Isolation and characterization of a new species of *Neisseria*, *Neisseria musculi*, from the wild house mouse

Running Title: *Neisseria musculi* sp. nov. from the wild house mouse

Nathan J. Weyand[1,2], Mancheong Ma[1], Megan Phifer-Rixey[3], Nyiawung A. Taku[1,4], María A. Rendón[1], Alyson M. Hockenberry[1], Won J. Kim[1], Al B. Agellon[5], Nicolas Biais[6], Taichi A. Suzuki[3], Lily G. Sait[7], Odile B. Harrison[7], Holly B. Bratcher[7], Michael W. Nachman[3], Martin C. J. Maiden[7] and Magdalene So[1]

[1]Department of Immunobiology and BIO5 Institute, University of Arizona, Tucson, AZ 85719, USA; [2]Present address: Department of Biological Sciences, Ohio University, Athens, OH 45701, USA; [3]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA; present address: Museum of vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, Berkeley, CA 94720, USA; [4]Present address: Shiley School of Engineering, University of Portland, Portland, OR 97203 USA; [5]School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, AZ 85721, USA; [6]Department of Biology, Brooklyn College of the City University of New York, Brooklyn, NY 11210, USA; [7]Department of Zoology, University of Oxford, Oxford OX1 3PS, UK.

Correspondance: Nathan J. Weyand; weyand@ohio.edu; 740-593-0191

Keywords: *Neisseria*, *Mus musculus domesticus*, genetic competence, DNA uptake sequence, Type IV pilus.

Subject category: New taxa - Proteobacteria

Word count: abstract, 162; main text, 3,558.

Abbreviations: DUS, DNA uptake sequence; cfu, colony forming units; gDNA, genomic DNA; MLST, multilocus sequence typing; RifR, rifampicin resistant, SEM, scanning electron microscopy, SmR, streptomycin resistant; Tfp, Type IV pili.

The GenBank accession numbers for the partial *rplF* gene sequences of strains AP2031^T, AP2104, AP2105 and AP2119 are KT997727, KT997728, KT997729 and KT997730, respectively. The European Nucleotide Archive run identifier for the AP2031 draft genome is ERR1121178. AP2031 genome data was also deposited in PubMLST databases and can be found at: www.pubmlst.org/neisseria/ or http://pubmlst.org/mlst/ (ID 29520).

Abstract

Neisseria have been isolated from or detected in a wide range of animals, from nonhuman primates and felids to a rodent, the guinea pig (Liu et al., 2015). By means of selective culture, biochemical testing, Gram staining and PCR screening for the *Neisseria*-specific Internal Transcribed Spacer region of the rRNA operon, we isolated four strains of *Neisseria* from the oral cavity of the wild house mouse, *Mus musculus* subspecies *domesticus.* The isolates are highly related and form a separate clade in the genus, as judged by rMLST and core gene tree analyses. One isolate, provisionally named *Neisseria musculi* sp. nov. (type strain AP2031^T = DSM 101846^T = CCUG 68283^T = LMG 29261^T), was studied further. AP2031/*N. musculi* grows well *in vitro*. It is naturally competent, taking up DNA in a DUS and *pilT*-dependent manner, and is amenable to genetic manipulation. These and other genomic attributes of *N. musculi* make it an ideal candidate for use in developing a mouse model for studying *Neisseria*-host interactions.

The *Neisseriaceae*, a large family of Gram-negative bacteria, are most well known for two members that cause diseases of importance to human health, *Neisseria meningitidis* and *Neisseria gonorrhoeae*. The other members of the family are commensals of humans and other animals. *N. meningitidis* and *N. gonorrhoeae* also exhibit commensal-like behavior in that they establish asymptomatic infections in man at high frequency (Caugant and Maiden, 2009, Read, 2014, Turner et al., 2002). This behavior may be due to traits they inherited from their commensal forebears as they evolved a more pathogenic lifestyle (Marri et al., 2010).

At least ten species of commensal *Neisseria* are known to colonize humans, and an equal number or more have been isolated from, or detected in, a wide range of animals (Bennett et al., 2014a, Liu et al., 2015). Commensal *Neisseria* have been isolated from the oropharynx and urogenital tract of man, the oral cavity and throat of guinea pigs, the liver and feces of ducks, the oral cavity and dental plaque of cows, the upper respiratory tract and lung of dogs, and the oral cavity of a rhesus macaques (see references (see Table 2 in Liu et al., 2015, Barrett and Sneath, 1994, Bennett et al., 2014a, Murphy et al., 2005, Weyand et al., 2013)).

Many aspects of *Neisseria*-host interactions can be more readily dissected using a small animal model, for example, by pairing the genetically tractable mouse, *Mus musculus*, with a species of *Neisseria* that naturally colonizes the animal. We therefore sought to isolate a mouse-dwelling *Neisseria* strain. Through selective culturing, biochemical testing and genotyping, we obtained four *Neisseria* isolates, representing a new species in the genus, from the oral cavity of wild mice (*Mus musculus* subspecies *domesticus*) trapped in two geographically distinct regions in North America. One isolate, AP2031, was chosen for further study. AP2031 grows well *in vitro*. It is naturally competent; DNA transformation is greatly enhanced by the presence of the neisserial DNA Uptake Sequence (DUS) and by an intact *pilT* gene. We propose that AP2031 be classified as *Neisseria musculi* sp. nov. (type strain AP2031).

Oral swabs were collected from 36 healthy wild mice (*M. m. domesticus*) live-trapped in two distinct regions of North America (Table S1). A subset of these mice (TAS216-TAS220) was sampled immediately for *Neisseria*, then humanely euthanized for use in another study. All other mice were brought back to the lab and housed individually or in breeding pairs in static microisolator cages, with food and water *ad libitum*. Cages were changed weekly or biweekly in a laminar flow hood following standard protocols to minimize cross-cage contamination. After initial breeding, all animals were housed individually or in same-sex pairs for at least six weeks to ensure eradication of mouse pathogens that pose a risk to animal care facilities. Serologically negative mice were made available for *Neisseria* sampling. All collection and husbandry activities were conducted in accordance with University of Arizona IACUC policies.

Mice that had passed the quarantine period were sampled for *Neisseria*. The oral cavity was swabbed using the BD^{TM} BBL^{TM} CultureSwab Plus Transport System (Fisher Scientific). To enrich for growth of *Neisseria* swab suspensions in GCB medium base (Beckton Dickinson) were plated on GCB agar containing Vancomycin (2 mg/L) and Trimethoprim (3 mg/L), and the plates were incubated for 48 hours at 37^{o}C, 5% CO_2.

To identify and speciate *Neisseria* we proceeded as follows. A portion from each colony growing on GCB vancomycin/trimethoprim agar was used for oxidase testing as described (Cheesbrough, 2006) using oxidase reagent (PML Microbiologicals). Oxidase^{+} colonies were Gram stained, and the internal transcribed spacer (ITS) region in the rRNA operon of Gram^{-} colonies was amplified by colony PCR, using primers specific to sequences that are highly conserved among *Neisseria* species (see Table S2 for primers). ITS^{+} colonies were streak-purified, assigned a strain number, and stored at -80^{o}C in GCB + glycerol (20% wt/vol). We identified four isolates that were oxidase^{+}, Gram^{-}, and *Neisseria* ITS^{+}. Two of these isolates, AP2031 and AP2119, were from mice trapped in Tucson, Arizona, USA (TUSA8 and TAS218, respectively), and two isolates, AP2104 and AP2105, were from mice trapped in a haystack near Parkland, Alberta, Canada (TAS116 and TAS118, respectively)(Table S1). To further characterize these four isolates, their *rplF* sequences were determined and compared to those of other *Neisseria* species. A neighbor-joining tree was drawn with *rplF* sequences using MEGA5 (Tamura et al., 2011). All three codon positions were included. Bootstrap

support values (2,000 replicates) greater than 70% were added to nodes. Sequence comparisons of *rplF*, an approach that distinguishes closely related species in the *Neisseria* genus (Bennett et al., 2014b), indicated the four isolates were tightly clustered in a separate clade (Fig. 1a). Surprisingly, they were not as related to *Neisseria animalis*, a guinea pig isolate, as to *Neisseria dentiae*, a cow dental plaque isolate (Sneath and Barrett, 1996, Berger, 1960).

We conducted ANI analysis for all four mouse isolates using BioEdit 7.2.5 (Hall, 1999). Only three polymorphic sites were detected in the *rplF* fragment sequenced. Since the four isolates share 99.4% sequence identity we conclude they belong to the same species, *N. musculi* sp. nov. This conclusion is based on the fact that *rplF* gene trees are more concordant with rMLST trees, the gold standard for *Neisseria* species identification, whereas 16S rRNA trees lack such resolution (Bennett et al., 2014b).

The genome of isolate AP2031, from Tucson mouse TUSA8, was sequenced at the Oxford Genomic Center, Wellcome Trust Centre for Human Genetics, University of Oxford, United Kingdom. Samples were quantified using PicoGreen (Invitrogen) and their integrity was assessed using 1% E-Gel EX (Invitrogen). DNA was fragmented using the Episonic system and sized using a Tapestation D1200 system (Agilent/Lab901). Libraries were constructed with NEBNext DNA Sample Prep Master Mix Set 1 Kit (New England BioLabs), and adaptors were ligated using Illumina Adapters (Multiplexing Sample Preparation Oligonucleotide Kit). Ligated libraries were size-selected using AMPure magnetic beads (Agencourt), PCR enriched for 10 cycles as per Illumina recommendations, and purified with AMPure XP beads. Size distribution was determined using a Tapestation 1DK system (Agilent/Lab901), and multiplex pool concentrations were determined by PicoGreen (Invitrogen). Pooled libraries were quantified using the quantitative PCR Library Quantification Kit and MX3005PTM instrument (Agilent). Sequencing was performed using the Illumina HiSeq 2000 system to generate 100 nucleotide paired-end reads. Resulting short-read sequences were assembled *de novo* using the VelvetOptimiser algorithm as part of an in-house pipeline developed at the Department of Zoology, University of Oxford. The minimum output contig size was set to 200bp with the scaffolding option switched off; all other program settings were left at default. No read trimming was performed. Draft genome sequences

(ID 29520) were deposited in the *Neisseria* pubmlst.org database (www.pubmlst.org/neisseria). The DNA G+C content of AP2031 was 53.3 mol% consistent with other species in the genus (Tonjum, 2005).

A detailed analysis of AP2031 genome content will be presented in a future paper. Genomic data from AP2031 was compared with a representative isolate dataset containing all known *Neisseria* species for which WGS data were available (Bennett et al., 2012). Multi-locus sequence typing (MLST) of 51 ribosomal genes (rMLST)(Fig. 1b) and 246 *Neisseria* core genes (cgMLST)(Fig. S1) was used to generate trees as described previously enabling relationships between isolates to be determined and speciation (Bratcher et al., 2014, Jolley et al., 2012). Sequences were extracted from the bacterial isolate genome sequence database (BIGSdb) (Jolley and Maiden, 2010), aligned using MAFFT and imported into Splitstree thereby generating NeighborNet trees. Nucleotide sequences used for rMLST and cgMLST are listed in Table S3. These analyses confirmed the *rplF* findings on the evolutionary relationship of AP2031 to other species of *Neisseria.*

rplF, rMLST and cgMLST phylogenetic analysis reveal AP2031 to be most closely related to *N. dentiae*. The *rplF* ANI between AP2031 and *N. dentiae* is 85.1%. Other members of the genus, as expected, have lower ANI values (See Table S4). The ANI calculator at http://enve-omics.ce.gatech.edu/ani/ was used to estimate the ANI between the AP2031 and *N. dentiae* draft genome data sets. This tool calculates ANI as described previously (Goris et al., 2007). Consistent with the *rplF* analysis, an ANI value of 86.8% was obtained for the draft genomes of AP2031 and *N. dentiae*. Given that ANI values lower than 95% delineate the genomes of different species (Goris et al., 2007), our findings further indicate AP2031 and the other mouse *Neisseria* isolates comprise a new species in the genus *Neisseria*.

During passage on agar plates, *Neisseria musculi* sp. nov. (type strain AP2031) gave rise to colonies with smooth and rough phenotypes (Fig. S2). Images of colony morphology were acquired using a Cannon SLR digital camera attached to a Wild M7 zoom microscope (Heerbrugg, Switzerland) mounted on a Bausch and Lomb stage. All colonies of AP2031 were circular and convex, and varied in opacity. Smooth colonies

had shiny, unwrinkled surfaces, while rough colonies had bumpy, undulating surfaces. These phenotypes were interchangeable: smooth colony variants would give rise to rough colony variants and *vice versa* (data not shown). This process is reminiscent of colony phase variation observed for *N. gonorrhoeae*, which is caused by variable expression of surface components (Bhat et al., 1991, Long et al., 1998, Swanson and Barrera, 1983, Swanson et al., 1971, Walstad et al., 1977). The mechanism underlying AP2031 colony phase variation is under investigation.

Smooth and rough variants formed biofilms when cultured on glass, and Scanning Electron Microscopy revealed their distinct morphologies. Scanning Electron Microscopy was performed on AP2031 cells as described previously (Higashi et al., 2011). Biofilms formed by the rough variant were taller and contained multiple layers of cells (Fig. 2). Rough variant cells are best described as diplococcobacilli: short rods often found in pairs. The rods were approximately 0.4 microns in width and 0.5-0.8 microns in length (Fig. 2a, b, e, f). The shape of smooth variant cells was difficult to discern. Many appeared to be encased in sheaths of varying length with occasional indentations along their sides (Fig. 2c, d, g, h, i; Fig. S3). Several species of commensal *Neisseria* are rod-shaped (e.g., *Neisseria elongata*, *Neisseria weaveri*, *Neisseria shayeganii, Neisseria bacilliformis, Neisseria zoodegmatis, Neisseria animaloris*) (Ganiere et al., 1995, Liu et al., 2015, Veyrier et al., 2015), but there are no reports, to our knowledge, of sheath-like structures associated with *Neisseria* cells. The surfaces of smooth and rough variants were decorated with fibers, with rough variant fibers being shorter and less abundant. The nature of these fibers remains to be determined.

All species of *Neisseria* studied to date are competent for DNA transformation. Playing important roles in this process are the Type IV pilus (Tfp) and the DNA Uptake Sequence (DUS), a 10-nucleotide sequence, 5' -GCCGTCTGAA- 3', that is highly enriched in neisserial genomes. Tfp-mediated DNA transformation involves, among other things, the binding of the DUS to Tfp-associated protein ComP, and PilT-driven retraction of the Tfp (Cehovin et al., 2013, Hamilton and Dillard, 2006, Wolfgang et al., 1998). Two DUS variants, distinguished by two bases immediately preceding the 10-mer, are abundant in neisserial genomes: 5' -agGCCGTCTGAA- 3' and 5' -atGCCGTCTGAA- 3'. agDUS is most abundant in human commensals *N. elongata, N.*

bacilliformis and *Neisseria subflava*, and dog commensal *N. weaveri* (Frye et al., 2013, Mell and Redfield, 2014), while atDUS is most abundant in human commensals *Neisseria cinerea, Neisseria lactamica* and *Neisseria polysaccharea*, and human pathogens *N. gonorrhoeae* and *N. meningitidis*.

To determine the number of copies of DUS in the AP2031 genome, and to extract the DUS and surrounding sequences into a fasta file for constructing the WebLogo, Perl regular expression pattern matching in combination with the fuzznuc program from EMBOSS were used (Crooks et al., 2004, Rice et al., 2000). There are 3893 copies of the 10-mer DUS in AP2031, of which 2614 are agDUS and 119 are atDUS (Fig. S4). AP2031 has a complete set of genes for the biogenesis of the Tfp including *pilT,* which encodes the motor protein for the Tfp fiber (Table S5). The abundance of the neisserial DUS and the presence of Tfp genes strongly suggest that AP2031 is naturally competent for DNA transformation.

To test the competence of AP2031, we determined the transformation frequency of the smooth variant in the presence of genomic DNA from AP2098, a naturally occurring streptomycin resistant (Sm^R) variant. AP2098 was isolated by plating the original AP2031 strain on GCB plates containing streptomycin (100 mg/L). Genomic DNA from AP2098 was used for transformation assays using previous described procedures (Weyand et al., 2013). Transformation mixtures were plated on GCB agar with and without streptomycin (100 mg/L). Transformation frequency was expressed as the number of Sm^R colony forming units (cfu)/total cfu/μg of DNA. For some negative controls, DNA was incubated for 20 minutes with two units of DNaseI prior to transformation (New England Biolabs). The frequency was 1.7×10^{-4} (SEM $\pm$ 4.8×10^{-5}) in the presence of DNA, and >3 logs lower in its absence or when the DNA was predigested with DNAse (Table 1).

To determine the role of the DUS in AP2031 transformation, cells were incubated with synthetic DNA amplified by PCR and encoding *rpsL* from AP2098 (Sm^R), with or without the 10-mer DUS 5' -GCCGTCTGAA- 3' (See Table S2 for primers). DNA (1 μg) was used for transformation assays, and transformation frequencies were calculated as

described above. The transformation frequency was nearly 3 logs higher when the transforming DNA contained the DUS (Table 1).

Finally, we determined the role of *pilT* in AP2031 competence. Primers NP242F and NP242R (Table S2) containing flanking sequences for the AP2031 *pilT* gene were used to amplify a kanamycin cassette from plasmid pNBNeiKan (synthesized by Genescript) for use in constructing gene deletions (Datsenko and Wanner, 2000). A DUS and an *Escherichia coli* σ70-dependent consensus promoter upstream of the cassette promote transformation and expression of the kanamycin resistance gene, respectively. The length of the *pilT* flanking sequences in the amplicon was extended by PCR using primers NP252F and NP252R. The amplification products were purified and used in spot or liquid transformations (Dillard, 2006) of AP2031 to replace the *pilT* gene with a kanamycin resistance marker. Primers NP245F and NP245R were used to amplify the mutated region (Δ*pilT*::*kan*) and the amplification product was cloned into pGEMT (Promega). The resulting plasmid DNA was used for transformations aimed at replacing the *pilT* locus with Δ*pilT*::*kan* in a rifampicin resistant strain background (see AP2365, Table 1). Transformants were selected and maintained on GCB containing kanamycin (50 mg/L). *pilT* mutants were confirmed by Sanger sequencing following PCR with primers NP245F and NP245R. The Δ*pilT* mutant transformed at a frequency approximately 3 logs lower than its wt parent (Table 1). Taken together, these results indicate AP2031 is naturally competent, and DNA uptake involves the neisserial DUS and Tfp.

By means of selective culture, oxidase testing, Gram staining and PCR testing for the *Neisseria* ITS, we isolated four strains of *Neisseria* from the oral cavity of healthy wild mice (*Mus musculus* subspecies *domesticus*) trapped in Tucson, Arizona, USA, and Parkland, Alberta Province, Canada. *rplF* sequence comparisons indicate they are very closely related to each other and form a distinct clade in the *Neisseria* genus. Their closest neighbor is *Neisseria dentiae*, a bovine isolate (Fig. 1)(Sneath and Barrett, 1996). This analysis is confirmed by trees of the Tucson mouse isolate, AP2031, based on MLST (Fig. 1b, Fig. S1). We propose that the mouse *Neisseria* strains be classified as *Neisseria musculi* sp. nov. (type strain AP2031). The rMLST approach

unambiguously identifies members of the *Neisseria* genus (Bennett et al., 2012). Our rMLST and core gene analyses indicate AP2031/*N. musculi* is a member of the *Neisseria* genus. This assignment is supported by the DUS sequence analysis of AP2031/*N. musculi* (Fig. S4).

AP2031/*N. musculi* is naturally competent; DNA transformation is greatly enhanced by the DUS and *pilT,* which encodes the Tfp motor protein (Table S5). Tfp is a hallmark of the *Neisseria*. Tfp consists of a fiber, composed mainly of pilin subunits, that extends from the cell into the extracellular milieu, and machinery for its assembly and many biological activities (Carbonnelle et al., 2006, Giltner et al., 2012). In addition to DNA tranformation (Cehovin et al., 2013, Wolfgang et al., 1998), Tfp plays important roles in other aspects of *Neisseria* biology, including bacterial attachment, motility, microcolony formation and signaling to the host cell (Eriksson et al., 2012, Helaine et al., 2005, Higashi et al., 2007, Howie et al., 2005, Kurre et al., 2012, Lee et al., 2005, Merz et al., 1999, Merz et al., 1996, Merz et al., 2000, Nassif et al., 1994, Pujol et al., 1999, Winther-Larsen et al., 2001, Wolfgang et al., 1998). That AP2031/*N. musculi* harbors the genes necessary for the biosynthesis, assembly and anchoring of the Tfp fiber; for the retraction and mechanotransductive properties of the fiber; and for DNA transformation (Table S5) strongly suggest that it expresses Tfp that functions similarly to those of other species in the genus.

Pilins in Tfp of pathogenic *Neisseria* are decorated with mono-, di- or tri-saccharides, and/or phosphorylated residues (Aas et al., 2006, Forest et al., 1999, Hartley et al., 2011, Naessan et al., 2008). These post-translational modifications are thought to contribute to Tfp bundling, neisserial interactions with human cells and protection of the bacteria from immune defenses (Chamot-Rooke et al., 2011, Jennings et al., 2011, Marceau et al., 1998). AP2031/*N. musculi* encodes enzymes for the biosynthesis and covalent linkage of the basal monosaccharide to pilin (*pglB, pglB2, pglC* and *pglD*), but lacks the genes that add additional sugars to the monosaccharide (*pglA* and *pglE*). This suggests its pilin may be decorated with monosaccharides. It should be noted that pilins in *N. elongata* Tfp are not glycosylated even though the commensal harbors *pgl* genes (Anonsen et al., 2015). Whether this lack of glycosylation is due to control of gene

expression is unclear. Finally, AP2031/*N. musculi* lacks the genes (*pptA* and *pptB*) that phosphorylate pilin (Anonsen et al., 2012, Naessan et al., 2008).

In summary, we have isolated a new species in the *Neisseria* genus from the oral cavity of wild mice. We determined the ability of one of these isolates, Type strain AP2031, to take up neisserial DNA, and sequenced its genome. AP2031 had the following characteristics that support its inclusion in the genus *Neisseria*. First, it stains Gram-negative. Second, it has oxidase activity. Third, genus specific primers successfully amplified the 16S internal transcribed spacer region. Fourth, *rplF* and MLST phylogenic analysis determined that AP2031 belongs in the genus and is most closely related to *N. dentiae*, an established *Neisseria* species. Fifth, AP2031 genome sequencing established that this isolate has over 3000 copies of the best characterized *Neisseria*-specific repetitive element, the DNA uptake sequence.

AP2031/*N. musculi* grows readily in the lab and is amenable to genetic manipulation. These qualities of AP2031, coupled with its genetic relatedness to other *Neisseria* species, make the isolate an excellent candidate for use in developing a mouse model for probing *Neisseria*-host interactions. Such a model will circumvent roadblocks imposed by host tropism, an issue that makes studying human-dwelling *Neisseria* problematical. It will allow assessment of the role of the Type IV pilus in *Neisseria* colonization, and dissection of immune responses mounted by the host (the mouse) to the commensal *Neisseria*.

Description of *Neisseria musculi* sp. nov.

Neisseria musculi (mus'cu.li. L. gen. n. *musculi* of a mouse).

Cells are diplococcobacilli (approx. 0.5 to 0.8 mm in length) Gram-negative and oxidase-positive. Good growth occurs on chocolate agar, TSA with 5% sheep blood and GCB agar. Colonies are small, circular, convex and vary between colony morphologies with margins that are entire (smooth colonies) or undulate (rough colonies). Colonies vary in opacity and are 0.5 to 1 mm in diameter after 48 h of growth in 5% CO_2 at 37 °C on GCB agar. The DNA uptake sequence, GCCGTCTGAA, enriched in the type strain's

339 genome confirms membership in the genus. The type strain has a DNA G+C content of
340 53.3 mol%.
341
342 Type strain is AP2031^{T} (=DSM 101846^{T} =CCUG 68283^{T} =LMG 29261^{T}) was isolated
343 from an oral swab of *Mus musculus* subspecies *domesticus* in Tucson USA.
344

345 Acknowledgements
346 We wish to thank M. Goytia for assistance with DNA preparations and D. Yoder (BIO5
347 Institute Media Facility) for providing excellent media services. We thank J. Bray for
348 database deposition of the AP2031 Sequence Read Archive. This work was supported
349 by grants from the University of Arizona College of Medicine and the BIO5 Institute to
350 M.S., and NIH grant RO1 GM074245 to M.W.N.
351

Figure Legends

Fig. 1(a) Evolutionary relationship of mouse *Neisseria* isolates AP2031 (in grey box), AP2119, AP2104, and AP2105, to 50 isolates of *Neisseria*, deduced from a Neighbor-joining tree constructed from their *rplF* sequences. Type strains are indicated with "-T". Bootstrap values ≥70% are noted (2,000 replications). Bar: 0.06 substitutions per nucleotide position. PubMLST ID or GenBank accession numbers proceed the isolate and species notations. Taxa included for *rplF* comparisons were previously analyzed by Bennett and colleagues (Bennett et al., 2014b). (b) Neighbor-net tree reconstructed from 51 concatenated ribosomal gene sequences, with AP2031 highlighted in grey box.

Fig. 2. Scanning Electron Microscopy images of AP2031 shown as side (a, b, c, d) and aerial (e, f, g, h, i) views. Images of rough (a, b, e, f) and smooth (c, d, g, h, i) colony variants after 3 hours of growth on glass slides are shown. Boxed sections of images are shown at higher magnification in adjoining panels (b, d, f, h, i). White arrows in panel (i) indicate indentations along sheaths. Acquisition voltage, distance (μm) and scale bar are listed.

Supplementary Fig. S1. Evolutionary relationships of *Neisseria* species, with AP2031 highlighted in grey box, as shown in a Neighbor-Net splits tree constructed using concatenated sequences of 246 *Neisseria* core genes. See Materials and Methods for tree construction.

Supplementary Fig. S2. AP2031 smooth (a) and rough (b) colony variants on 48-hour old plates. Scale bar: 0.5 mm.

Supplementary Fig. S3. Scanning Electron Microscopy images of AP2031 smooth colony variant after 3 hours of growth on glass slides. Boxed section of image is shown at higher magnification in adjoining panel. Images of long filaments in the smooth variant. Acquisition voltage, distance (μm) and scale bar are listed.

Supplementary Fig. S4. Comparison of the most abundant DNA Uptake Sequence (DUS) of mouse *Neisseria* isolate AP2031 with DUSs prevalent in other *Neisseria*

species. (a) Sequence logo of the most abundant DUS in AP2031 (3893 copies). Two nucleotide positions on both sides of the DUS 10-mer are also shown. Nucleotide positions of the 10-mer are indicated. (b) The most abundant 12-mer agDUS in commensals such as *N. elongata* (Nel) subsp. *glycolytica* ATCC 29315*, N. bacilliformis* (Nba) ATCC BAA120, *N. weaveri* (Nwe) ATCC 51223 and *N. subflava* (Nsu) NJ9703. (c) The most abundant 12-mer atDUS in commensals such as *N. cinerea* (Nci) ATCC 14685, *N. lactamica* (Nla) 020-06*, N. polysaccharea* (Npo) ATCC 43768, and pathogens *N. gonorrhoeae* (Ngo) FA1090 *and N. meningitidis* (Nme) MC58*.*

References

Aas, F. E., Egge-Jacobsen, W., Winther-Larsen, H. C., Lovold, C., Hitchen, P. G., Dell, A. & Koomey, M. (2006). *Neisseria gonorrhoeae* type IV pili undergo multisite, hierarchical modifications with phosphoethanolamine and phosphocholine requiring an enzyme structurally related to lipopolysaccharide phosphoethanolamine transferases. *J Biol Chem* **281**, 27712-27723.

Anonsen, J. H., Egge-Jacobsen, W., Aas, F. E., Borud, B., Koomey, M. & Vik, A. (2012). Novel protein substrates of the phospho-form modification system in *Neisseria gonorrhoeae* and their connection to O-linked protein glycosylation. *Infection and immunity* **80**, 22-30.

Anonsen, J. H., Vik, A., Borud, B., Viburiene, R., Aas, F. E., Kidd, S. W., Aspholm, M. & Koomey, M. (2015). Characterization of a unique tetrasaccharide and distinct glycoproteome in the O-linked protein glycosylation system of *Neisseria elongata* subspecies *glycolytica*. *Journal of bacteriology*.

Barrett, S. J. & Sneath, P. H. (1994). A numerical phenotypic taxonomic study of the genus *Neisseria*. *Microbiology* **140 (Pt 10)**, 2867-2891.

Bennett, J. S., Bratcher, H. B., Brehony, C., Harrison, O. B. & Maiden, M. C. (2014a). The Genus *Neisseria*. In The Prokaryotes - *Alphaproteobacteria and Betaproteobacteria*, pp. 881-900: Springer-Verlag.

Bennett, J. S., Jolley, K. A., Earle, S. G., Corton, C., Bentley, S. D., Parkhill, J. & Maiden, M. C. (2012). A genomic approach to bacterial taxonomy: an examination and proposed reclassification of species within the genus *Neisseria*. *Microbiology* **158**, 1570-1580.

Bennett, J. S., Jolley, K. A. & Maiden, M. C. (2013). Genome sequence analyses show that *Neisseria oralis* is the same species as *'Neisseria mucosa* var. *heidelbergensis'*. *International journal of systematic and evolutionary microbiology* **63**, 3920-3926.

Bennett, J. S., Watkins, E. R., Jolley, K. A., Harrison, O. B. & Maiden, M. C. (2014b). Identifying *Neisseria* species by use of the 50S ribosomal protein L6 (*rplF*) gene. *Journal of clinical microbiology* **52**, 1375-1381.

Berger, U. (1960). *Neisseria animalis* nov. spec. . *Z Hyg Infektionskr* **147**, 158-161.

Bhat, K. S., Gibbs, C. P., Barrera, O., Morrison, S. G., Jahnig, F., Stern, A., Kupsch, E. M., Meyer, T. F. & Swanson, J. (1991). The opacity proteins of *Neisseria gonorrhoeae* strain MS11 are encoded by a family of 11 complete genes. *Molecular microbiology* **5**, 1889-1901.

Bratcher, H. B., Corton, C., Jolley, K. A., Parkhill, J. & Maiden, M. C. (2014). A gene-by-gene population genomics platform: de novo assembly, annotation and genealogical analysis of 108 representative *Neisseria meningitidis* genomes. *BMC Genomics* **15**, 1138.

Carbonnelle, E., Helaine, S., Nassif, X. & Pelicic, V. (2006). A systematic genetic analysis in *Neisseria meningitidis* defines the Pil proteins required for assembly, functionality, stabilization and export of type IV pili. *Molecular microbiology* **61**, 1510-1522.

Caugant, D. A. & Maiden, M. C. (2009). Meningococcal carriage and disease--population biology and evolution. *Vaccine* **27 Suppl 2**, B64-70.

Cehovin, A., Simpson, P. J., McDowell, M. A., Brown, D. R., Noschese, R., Pallett, M., Brady, J., Baldwin, G. S., Lea, S. M. & other authors (2013). Specific DNA recognition mediated by a type IV pilin. *Proc Natl Acad Sci U S A* **110**, 3065-3070.

Chamot-Rooke, J., Mikaty, G., Malosse, C., Soyer, M., dumont, A., Gault, J., Imhaus, A. F., Martin, P., Trellet, M. & other authors (2011). Posttranslational modification of pili upon cell contact triggers *N. meningitidis* dissemination. *Science* **331**, 778-782.

Cheesbrough, M. (2006). Oxidase Test. In District Laboratory Practice in Tropical Countries Part 2, pp. 69-70. New York: Cambridge University Press.

Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190.

Datsenko, K. A. & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6640-6645.

Dillard, J. P. (2006). Genetic manipulation of *Neisseria gonorrhoeae*. In Current Protocols in Microbiology.

Eriksson, J., Eriksson, O. S. & Jonsson, A. B. (2012). Loss of meningococcal PilU delays microcolony formation and attenuates virulence in vivo. *Infection and immunity* **80**, 2538-2547.

Forest, K. T., Dunham, S. A., Koomey, M. & Tainer, J. A. (1999). Crystallographic structure reveals phosphorylated pilin from *Neisseria*: phosphoserine sites modify type IV pilus surface chemistry and fibre morphology. *Molecular microbiology* **31**, 743-752.

Frye, S. A., Nilsen, M., Tonjum, T. & Ambur, O. H. (2013). Dialects of the DNA uptake sequence in *Neisseriaceae*. *PLoS genetics* **9**, e1003458.

Ganiere, J. P., Escande, F., Andre-Fontaine, G., Larrat, M. & Filloneau, C. (1995). Characterization of group EF-4 bacteria from the oral cavity of dogs. *Veterinary microbiology* **44**, 1-9.

Giltner, C. L., Nguyen, Y. & Burrows, L. L. (2012). Type IV pilin proteins: versatile molecular modules. *Microbiol Mol Biol Rev* **76**, 740-772.

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P. & Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International journal of systematic and evolutionary microbiology* **57**, 81-91.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In Nucleic Acids Symposium Series, pp. 95-98.

Hamilton, H. L. & Dillard, J. P. (2006). Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. *Molecular microbiology* **59**, 376-385.

Hartley, M. D., Morrison, M. J., Aas, F. E., Borud, B., Koomey, M. & Imperiali, B. (2011). Biochemical characterization of the O-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of protein glycans containing N,N'-diacetylbacillosamine. *Biochemistry* **50**, 4936-4948.

Helaine, S., Carbonnelle, E., Prouvensier, L., Beretti, J. L., Nassif, X. & Pelicic, V. (2005). PilX, a pilus-associated protein essential for bacterial aggregation, is a key to pilus-facilitated attachment of *Neisseria meningitidis* to human cells. *Molecular microbiology* **55**, 65-77.

Higashi, D. L., Biais, N., Weyand, N. J., Agellon, A., Sisko, J. L., Brown, L. M. & So, M. (2011). *N. elongata* produces type IV pili that mediate interspecies gene transfer with *N. gonorrhoeae*. *PloS one* **6**, e21373.

Higashi, D. L., Lee, S. W., Snyder, A., Weyand, N. J., Bakke, A. & So, M. (2007). Dynamics of *Neisseria gonorrhoeae* attachment: microcolony development, cortical plaque formation, and cytoprotection. *Infection and immunity* **75**, 4743-4753.

Howie, H. L., Glogauer, M. & So, M. (2005). The *N. gonorrhoeae* type IV pilus stimulates mechanosensitive pathways and cytoprotection through a pilT-dependent mechanism. *PLoS Biol* **3**, e100.
Jennings, M. P., Jen, F. E., Roddam, L. F., Apicella, M. A. & Edwards, J. L. (2011). *Neisseria gonorrhoeae* pilin glycan contributes to CR3 activation during challenge of primary cervical epithelial cells. *Cell Microbiol* **13**, 885-896.
Jolley, K. A., Bliss, C. M., Bennett, J. S., Bratcher, H. B., Brehony, C., Colles, F. M., Wimalarathna, H., Harrison, O. B., Sheppard, S. K. & other authors (2012). Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. *Microbiology* **158**, 1005-1015.
Jolley, K. A. & Maiden, M. C. (2010). BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11**, 595.
Kurre, R., Hone, A., Clausen, M., Meel, C. & Maier, B. (2012). PilT2 enhances the speed of gonococcal type IV pilus retraction and of twitching motility. *Molecular microbiology* **86**, 857-865.
Lee, S. W., Higashi, D. L., Snyder, A., Merz, A. J., Potter, L. & So, M. (2005). PilT is required for PI(3,4,5)P3-mediated crosstalk between *Neisseria gonorrhoeae* and epithelial cells. *Cell Microbiol* **7**, 1271-1284.
Liu, G., Tang, C. M. & Exley, R. M. (2015). Non-pathogenic *Neisseria*: members of an abundant, multi-habitat, diverse genus. *Microbiology* **161**, 1297-1312.
Long, C. D., Madraswala, R. N. & Seifert, H. S. (1998). Comparisons between colony phase variation of *Neisseria gonorrhoeae* FA1090 and pilus, pilin, and S-pilin expression. *Infection and immunity* **66**, 1918-1927.
Marceau, M., Forest, K., Beretti, J. L., Tainer, J. & Nassif, X. (1998). Consequences of the loss of O-linked glycosylation of meningococcal type IV pilin on piliation and pilus-mediated adhesion. *Molecular microbiology* **27**, 705-715.
Marri, P. R., Paniscus, M., Weyand, N. J., Rendon, M. A., Calton, C. M., Hernandez, D. R., Higashi, D. L., Sodergren, E., Weinstock, G. M. & other authors (2010). Genome sequencing reveals widespread virulence gene exchange among human *Neisseria* species. *PloS One* **5**, e11835.
Mell, J. C. & Redfield, R. J. (2014). Natural competence and the evolution of DNA uptake specificity. *Journal of bacteriology* **196**, 1471-1483.
Merz, A. J., Enns, C. A. & So, M. (1999). Type IV pili of pathogenic *Neisseriae* elicit cortical plaque formation in epithelial cells. *Molecular microbiology* **32**, 1316-1332.
Merz, A. J., Rifenbery, D. B., Arvidson, C. G. & So, M. (1996). Traversal of a polarized epithelium by pathogenic *Neisseriae*: facilitation by type IV pili and maintenance of epithelial barrier function. *Mol Med* **2**, 745-754.
Merz, A. J., So, M. & Sheetz, M. P. (2000). Pilus retraction powers bacterial twitching motility. *Nature* **407**, 98-102.
Murphy, J., Devane, M. L., Robson, B. & Gilpin, B. J. (2005). Genotypic characterization of bacteria cultured from duck faeces. *Journal of applied microbiology* **99**, 301-309.
Naessan, C. L., Egge-Jacobsen, W., Heiniger, R. W., Wolfgang, M. C., Aas, F. E., Rohr, A., Winther-Larsen, H. C. & Koomey, M. (2008). Genetic and functional analyses of PptA, a phospho-form transferase targeting type IV pili in *Neisseria gonorrhoeae*. *Journal of bacteriology* **190**, 387-400.
Nassif, X., Beretti, J. L., Lowy, J., Stenberg, P., O'Gaora, P., Pfeifer, J., Normark, S. & So, M. (1994). Roles of pilin and PilC in adhesion of *Neisseria meningitidis* to human

epithelial and endothelial cells. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 3769-3773.
Pujol, C., Eugene, E., Marceau, M. & Nassif, X. (1999). The meningococcal PilT protein is required for induction of intimate attachment to epithelial cells following pilus-mediated adhesion. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 4017-4022.
Read, R. C. (2014). *Neisseria meningitidis*; clones, carriage, and disease. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **20**, 391-395.
Rice, P., Longden, I. & Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276-277.
Sneath, P. H. & Barrett, S. J. (1996). A new species of *Neisseria* from the dental plaque of the domestic cow, *Neisseria dentiae* sp. nov. *Letters in applied microbiology* **23**, 355-358.
Swanson, J. & Barrera, O. (1983). Gonococcal pilus subunit size heterogeneity correlates with transitions in colony piliation phenotype, not with changes in colony opacity. *The Journal of experimental medicine* **158**, 1459-1472.
Swanson, J., Kraus, S. J. & Gotschlich, E. C. (1971). Studies on gonococcus infection. I. Pili and zones of adhesion: their relation to gonococcal growth patterns. *The Journal of experimental medicine* **134**, 886-906.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution* **28**, 2731-2739.
Tonjum, T. (2005). Genus I. *Neisseria* Trevisan 1885, 105AL. In Bergey's Manual of Systematic Bacteriology, pp. 777-798. Edited by D. J. Brenner, N. R. Krieg, J. T. Staley and G. M. Garrity. New York: Springer.
Turner, C. F., Rogers, S. M., Miller, H. G., Miller, W. C., Gribble, J. N., Chromy, J. R., Leone, P. A., Cooley, P. C., Quinn, T. C. & Zenilman, J. M. (2002). Untreated gonococcal and chlamydial infection in a probability sample of adults. *Jama* **287**, 726-733.
Veyrier, F. J., Biais, N., Morales, P., Belkacem, N., Guihen, C., Ranjeva, S., Sismeiro, O., Pehau-Arnaudet, G., Rocha, E. P. & other authors (2015). Common Cell Shape Evolution of Two Nasopharyngeal Pathogens. *PLoS Genet* **11**, e1005338.
Walstad, D. L., Guymon, L. F. & Sparling, P. F. (1977). Altered outer membrane protein in different colonial types of *Neisseria gonorrhoeae*. *Journal of bacteriology* **129**, 1623-1627.
Weyand, N. J., Wertheimer, A. M., Hobbs, T. R., Sisko, J. L., Taku, N. A., Gregston, L. D., Clary, S., Higashi, D. L., Biais, N. & other authors (2013). *Neisseria* infection of rhesus macaques as a model to study colonization, transmission, persistence, and horizontal gene transfer. *Proc Natl Acad Sci U S A* **110**, 3059-3064.
Winther-Larsen, H. C., Hegge, F. T., Wolfgang, M., Hayes, S. F., van Putten, J. P. & Koomey, M. (2001). *Neisseria gonorrhoeae* PilV, a type IV pilus-associated protein essential to human epithelial cell adherence. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 15276-15281.
Wolfgang, M., Lauer, P., Park, H. S., Brossay, L., Hebert, J. & Koomey, M. (1998). PilT mutations lead to simultaneous defects in competence for natural transformation and twitching motility in piliated *Neisseria gonorrhoeae*. *Molecular microbiology* **29**, 321-330.

Table 1. Transformation frequency of mouse *Neisseria* strains.

Recipient strain	DNA	Transformation frequency* (x 10^{-4})
AP2031	AP2031-SmR gDNA	1.7 ± 0.48
AP2031	AP2031-SmR gDNA + DNAse	< 0.00024 ± 0.000071
AP2031	No DNA	< 0.00034 ± 0.00013
AP2031	*rpsL*^{SmR} + DUS	0.43 ± 0.4
AP2031	*rpsL*^{SmR}	0.0016 ± 0.00079
AP2031	No DNA	< 0.00016 ± 0.000031
AP2365 (AP2031-RifR)	AP2031-SmR gDNA	1.249 ± 0.55
AP2365 Δ*pilT*::Kan	AP2031-SmR gDNA	< 0.0015 ± 0.00037
AP2365	No DNA	< 0.00091 ± 0.00023
AP2365 Δ*pilT*::Kan	No DNA	< 0.0019 ± 0.00043

*Transformation frequencies are expresses as the number of streptomycin resistant (SmR) CFU/total CFU. Values are averaged from three independent experiments ± SEM. "<": transformation frequency below the limit of detection. gDNA: genomic DNA; *rpsL*: PCR amplicon of *rpsL* locus with SmR mutation; DUS: DNA uptake sequence GCCGTCTGAA; RifR, rifampicin resistant.

590
591