1997; Oliver et al. 2000; Denamur et al. 2002; Richardson et al. 2002; Prunier et al. 2003), but the extent of this variation pales in comparison to the six orders of magnitude variation in transformation in *H. influenzae*.

Transformation may also have been selected because it causes genetic exchange between cells. Extensive research on the evolution of sexual recombination in eukaryotes has identified many potential genetic costs and benefits of recombination but has not provided a clear evolutionary explanation (Kondrashov 1993; Otto and Gerstein 2006). Although most of these costs and benefits also apply to transformational recombination, selection on transformational recombination is further complicated by the predicted excess of deleterious mutations in transforming DNA and by the asymmetric potential for loss of competence-promoting alleles through recombination (Redfield 1988; Redfield et al. 1997). Because these costs and benefits depend on both the genotype of the competent cell and the alleles carried by individual incoming DNA fragments, they are likely to vary much more dramatically than the direct costs and benefits of DNA uptake.

Little is known about *H. influenzae*'s life in its natural environment, partly because the very small scales relevant to bacterial ecology are difficult to investigate, and partly because *H. influenzae* is a moderately fastidious human commensal. However, the ability to take up DNA is clearly important, because although competence itself is often reduced or lost, the genes responsible are not. The genetic and phenotypic variation we have documented thus supports the model of competence evolution proposed by Redfield et al. (2006). In this model, the ancestral Pasteurellacean and the ancestral *H. influenzae* were naturally competent, with substantial variation arising in their descendants through mutation. During subsequent evolution noncompetent lineages were not always efficiently eliminated, and may sometimes have been actively selected. However, these lineages did not persist indefinitely, and all extant lineages are descended from relatively recent ancestors who were competent.

ACKNOWLEDGMENTS

The authors wish to thank members of the Redfield Laboratory for discussions of this work and R. Moxon, G. Ehrlich, R. Munson, A. Smith, and A. Erwin for kindly shipping strains to us and for comments on an earlier version of this article. HM was supported by a National Institutes of Health Ruth L. Kirschstein postdoctoral fellowship and a Killam postdoctoral fellowship.

LITERATURE CITED

Albritton, W. L., J. K. Setlow, M. Thomas, F. Sottnek, and A. G. Steigerwalt. 1984. Heterospecific transformation in the genus *Haemophilus*. Mol. Gen. Genet. 193:358–363.

Alexander, H. E., and G. Leidy. 1951. Determination of inherited traits of *H. influenzae* by desoxyribonucleic acid fractions isolated from type-specific cells. J. Exp. Med. 93:345–359.

Bakaletz, L. O., B. D. Baker, J. A. Jurcisek, A. Harrison, L. A. Novotny, J. E. Bookwalter, R. Mungur, and R. S. Munson Jr. 2005. Demonstration of Type IV pilus expression and a twitching phenotype by *Haemophilus influenzae*. Infect. Immun. 73:1635–1643.

Berge, M., I. Mortier-Barriere, B. Martin, and J. P. Claverys. 2003. Transformation of *Streptococcus pneumoniae* relies on DprA- and RecA-dependent protection of incoming DNA single strands. Mol. Microbiol. 50:527–536.

Cameron, A. D., M. Volar, L. A. Bannister, and R. J. Redfield. 2008. RNA secondary structure regulates the translation of *sxy* and competence development in *Haemophilus influenzae*. Nucleic. Acids. Res. 36:10–20.

Chen, I., and D. Dubnau. 2004. DNA uptake during bacterial transformation. Nat. Rev. Microbiol. 2:241–249.

Clark, A. G., and T. S. Whittam. 1992. Sequencing errors and molecular evolutionary analysis. Mol. Biol. Evol. 9:744–752.

Coupat, B., F. Chaumeille-Dole, S. Fall, P. Prior, P. Simonet, X. Nesme, and F. Bertolla. 2008. Natural transformation in the *Ralstonia solanacearum* species complex: number and size of DNA that can be transferred. FEMS Microbiol. Ecol. 66:14–24.

Dargis, M., P. Gourde, D. Beauchamp, B. Foiry, M. Jacques, and F. Malouin. 1992. Modification in penicillin-binding proteins during in vivo development of genetic competence of *Haemophilus influenzae* is associated with a rapid change in the physiological state of cells. Infect. Immun. 60:4024–4031.

Denamur, E., S. Bonacorsi, A. Giraud, P. Duriez, F. Hilali, C. Amorin, E. Bingen, A. Andremont, B. Picard, F. Taddei, et al. 2002. High frequency of mutator strains among human uropathogenic *Escherichia coli* isolates. J. Bacteriol. 184:605–609.

Dhooge, I., M. Vaneechoutte, G. Claeys, G. Verschraegen, and P. Van Cauwenberge. 2000. Turnover of *Haemophilus influenzae* isolates in otitis-prone children. Int. J. Pediatr. Otorhinolaryngol. 54:7–12.

Erwin, A. L., S. A. Sandstedt, P. J. Bonthuis, J. L. Geelhood, K. L. Nelson, W. C. Unrath, M. A. Diggle, M. J. Theodore, C. R. Pleatman, E. A. Mothershed, et al. 2008. Analysis of genetic relatedness of *Haemophilus influenzae* isolates by multilocus sequence typing. J. Bacteriol. 190:1473–1483.

Faden, H., L. Duffy, A. Williams, D. A. Krystofik, and J. Wolf. 1995. Epidemiology of nasopharyngeal colonization with nontypeable *Haemophilus influenzae* in the first 2 years of life. J. Infect. Dis. 172:132–135.

Farjo, R. S., B. Foxman, M. J. Patel, L. Zhang, M. M. Pettigrew, S. I. McCoy, C. F. Marrs, and J. R. Gilsdorf. 2004. Diversity and sharing of *Haemophilus influenzae* strains colonizing healthy children attending day-care centers. Pediatr. Infect. Dis. J. 23:41–46.

Fu, Y. X., and W. H. Li. 1993. Statistical tests of neutrality of mutations. Genetics 133:693–709.

Fujise, O., L. Lakio, Y. Wang, S. Asikainen, and C. Chen. 2004. Clonal distribution of natural competence in *Actinobacillus actinomycetemcomitans*. Oral Microbiol. Immunol. 19:340–342.

Goodgal, S. H., and R. M. Herriott. 1961. Studies on transformations of *Haemophilus influenzae*. I. Competence. J. Gen. Physiol. 44:1201–1227.

Gromkova, R. C., T. C. Mottalini, and M. G. Dove. 1998. Genetic transformation in *Haemophilus parainfluenzae* clinical isolates. Curr. Microbiol. 37:123–126.

Gwinn, M. L., R. Ramanathan, H. O. Smith, and J. F. Tomb. 1998. A new transformation-deficient mutant of *Haemophilus influenzae* Rd with normal DNA uptake. J. Bacteriol. 180:746–748.

Herriott, R. M., E. M. Meyer, and M. Vogt. 1970. Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*. J. Bacteriol. 101:517–524.

Hogg, J. S., F. Z. Hu, B. Janto, R. Boissy, J. Hayes, R. Keefe, J. C. Post, and G. D. Ehrlich. 2007. Characterization and modeling of the *Haemophilus influenzae* core and supragenomes based on the complete genomic sequences of Rd and 12 clinical nontypeable strains. Genome Biol. 8:R103.

Hughes, A. L., R. Friedman, P. Rivailler, and J. O. French. 2008. Synonymous and nonsynonymous polymorphisms versus divergences in bacterial genomes. Mol. Biol. Evol. 25:2199–2209.

Jurcisek, J. A., J. E. Bookwalter, B. D. Baker, S. Fernandez, L. A. Novotny, R. S. Munson, Jr., and L. O. Bakaletz. 2007. The PilA protein of nontypeable *Haemophilus influenzae* plays a role in biofilm formation, adherence to epithelial cells and colonization of the mammalian upper respiratory tract. Mol. Microbiol. 65:1288–1299.

Kahn, M. E., and R. Gromkova. 1981. Occurrence of pili on and adhesive properties of *Haemophilus parainfluenzae*. J. Bacteriol. 145:1075–1078.

Karudapuram, S., X. Zhao, and G. J. Barcak. 1995. DNA sequence and characterization of *Haemophilus influenzae dprA+*, a gene required for chromosomal but not plasmid DNA transformation. J. Bacteriol. 177:3235–3240.

Kondrashov, A. S. 1993. Classification of hypotheses on the advantage of amphimixis. J. Hered. 84:372–387.

Kramer, N., J. Hahn, and D. Dubnau. 2007. Multiple interactions among the competence proteins of *Bacillus subtilis*. Mol. Microbiol. 65:454–464.

Kuzminov, A. 1999. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. Microbiol. Mol. Biol. Rev. 63:751–813, table of contents.

MacFadyen, L. P., D. Chen, H. C. Vo, D. Liao, R. Sinotte, and R. J. Redfield. 2001. Competence development by *Haemophilus influenzae* is regulated by the availability of nucleic acid precursors. Mol. Microbiol. 40:700–707.

Majewski, J., and F. M. Cohan. 1999. DNA sequence similarity requirements for interspecific recombination in *Bacillus*. Genetics 153:1525–1533.

Martin, B., Y. Quentin, G. Fichant, and J. P. Claverys. 2006. Independent evolution of competence regulatory cascades in streptococci? Trends Microbiol. 14:339–345.

Matic, I., M. Radman, F. Taddei, B. Picard, C. Doit, E. Bingen, E. Denamur, and J. Elion. 1997. Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. Science 277:1833–1834.

Matthews, L. W., S. Spector, J. Lemm, and J. L. Potter. 1963. Studies on pulmonary secretions. I. the over-all chemical composition of pulmonary secretions from patients with cystic fibrosis, bronchiectasis, and laryngectomy. Am. Rev. Respir. Dis. 88:199–204.

Maughan, H., S. Sinha, L. Wilson, and R. J. Redfield. 2008. Competence, DNA uptake, and transformation in *Pasteurellaceae*. Pp. 81–100 *in* P. Kuhnert and H. Christensen, eds. *Pasteurellaceae*: biology, genomics, and molecular aspects. Caister Academic Press, Norfolk, U.K.

Meats, E., E. J. Feil, S. Stringer, A. J. Cody, R. Goldstein, J. S. Kroll, T. Popovic, and B. G. Spratt. 2003. Characterization of encapsulated and noncapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. J. Clin. Microbiol. 41:1623–1636.

Michod, R. E., H. Bernstein, and A. M. Nedelcu. 2008. Adaptive value of sex in microbial pathogens. Infect. Genet. Evol. 8:267–285.

Moller, L. V., A. G. Regelink, H. Grasselier, J. E. Dankert-Roelse, J. Dankert, and L. van Alphen. 1995. Multiple *Haemophilus influenzae* strains and strain variants coexist in the respiratory tract of patients with cystic fibrosis. J. Infect. Dis. 172:1388–1392.

Mortier-Barriere, I., M. Velten, P. Dupaigne, N. Mirouze, O. Pietrement, S. McGovern, G. Fichant, B. Martin, P. Noirot, E. Le Cam, et al. 2007. A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. Cell 130:824–836.

Mukundan, D., Z. Ecevit, M. Patel, C. F. Marrs, and J. R. Gilsdorf. 2007. Pharyngeal colonization dynamics of *Haemophilus influenzae* and *Haemophilus haemolyticus* in healthy adult carriers. J. Clin. Microbiol. 45:3207–3217.

Nyhan, W. L. 2005. Encyclopedia of life sciences. John Wiley and Sons, Ltd., Chichester, U.K.

Oliver, A., R. Canton, P. Campo, F. Baquero, and J. Blazquez. 2000. High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. Science 288:1251–1254.

Otto, S. P., and A. C. Gerstein. 2006. Why have sex? The population genetics of sex and recombination. Biochem. Soc. Trans. 34:519–522.

Palchevskiy, V., and S. E. Finkel. 2006. *Escherichia coli* competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient. J. Bacteriol. 188:3902–3910.

Poje, G., and R. J. Redfield. 2003. Transformation of *Haemophilus influenzae*. Methods Mol. Med. 71:57–70.

Pozzi, G., L. Masala, F. Iannelli, R. Manganelli, L. S. Havarstein, L. Piccoli, D. Simon, and D. A. Morrison. 1996. Competence for genetic transformation in encapsulated strains of *Streptococcus pneumoniae*: two allelic variants of the peptide pheromone. J. Bacteriol. 178:6087–6090.

Prunier, A. L., B. Malbruny, M. Laurans, J. Brouard, J. F. Duhamel, and R. Leclercq. 2003. High rate of macrolide resistance in *Staphylococcus aureus* strains from patients with cystic fibrosis reveals high proportions of hypermutable strains. J. Infect. Dis. 187:1709–1716.

Ramirez, M., D. A. Morrison, and A. Tomasz. 1997. Ubiquitous distribution of the competence related genes *comA* and *comC* among isolates of *Streptococcus pneumoniae*. Microb. Drug. Resist. 3:39–52.

Rayner, M. G., Y. Zhang, M. C. Gorry, Y. Chen, J. C. Post, and G. D. Ehrlich. 1998. Evidence of bacterial metabolic activity in culture-negative otitis media with effusion. JAMA 279:296–299.

Redfield, R. J. 1988. Evolution of bacterial transformation: is sex with dead cells ever better than no sex at all? Genetics 119:213–221.

———. 1991. *sxy-1*, a *Haemophilus influenzae* mutation causing greatly enhanced spontaneous competence. J. Bacteriol. 173:5612–5618.

Redfield, R. J., M. R. Schrag, and A. M. Dean. 1997. The evolution of bacterial transformation: sex with poor relations. Genetics 146:27–38.

Redfield, R. J., A. D. Cameron, Q. Qian, J. Hinds, T. R. Ali, J. S. Kroll, and P. R. Langford. 2005. A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. J. Mol. Biol. 347:735–747.

Redfield, R. J., W. A. Findlay, J. Bosse, J. S. Kroll, A. D. Cameron, and J. H. Nash. 2006. Evolution of competence and DNA uptake specificity in the *Pasteurellaceae*. BMC Evol. Biol. 6:82.

Richardson, A. R., Z. Yu, T. Popovic, and I. Stojiljkovic. 2002. Mutator clones of *Neisseria meningitidis* in epidemic serogroup A disease. Proc. Natl. Acad. Sci. USA 99:6103–6107.

Rowji, P., R. Gromkova, and H. Koornhof. 1989. Genetic transformation in encapsulated clinical isolates of *Haemophilus influenzae* type b. J. Gen. Microbiol. 135:2775–2782.

Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497.

Scocca, J. J., R. L. Poland, and K. C. Zoon. 1974. Specificity in deoxyribonucleic acid uptake by transformable *Haemophilus influenzae*. J. Bacteriol. 118:369–373.

Setlow, J. K., M. E. Boling, D. P. Allison, and K. L. Beattie. 1973. Relationship between prophage induction and transformation in *Haemophilus influenzae*. J. Bacteriol. 115:153–161.

Seto, H., and A. Tomasz. 1975. Selective release of a deoxyribonucleic acid-binding factor from the surface of competent pneumococci. J. Bacteriol. 124:969–976.

Shen, P., and H. V. Huang. 1986. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. Genetics 112:441–457.

Sikorski, J., N. Teschner, and W. Wackernagel. 2002. Highly different levels of natural transformation are associated with genomic subgroups within a local population of *Pseudomonas stutzeri* from soil. Appl. Environ. Microbiol. 68:865–873.

Smith, H. O., J. F. Tomb, B. A. Dougherty, R. D. Fleischmann, and J. C. Venter. 1995. Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. Science 269:538–540.

Solomon, J. M., and A. D. Grossman. 1996. Who's competent and when: regulation of natural genetic competence in bacteria. Trends Genet. 12:150–155.

Stuy, J. H., and R. B. Walter. 1986. Homology-facilitated plasmid transfer in *Haemophilus influenzae*. Mol. Gen. Genet. 203:288–295.

Swords, W. E., B. A. Buscher, K. Ver Steeg Ii, A. Preston, W. A. Nichols, J. N. Weiser, B. W. Gibson, and M. A. Apicella. 2000. Non-typeable *Haemophilus influenzae* adhere to and invade human bronchial epithelial cells via an interaction of lipooligosaccharide with the PAF receptor. Mol. Microbiol. 37:13–27.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Trottier, S., K. Stenberg, and C. Svanborg-Eden. 1989. Turnover of nontypable *Haemophilus influenzae* in the nasopharynges of healthy children. J. Clin. Microbiol. 27:2175–2179.

van Alphen, L., D. A. Caugant, B. Duim, M. O'Rourke, and L. D. Bowler. 1997. Differences in genetic diversity of nonecapsulated *Haemophilus influenzae* from various diseases. Microbiology 143(Pt 4):1423–1431.

VanWagoner, T. M., P. W. Whitby, D. J. Morton, T. W. Seale, and T. L. Stull. 2004. Characterization of three new competence-regulated operons in *Haemophilus influenzae*. J. Bacteriol. 186:6409–6421.

Vlassov, V. V., P. P. Laktionov, and E. Y. Rykova. 2007. Extracellular nucleic acids. Bioessays 29:654–667.

Zawadzki, P., M. S. Roberts, and F. M. Cohan. 1995. The log-linear relationship between sexual isolation and sequence divergence in *Bacillus* transformation is robust. Genetics 140:917–932.

Associate Editor: M. Travisano

Supporting Information

The following supporting information is available for this article:

Table S1. Transformation frequencies in rich medium.

Table S2. Competence genes containing frameshifts or premature stop codons in sequences from GenBank.

Supporting Information may be found in the online version of this article.
(This link will take you to the article abstract).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.
Additional results and discussion can be found in a document at http://www.repository.naturalis.nl/record/289893.