RESEARCH LETTER

Comparative genomics of *Neisseria weaveri* clarifies the taxonomy of this species and identifies genetic determinants that may be associated with virulence

Hana Yi[1], Yong-Joon Cho[2], Seok-Hwan Yoon[3], Sang-Cheol Park[3] & Jongsik Chun[1,2,3,4]

[1]Institute of Molecular Biology and Genetics, Seoul, Korea; [2]Chunlab, Inc., Seoul, Korea; [3]School of Biological Sciences, Seoul National University, Seoul, Korea; and [4]AICT, Gyeonggi-do, Korea

Correspondence: Jongsik Chun, School of Biological Sciences, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul 151-742, Korea. Tel.: +82 2 880 8153; fax: +82 2 874 8153; e-mail: jchun@snu.ac.kr

Received 1 December 2011; revised 9 December 2011; accepted 9 December 2011. Final version published online 6 January 2012.

DOI: 10.1111/j.1574-6968.2011.02485.x

Editor: Aharon Oren

Keywords
Neisseria weaveri; comparative genomics; reclassification.

Abstract

A group of bacterial strains formerly known as CDC group M-5 are opportunistic pathogens to humans. In 1993, a name, *Neisseria weaveri*, was proposed by two independent studies to harbor CDC group M-5 strains, namely *N. weaveri* Holmes *et al.* 1993 and *N. weaveri* Andersen *et al.* 1993, with two different 'type' strains. However, no study has been conducted on to the relatedness of the two 'type' strains, although the close relationship of the two taxa has long been accepted unofficially. Formally, the status of the name *N. weaveri* Andersen *et al.* 1993 is illegitimate because it is a later homonym of *N. weaveri* Holmes *et al.*, 1993; but the name of the strain is still validly published. In this study, we attempt to resolve the confusion caused by the apparent duplication of the species *N. weaveri* (with different type strains) using whole genome shotgun sequencing. We also sought to gain insight into the genetic characteristics of *N. weaveri* by conducting comparative genomics. On the basis of genomic similarities revealed through a comparative genomic study, we propose that *N. weaveri* Andersen *et al.* 1993 should be re-classified as a later heterotypic synonym of *N. weaveri* Holmes *et al.*, 1993.

Introduction

The genus *Neisseria* is composed of commensal bacteria that colonize the mucus membranes of mammals. *Neisseria* encompasses two important pathogens – *Neisseria meningitidis* and *Neisseria gonorrhoeae* – as well as many other opportunistic pathogens (Janda & Knapp, 2003; Han *et al.*, 2006). Extensive genomic analyses have been successfully applied to reveal pathogenic mechanisms and vaccine candidates (Sun *et al.*, 2000; Dunning Hotopp *et al.*, 2006; Kawai *et al.*, 2006; Maiden, 2008; Schoen *et al.*, 2008; Aspholm *et al.*, 2010; Marri *et al.*, 2010; Joseph *et al.*, 2011), and as a result, 46 *Neisseria* genome sequences of human origin are available from public databases. However, none of the genomes from strains that originated from animals have been studied.

A group of bacterial strains previously known as CDC group M-5 are part of the normal canine oral flora, but it has known opportunistic pathogenicity in human peritonitis (Kocyigit *et al.*, 2010), lower respiratory tract infections (Panagea *et al.*, 2002), cervical and vaginal infections (Flores-Paz *et al.*, 2003), wound infections (Capitini *et al.*, 2002), and septicemia (Carlson *et al.*, 1997). In 1993, a name, *Neisseria weaveri*, has been proposed by two independent studies to harbor CDC group M-5 strains, namely *N. weaveri* Holmes *et al.* 1993 and *N. weaveri* Andersen *et al.* 1993 (Andersen *et al.*, 1993a, b). The type strain defined by Holmes *et al.* (1993) was isolated from human dog bite wound in Göteborg, Sweden (1974) and that defined by Andersen *et al.* (1993a) was isolated from same type of wound in New York, USA (1962). Even though both taxa were published in the same volume of *International Journal of Systematic Bacteriology* (IJSB), *N. weaveri* Holmes *et al.* 1993 has page precedence over *N. weaveri* Andersen *et al.* 1993 according to Bacteriological Code Rules 51b (4) and 24b (2)

FEMS MICROBIOLOGY LETTERS

(Lagage *et al.*, 1992). Thus, *N. weaveri* Andersen *et al.* 1993 is illegitimate because it is a later homonym of *N. weaveri* Holmes *et al.* 1993 (Euzéby, 1998). Although different type strains were deposited as representing *N. weaveri*, the close relationship of the two taxa has long been accepted because of the similar pathogenic characteristics of the two 'type' strains. However, there has been no systematic investigation about the relatedness of these two 'type' strains and thus the illegitimate name has remained as a homonym and not transferred as a synonym of *N. weaveri* Holmes *et al.* 1993.

Thus, in this study, we attempt to resolve the confusion caused by two *N. weaveri* species with different 'type' strains using whole genome shotgun sequencing. We also sought to gain insight into the genetic characteristics of *N. weaveri* by conducting comparative genomics.

Materials and methods

Bacterial strains and DNA extraction

The 'type' strains of *N. weaveri* Holmes *et al.* 1993 (LMG 5135^T) and *N. weaveri* Andersen *et al.* 1993 (ATCC 51223^T) were obtained from LMG and ATCC, respectively, and the genomic DNAs were extracted using DNeasy Blood & Tissue kit (Qiagen).

Genome sequencing

The draft genome sequences of strains LMG 5135^T and ATCC 51223^T were determined by paired-end shotgun sequencing using the Genome Analyzer IIx (Illumina) with > 1000× fold coverage. The sequencing reads were assembled using the CLC genomics wb4 (CLCbio) and CodonCode Aligner (CodonCode Co.).

Annotation and comparative genomics

Gene finding and annotation were achieved using the RAST server (Aziz *et al.*, 2008). Orthologous gene prediction and comparative genomic analyses were conducted as described previously (Chun *et al.*, 2009). In brief, a segment on target contig, which is homologous to a query open reading frame (ORF), was identified using the BLASTN program. This potentially homologous region was expanded in both directions by 2000 bp. Nucleotide sequences of the query ORF and selection of target homologous region were then aligned using a pairwise global alignment algorithm (Myers & Miller, 1988), and the resultant matched region in the subject contig was extracted and saved as a homolog. Orthologs and paralogs were differentiated by reciprocal comparison.

Phylogenetic tree construction

A set of orthologous ORFs (327 total, 118 543 bp) showing > 70% similar to *N. meningitidis* MC58 (NC_003112) was selected as highly conserved proteins of the genus *Neisseria* and then aligned using the CLUSTALX (Thompson *et al.*, 2002). The resultant multiple alignments were concatenated and then used to construct a genome tree using the neighbor-joining (Saitou & Nei, 1987) method implemented in MEGA program (Kumar *et al.*, 2008). An evolutionary distance matrix for the neighbor-joining tree was generated according to the model of Jukes & Cantor (1969).

Average nucleotide identity calculation

The average nucleotide identity (ANI) was calculated using BLAST as previously described (Goris *et al.*, 2007). In a given pair of genomes, the query genome is spliced into 1020-nt fragments and then blasted against the subject genome. The average of reciprocal results was represented as an ANI value.

Results and discussion

The genome sequences of strains LMG 5135^T and ATCC 51223^T were assembled into 46 and 40 contigs (> 1 kb long), respectively, and deposited into GenBank as accession numbers AFWQ00000000 and AFWR00000000, respectively. Each genome was 2.1 Mb in size (excluding the gaps) and had a G + C content of 49.0%. The genomic contents of the two *N. weaveri* strains were very similar, containing 2233 and 2099 predicted coding sequences (CDSs), respectively.

The genome tree based on the highly conserved orthologous ORFs showed that the two different *N. weaveri* species were closely related, forming a monophyletic clade within the radiation of *Neisseria* (Fig. 1). This phylogenetic closeness of the two species was also supported by the 16S rRNA gene tree (Supporting Information, Fig. S1), in which they have identical 16S rRNA gene sequences. The 16S rRNA gene sequence obtained from the genome sequence was albeit different (3/1488 nt) from the previously known PCR-derived sequence (L10738). The genomic relatedness of the two *N. weaveri* species was calculated by ANI (Konstantinidis & Tiedje, 2005). It is known that 94%–96% of the ANI between a pair of genome sequences may substitute for 70% of the DNA–DNA hybridization value (Konstantinidis & Tiedje, 2005; Goris *et al.*, 2007; Richter & Rossello-Mora, 2009; Auch *et al.*, 2010). The ANI between the two *N. weaveri* strains was 99.1%, clearly indicating that the two strains belong to the same species.

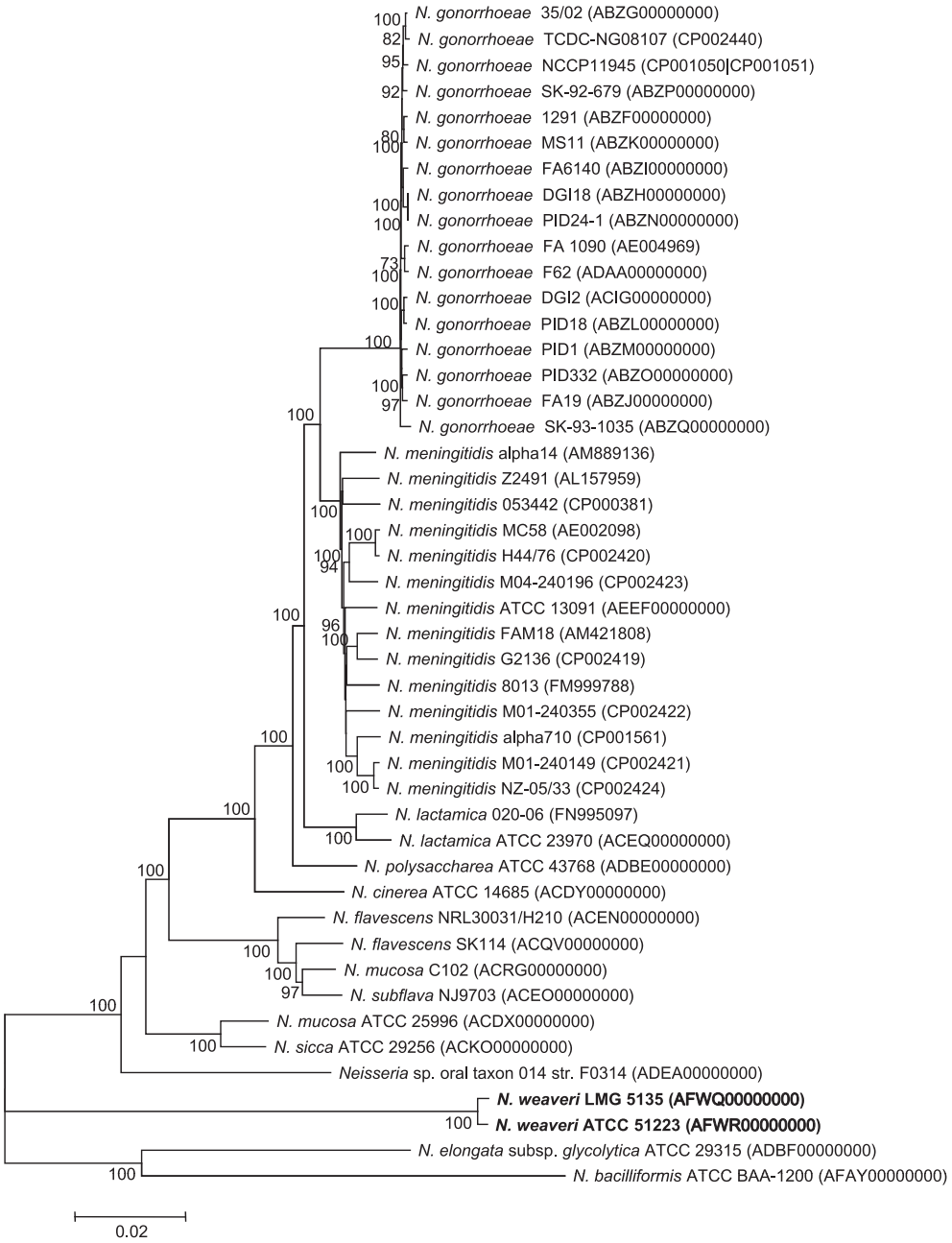

Fig. 1. Genome-based neighbor-joining tree showing the evolutionary relationships among 48 *Neisseria* strains. Each node number represents the percentage of bootstrap support (> 70%) from 1000 resampled datasets. The bar represents 0.02 substitutions per site.

Despite the close relatedness of the two *N. weaveri* strains, they could be distinguished by several genetic elements. Compared with strain ATCC 51223, strain LMG 5135 contains one unique prophage region, one integrative element, and six nonhypothetical genes, but lacks five genes (Table S1). Compared with other *Neisseria* strains, both *N. weaveri* strains contain a unique prophage region, five unique integrative elements, and 21 unique nonhypothetical genes (Table S2).

Many putative virulence genes (Marri *et al.*, 2010) and repeat elements (Parkhill *et al.*, 2000; Snyder & Saunders, 2006; Snyder *et al.*, 2007; Marri *et al.*, 2010) were also detected in *N. weaveri* (Table 1), which are known to play key roles in *Neisseria* virulence and are exchanged *via* genetic transformation, gene expression, and genome rearrangements (Marri *et al.*, 2010; Joseph *et al.*, 2011). The number of DNA uptake sequences [DUS; function in DNA uptake/transformation (Goodman & Scocca, 1988;

Table 1. Repetitive elements and virulence genes in the *Neisseria* genomes

Genome	dRS3	CR	DUS	DUS1	Virulence genes
N. gonorrhoeae (*n* = 17)	138–215	63–232	1897–1967	103–129	111–120
N. meningitidis (*n* = 14)	476–759	417–528	1851–1947	148–180	131–152
N. lactamica (*n* = 2)	203–253	173–174	2213–2245	91–111	111–113
N. polysaccharea (*n* = 1)	152	148	2183	138	106
N. cinerea (*n* = 1)	5	24	1923	148	101
N. flavescens (*n* = 2)	6–9	15–52	2767–2830	167–247	83–84
N. mucosa C102	3	6	2964	155	85
N. subflava (*n* = 1)	156	157	2235	136	82
N. mucosa ATCC 25996	61	313	260	3471	81
N. sicca (*n* = 1)	2	2	2888	173	79
N. sp. oral taxon 014 str. F0314	90	5	3236	467	86
N. weaveri (*n* = 2)	0	0	2841–2878	15	78
N. elongata ssp. *glycolytica* (*n* = 1)	24	106	3273	174	69
N. bacilliformis (*n* = 1)	102	0	4265	140	67

Qvarnstrom & Swedberg, 2006)] and the number of virulence genes were also within the known range of the commensal *Neisseria* genome (Marri *et al.*, 2010; Joseph *et al.*, 2011). The absence of the Opa family [opacity outer membrane proteins for attachment, invasion, immune cell signaling, and inflammation (Dehio *et al.*, 1998; Marri *et al.*, 2010)] and certain iron scavenging genes (Marri *et al.*, 2010) (Table S3) also reflect the genetic characteristics of *N. weaveri* as a member of the commensal *Neisseria*. However, the number of DUS1 was markedly lower in *N. weaveri* compared with other *Neisseria* strains from humans. In contrast to human commensal *Neisseria*, neither the dRS3 element (Parkhill *et al.*, 2000; Bentley *et al.*, 2007) nor Correia elements [CR; (Correia *et al.*, 1986; Snyder *et al.*, 2009)], which function in gene regulation and sequence variation in pathogenic *Neisseria*, were detected in either of the *N. weaveri* genomes (Table 1). Instead, *N. weaveri* strains exclusively contain *vapBC* loci: a type II toxin–antitoxin system (Robson *et al.*, 2009) in which *vapC* encodes a toxin (PilT N-terminus) and *vapB* encodes a matching antitoxin (Cooper *et al.*, 2009). The absence of these loci in other *Neisseria* strains and the homology of these loci to genes in distantly related bacteria suggest that this toxin-related operon was acquired relatively recently *via* horizontal gene transfer. The overall pattern of virulence factors associated with *N. weaveri* suggests that its pathogenicity may differ from other *Neisseria*.

On the basis of the high genomic relatedness (99.1% ANI value) and the identical 16S rRNA gene sequences discovered in this study, we propose that the two *N. weaveri* species should be united as a single species. On the basis of time of publication and established rules of nomenclatural priority (Lagage *et al.*, 1992), we propose to reclassify *N. weaveri* Andersen *et al.* 1993 as a later heterotypic synonym of *N. weaveri* Holmes *et al.*, 1993. The results of this study demonstrate the effectiveness of applying genome-based analyses to microbial taxonomy, including the reclassification of bacterial strains.

Acknowledgements

We thank Dr J.P. Euzéby for his advice on nomenclature. This work was supported by Priority Research Centers Program (#2010-0094020) and a National Research Foundation grant (#2011-0016498) through the National Research Foundation of Korea, funded by the Ministry of Education, Science, and Technology, Republic of Korea.

Statement

The GenBank accession numbers for the genome sequences of strains LMG 5135^T and ATCC 51223^T are AFWQ00000000 and AFWR00000000, respectively.

References

Andersen BM, Steigerwalt AG, O'Connor SP, Hollis DG, Weyant RS, Weaver RE & Brenner DJ (1993a) *Neisseria weaveri* sp. nov., formerly CDC group M-5, a gram-negative bacterium associated with dog bite wounds. *J Clin Microbiol* **31**: 2456–2466.

Andersen BM, Steigerwalt AG, O'Connor SP, Hollis DG, Weyant RS, Weaver RE & Brenner DJ (1993b) *Neisseria weaveri* sp. nov. Validation of the Publication of New Names and New Combinations Previously Effectively Published Outside the IJSB, List no. 47. *Int J Syst Bacteriol* **43**: 864–865.

Aspholm M, Aas FE, Harrison OB, Quinn D, Vik A, Viburiene R, Tonjum T, Moir J, Maiden MC & Koomey M (2010) Structural alterations in a component of cytochrome c oxidase and molecular evolution of pathogenic *Neisseria* in humans. *PLoS Pathog* **6**: e1001055.

Auch AF, von Jan M, Klenk HP & Goker M (2010) Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* **2**: 117–134.

Aziz RK, Bartels D, Best AA *et al.* (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**: 75.

Bentley SD, Vernikos GS, Snyder LA *et al.* (2007) Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18. *PLoS Genet* **3**: e23.

Capitini CM, Herrero IA, Patel R, Ishitani MB & Boyce TG (2002) Wound infection with *Neisseria weaveri* and a novel subspecies of *Pasteurella multocida* in a child who sustained a tiger bite. *Clin Infect Dis* **34**: E74–E76.

Carlson P, Kontiainen S, Anttila P & Eerola E (1997) Septicemia caused by *Neisseria weaveri*. *Clin Infect Dis* **24**: 739.

Chun J, Grim CJ, Hasan NA *et al.* (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *P Natl Acad Sci USA* **106**: 15442–15447.

Cooper CR, Daugherty AJ, Tachdjian S, Blum PH & Kelly RM (2009) Role of vapBC toxin-antitoxin loci in the thermal stress response of *Sulfolobus solfataricus*. *Biochem Soc Trans* **37**: 123–126.

Correia FF, Inouye S & Inouye M (1986) A 26-base-pair repetitive sequence specific for *Neisseria gonorrhoeae* and *Neisseria meningitidis* genomic DNA. *J Bacteriol* **167**: 1009–1015.

Dehio C, Gray-Owen SD & Meyer TF (1998) The role of neisserial Opa proteins in interactions with host cells. *Trends Microbiol* **6**: 489–495.

Dunning Hotopp JC, Grifantini R, Kumar N *et al.* (2006) Comparative genomics of *Neisseria meningitidis*: core genome, islands of horizontal transfer and pathogen-specific genes. *Microbiology* **152**: 3733–3749.

Euzéby JP (1998) List of prokaryotic names with standing in nomenclature. Available at: http://www.bacterio.cict.fr/n/neisseria.html.

Flores-Paz R, Rivera-Sánchez R, García-Jiménez E & Arriaga-Alba M (2003) Etiology of cervical vaginal infection among patients of the Juárez Hospital of Mexico. *Salud Publica Mex* **45**: S694–S697.

Goodman SD & Scocca JJ (1988) Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. *P Natl Acad Sci USA* **85**: 6982–6986.

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P & Tiedje JM (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**: 81–91.

Han XY, Hong T & Falsen E (2006) *Neisseria bacilliformis* sp. nov. isolated from human infections. *J Clin Microbiol* **44**: 474–479.

Holmes B, Costas M, On SL, Vandamme P, Falsen E & Kersters K (1993) *Neisseria weaveri* sp. nov. (formerly CDC group M-5), from dog bite wounds of humans. *Int J Syst Bacteriol* **43**: 687–693.

Janda WM & Knapp JS (2003) *Neisseria* and *Moraxella catarrhalis*. *Manual of Clinical Microbiology*, 8th edn (Murray PR, Baron EJ, Jorgensen JH, Pfaller MA & Yolken RH, eds), pp. 585–608. ASM Press, Washington, DC.

Joseph B, Schwarz RF, Linke B, Blom J, Becker A, Claus H, Goesmann A, Frosch M, Muller T, Vogel U & Schoen C (2011) Virulence evolution of the human pathogen *Neisseria meningitidis* by recombination in the core and accessory genome. *PLoS ONE* **6**: e18441.

Jukes TH & Cantor CR (1969) Evolution of protein molecules. *Mammalian Protein Metabolism*, Vol. 3 (Munro HN, ed), pp. 21–123. Academic Press, New York.

Kawai M, Nakao K, Uchiyama I & Kobayashi I (2006) How genomes rearrange: genome comparison within bacteria *Neisseria* suggests roles for mobile elements in formation of complex genome polymorphisms. *Gene* **383**: 52–63.

Kocyigit I, Unal A, Sipahioglu M, Tokgoz B, Oymak O & Utas C (2010) Peritoneal dialysis-related peritonitis due to *Neisseria weaveri*: the first case report. *Perit Dial Int* **30**: 116–117.

Konstantinidis KT & Tiedje JM (2005) Genomic insights that advance the species definition for prokaryotes. *P Natl Acad Sci USA* **102**: 2567–2572.

Kumar S, Nei M, Dudley J & Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* **9**: 299–306.

Lagage SP, Sneath PHA, Lessel EF, Skerman VBD, Seeliger HPR & Clark WA (eds) (1992) *International Code of Nomenclature of Bacteria*. American Society for Microbiology, Washington, DC.

Maiden MC (2008) Population genomics: diversity and virulence in the *Neisseria*. *Curr Opin Microbiol* **11**: 467–471.

Marri PR, Paniscus M, Weyand NJ, Rendon MA, Calton CM, Hernandez DR, Higashi DL, Sodergren E, Weinstock GM, Rounsley SD & So M (2010) Genome sequencing reveals widespread virulence gene exchange among human *Neisseria* species. *PLoS ONE* **5**: e11835.

Myers EW & Miller W (1988) Optimal alignments in linear space. *Comput Appl Biosci* **4**: 11–17.

Panagea S, Bijoux R, Corkill JE, Al Rashidi F & Hart CA (2002) A case of lower respiratory tract infection caused by *Neisseria weaveri* and review of the literature. *J Infect* **44**: 96–98.

Parkhill J, Achtman M, James KD *et al.* (2000) Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. *Nature* **404**: 502–506.

Qvarnstrom Y & Swedberg G (2006) Variations in gene organization and DNA uptake signal sequence in the folP region between commensal and pathogenic *Neisseria* species. *BMC Microbiol* **6**: 11.

Richter M & Rossello-Mora R (2009) Shifting the genomic gold standard for the prokaryotic species definition. *P Natl Acad Sci USA* **106**: 19126–19131.

Robson J, McKenzie JL, Cursons R, Cook GM & Arcus VL (2009) The vapBC operon from *Mycobacterium smegmatis* is

an autoregulated toxin-antitoxin module that controls growth via inhibition of translation. *J Mol Biol* **390**: 353–367.

Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.

Schoen C, Blom J, Claus H *et al.* (2008) Whole-genome comparison of disease and carriage strains provides insights into virulence evolution in *Neisseria meningitidis*. *P Natl Acad Sci USA* **105**: 3473–3478.

Snyder LA & Saunders NJ (2006) The majority of genes in the pathogenic *Neisseria* species are present in non-pathogenic *Neisseria lactamica*, including those designated as 'virulence genes'. *BMC Genomics* **7**: 128.

Snyder LA, Cole JA & Pallen MJ (2009) Comparative analysis of two *Neisseria gonorrhoeae* genome sequences reveals evidence of mobilization of Correia Repeat Enclosed Elements and their role in regulation. *BMC Genomics* **10**: 70.

Snyder LA, McGowan S, Rogers M, Duro E, O'Farrell E & Saunders NJ (2007) The repertoire of minimal mobile elements in the *Neisseria* species and evidence that these are involved in horizontal gene transfer in other bacteria. *Mol Biol Evol* **24**: 2802–2815.

Sun YH, Bakshi S, Chalmers R & Tang CM (2000) Functional genomics of *Neisseria meningitidis* pathogenesis. *Nat Med* **6**: 1269–1273.

Thompson JD, Gibson TJ & Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics*, Chapter 2, Unit 2.3.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. 16S rRNA gene-based neighbor-joining tree showing the evolutionary relationships among 48 *Neisseria* strains.

Table S1. Genetic elements that distinguish strain LMG 5135 from strain ATCC 51223.

Table S2. Genetic elements that distinguish the *N. weaveri* genome from other *Neisseria* genomes.

Table S3. Composition of the iron utilization system in the *Neisseria* genomes.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.