Opportunity and Means: Horizontal Gene Transfer from the Human Host to a Bacterial Pathogen

Mark T. Anderson and H. Steven Seifert

Updated information and services can be found at:
http://mbio.asm.org/content/2/1/e00005-11.full.html

SUPPLEMENTAL MATERIAL

http://mbio.asm.org/content/2/1/e00005-11.full.html#SUPPLEMENTAL

REFERENCES

This article cites 25 articles, 13 of which can be accessed free at:
http://mbio.asm.org/content/2/1/e00005-11.full.html#ref-list-1

CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more>>
Opportunity and Means: Horizontal Gene Transfer from the Human Host to a Bacterial Pathogen

Mark T. Anderson and H. Steven Seifert
Department of Microbiology-Immunology, Northwestern University, Feinberg School of Medicine, Chicago, Illinois, USA

ABSTRACT The acquisition and incorporation of genetic material between nonmating species, or horizontal gene transfer (HGT), has been frequently described for phylogenetically related organisms, but far less evidence exists for HGT between highly divergent organisms. Here we report the identification and characterization of a horizontally transferred fragment of the human long interspersed nuclear element L1 to the genome of the strictly human pathogen Neisseria gonorrhoeae. A 685-bp sequence exhibiting 98 to 100% identity to copies of the human L1 element was identified adjacent to the irg4 gene in some N. gonorrhoeae genomes. The L1 fragment was observed in ~11% of the N. gonorrhoeae population sampled but was not detected in Neisseria meningitidis or commensal Neisseria isolates. In addition, N. gonorrhoeae transcripts containing the L1 sequence were detected by reverse transcription-PCR, indicating that an L1-derived gene product may be produced. The high degree of identity between human and gonococcal L1 sequences, together with the absence of L1 sequences from related Neisseria species, indicates that this HGT event occurred relatively recently in evolutionary history. The identification of L1 sequences in N. gonorrhoeae demonstrates that HGT can occur between a mammalian host and a resident bacterium, which has important implications for the co-evolution of both humans and their associated microorganisms.

IMPORTANCE The interactions between microbes and their hosts are relevant to several aspects of biology, including evolution, development, immunity, and disease. Neisseria gonorrhoeae serves as a particularly informative model for this interaction because it has exclusively coevolved with humans and is not known to be found in any other environment. In addition, investigation of the evolutionary relationship between N. gonorrhoeae and humans has practical implications, since gonorrhea is a prevalent sexually transmitted infection worldwide. This study was undertaken to characterize the horizontal transfer of genetic information from humans to N. gonorrhoeae, an event that has been scarcely recognized between any mammalian host and bacterial pathogen. Here we provide evidence that this genetic exchange was the result of a recent evolutionary event that has been propagated within the gonococcal population.

Evolution of asexual organisms can occur by the accumulation and vertical transmission of mutations in an existing genetic repertoire or through the acquisition of novel genetic determinants by horizontal gene transfer (HGT). HGT has the capacity to facilitate the acquisition of complex physiological functions (1) in a single molecular event and has been estimated to account for up to 17% of the total genetic material found in Escherichia coli (2). HGT must therefore be regarded as an important force in the adaptation of species. HGT is particularly well known to have a diverse impact on bacterial physiology, with examples in the literature describing the horizontal acquisition of novel biosynthetic (3), virulence (1), and antibiotic resistance (4) functions by diverse bacterial species. Despite the countless interactions between commensal or pathogenic bacteria and their cognate hosts and the ever-increasing amount of available genomic sequence data, examples of bacterial integration of host genetic information are exceedingly rare. The rarity of HGT from host to bacterium may be due, in part, to potential barriers such as restriction and modification functions, the availability of genetic material, and the fitness cost of deleterious HGT events. However, such an event has the potential to impact the evolution of both the host and the microbe.

Neisseria gonorrhoeae is the causative agent of the sexually transmitted infection gonorrhea. N. gonorrhoeae is naturally competent for transformation and has incorporated DNA from related neisserial species (5, 6) and more divergent bacterial species (7) into its genome. Recombination of exogenous DNA occurs efficiently in N. gonorrhoeae, and it is postulated that this process plays an important role in the microevolution of the organism (8, 9). N. gonorrhoeae has a long and exclusive evolutionary history with its human host, and infection is typically localized to the urogenital tract. Therefore, sources of foreign DNA available to the gonococcus are limited to coinhabiting microorganisms or the host itself. Mammalian genomes contain numerous copies of the long interspersed nuclear element (LINE) L1. Full-length human L1 elements encode a nucleic acid binding protein and a multido-
RESULTS

The genome sequences of 14 *N. gonorrhoeae* clinical isolates were recently determined and made publically available by the Neisseria gonorrhoeae group Sequencing Project (http://www.broadinstitute.org/), adding to the existing genome sequences of FA1090 and NCCP11945 (12). Analysis of the unfinished assembled identifications of a 685-bp fragment present in strains FA6410 (supercontig 29, positions 23818 to 24502), DG18 (supercontig 24, positions 24037 to 24721), and PID24 (supercontig 27, positions 23854 to 24538) that exhibited 98 to 100% identity to copies of human retrotransposable element L1. This Neisseria L1 sequence (nL1) also exhibits strong similarity to LINEs from other primates (e.g., *Pan troglodytes*), but since humans are the only known natural host for *N. gonorrhoeae*, it is likely that the HGT event that yielded nL1 occurred in humans. The nL1 fragment is identical among the FA6410, DG18, and PID24 Neisseria strains and occupies the same insertion site (Fig. 1). The nL1 element includes the coding sequence for the first 164 amino acids of the L1 open reading frame 1 (ORF1) nucleic acid binding protein (13) and 192 bases of the L1 5′ untranslated region upstream of ORF1. This region of the ORF1 protein contains a coiled-coil domain that is required for trimerization of the murine ORF1 protein and the first 8 amino acids of a 96-residue RNA recognition motif (14). The insertion site in the gonococcal genome is located 14 bp upstream of a copy of the phage-associated transposase *irg* (15, 16). This genetic region also contains a 24-bp inverted repeat, one of which is interrupted by the nL1 insertion (Fig. 1). Although repeat sequences and mobile genetic elements are often associated with HGT, comparison of nL1-positive and nL1-negative genomes yields no evidence that either the putative NF4-G4 prophage or the inverted repeat was directly involved in the acquisition of nL1.

Contamination of DNA sequencing samples has previously accounted for the false identification of LINE sequences in human-associated microorganisms (17). To test whether the presence of nL1 was the result of contamination introduced during the sequencing or purification of gonococcal DNA, independently prepared genomic DNA from the sequenced *N. gonorrhoeae* isolates was assayed for the presence of nL1 by PCR. Multiple reactions containing primer combinations specific to the nL1 sequence and flanking DNA verified the presence of nL1 in the genomic location predicted by the sequences (Fig. 2A to D). Sequencing of nL1-containing PCR products confirmed this result and the composition of nL1 as reported in the genome sequences of all three strains (data not shown). DNA hybridizations using a probe generated from human L1 sequences also confirmed that strains FA6140, DG18, and PID24 contained nL1 and that this fragment was not located at an alternative site in the genomes of the remaining 12 strains from Fig. 2 (data not shown). Together, these analyses conclusively demonstrate that the presence of L1 DNA in strains PID24, DG18, and FA6140 is attributable to HGT and is not an artifact of genome sequencing.

To assess the frequency of nL1 in the population, a diverse collection of *Neisseria* strains, isolated at different times and locations, were screened for the presence of nL1 using a combination of molecular and bioinformatic techniques (see Table S1 in the supplemental material). Four additional nL1-positive *N. gonorrhoeae* strains were identified by this analysis (97G181, ATCC 49226, PID334, and WHO I) and determined to harbor a 685-bp sequence identical to that reported for the sequenced nL1-positive strains. Therefore, the nL1 allele occurred at a frequency of ~11% among the 62 *N. gonorrhoeae* strains that were included in this study. Interestingly, the different disease manifestations of *N. gonorrhoeae* infection (uncomplicated infection, pelvic inflammatory disease, disseminated gonococcal infection) were all represented among the nL1-positive strains identified here. Although closely related members of the genus Neisseria are also exclusively associated with human infection, examination of 212 *Neisseria meningitidis* isolates and 19 commensal *Neisseria* isolates produced no evidence of the nL1 fragment (see Table S1 in the supplemental material). Assuming that the frequency of occurrence is similar in *N. meningitidis* and *N. gonorrhoeae*, the probability of observing zero nL1 events in the *N. meningitidis* population tested is 1.9E−11, since the number should follow a binomial distribution with *n* = 212 and *P* = 0.11. The lack of nL1 sequences in the highly genetically related species *N. meningitidis* suggests that the acquisition of nL1 by *N. gonorrhoeae* was a recent evolutionary event that occurred subsequent to the divergence of *N. meningitidis* and *N. gonorrhoeae*. The identification of nL1 in gonococci but not meningococci may reflect differences in the species-specific interactions between bacteria and host cells, the availability of host
DNA during infection with *N. gonorrhoeae* versus *N. meningitidis*, or most likely an event that occurred after the two species diverged.

The absolute conservation of the nL1 sequence among the seven positive isolates is striking and suggests that this locus is under selective pressure or has not had sufficient time to degenerate. Reverse transcription (RT)-PCR using total RNA and internal L1 primers detected transcripts containing the nL1 sequence originating from both strands of DNA (Fig. 2E and F). Therefore, the HGT event that yielded nL1 has the potential to produce a novel gene product and may contribute to the observed sequence conservation. However, because the frequency of nL1 occurrence in the tested population is modest, any selective force conferred by nL1 may be small and difficult to identify experimentally. During laboratory culture, no consistent gross phenotypic differences between nL1-positive and nL1-negative strains were observed.

In addition to selective pressure, clonal expansion could also account for the observed lack of variation of the nL1 allele among *N. gonorrhoeae* isolates. Multilocus sequencing typing (MLST) (18) was used to assess the relatedness of four nL1-positive isolates compared to 12 confirmed L1-negative isolates for which allele sequences were available. Strains DGI18 and PID24 have identical MLST profiles belonging to sequence type 8418 and may represent clonal expansion of the nL1 allele (Table 1). However, strains FA6140 and PID334 have sequence types divergent from both sequence type 8418 and each other. Since three distinct sequence types are represented among the four nL1 strains analyzed, either the presence of the nL1 fragment in these strains was the result of multiple identical HGT events or, more likely, an initial acquisition of nL1 from humans was subsequently transferred horizontally between gonococcal strains.

**DISCUSSION**

The high level of identity between human L1 and nL1 sequences, combined with its absence in closely related *Neisseria* species, suggests that this HGT event happened relatively recently in evolutionary history and/or that the nL1 region is under strong selective pressure. In addition, the low penetrance of nL1 within the gonococcal population is consistent with the hypothesis that nL1 was acquired recently from humans. Since numerous different L1 copies exist in any one human genome, including ones that contain sequences identical to nL1 (GenBank accession no. AC013546), analysis based on the accumulation of synonymous and nonsynonymous substitutions relative to a donor sequence is not feasible. *N. gonorrhoeae* can be found intracellularly and extracellularly associated with neutrophils and epithelial cells and thus has the opportunity to encounter host DNA during infection (19, 20). The natural competence of *N. gonorrhoeae* provides the bacterium with the means to acquire exogenous DNA, and the fact that the signature *Neisseria* DNA uptake sequence is not absolutely required for internalization of exogenous DNA (21) increases the probability of this event’s occurrence. Although *N. gonorrhoeae* can invade and replicate within host cells, there is no evidence for nuclear association, making interaction with host DNA in live cells unlikely. A potential scenario in which HGT between humans and intracellular *N. gonorrhoeae* could more readily occur is after an apoptotic (22) or necrotic event in which the host DNA is

![FIG 2](https://www.mbio.asm.org/content/2/1/e00005-11.full.html)
fragmented or disrupted within the dying cell. Alternatively, the active secretion of neutrophil extracellular traps (23) provides a means by which host DNA could become accessible to extracel- lular \textit{N. gonorrhoeae}.

The mechanism by which this foreign genetic material was incorporated into the \textit{N. gonorrhoeae} genome is a matter of spec- ulation. Comparison of the insertion site between nL1-positive and nL1-negative isolates demonstrates that the recombined fragment is the result of a simple insertion with no loss or gain of flanking chromosomal information (Fig. 1). Mobilization of L1 elements within the nuclei of host cells occurs via target-primed RT (11), which requires the function of ORF1p and ORF2p and results in target site duplication. No such signature duplication is present near the nL1 site, and retrotransposition is unlikely to account for the insertion of nL1 into the \textit{N. gonorrhoeae} genome. There is no homology between the gonococcal insertion site and the L1 sequences flanking the nL1 fragment, suggesting that direct recombination after transformation is also an unlikely mecha- nism for the integration of nL1. Therefore, we postulate that non- homologous end joining (NHEJ) could have occurred between a fragmented copy of L1 and the recipient region of the \textit{N. gonor- rhoeae} genome. Although \textit{N. gonorrhoeae} lacks the NHEJ machin- ery possessed by some bacteria, alternative pathways of low- frequency NHEJ that have the potential to mediate HGT of exogenous sequences have been proposed (24).

Regardless of the mechanism by which HGT occurred, it is possible that nL1 has an effect on the physiology of \textit{N. gonorrhoeae} by way of residual activity of the truncated ORF1 protein (if pro- duced) or via polar effects of the insertion itself. More signifi- cantly, though, the identification of human sequences within the \textit{N. gonorrhoeae} genome has important implications for modeling of the generation of genetic diversity and evolution of this and other obligate human pathogens and establishes that HGT can occur between a mammalian host and an associated microorgan- ism.

\section*{MATERIALS AND METHODS}

\subsection*{Bacterial strains and culture conditions.}

The \textit{Neisseria} isolates used in this study and their sources are listed in Table S1 in the supplemental material. \textit{N. gonorrhoeae} and commensal \textit{Neisseria} species were grown on gonococcal medium base (Difco) containing Kellogg supplements as de- scribed previously (25). \textit{N. meningitidis} strains were cultivated on gono- coccal medium supplemented with 1\% IsoVitaleX (Becton Dickinson). Bacteria used for the preparation of RNA were grown to exponential phase as described previously (25) in GCRL broth (1.5\% proteose peptone no. 3 [Difco], 0.4\% K$_2$HPO$_4$, 0.1\% KH$_2$PO$_4$, 0.1\% NaCl) supplemented with 1\% IsoVitaleX.

\subsection*{Detection of the nL1 allele.}

All of the oligonucleotide primers used in this study are listed in Table S2 in the supplemental material. Genomic DNA was purified on QIAamp columns (Qiagen) and used for PCR and slot blot hybridizations. The nL1 genotype of the \textit{N. gonorrhoeae} strains listed in Fig. 2 was established using the following primer combinations: IS1106for and IRGrev, LINEfor and IRGrev, IS1106for and LINErev, and LINEfor and IRGrev. All other \textit{Neisseria} isolates screened by PCR as reported in Table S1 in the supplemental material were screened by the LINEfor and LINErev primer pair at minimum. PCR products were se- quenced by the Northwestern University Genomics Core Facility. A sam- pling of isolates of each \textit{Neisseria} species screened was analyzed by slot blot hybridization using a 541-bp probe PCR amplified from human chromo- somal DNA using primers LINEfor and LINErev. The product was labeled with digoxigenin-dUTP via the random-primed method (Roche), and hybridizations were performed using standard techniques. The binomial distribution was used to determine the probability of observing zero nL1 events in \textit{N. meningitidis} based on the \textit{N. gonorrhoeae} frequency and as- suming equivalent rates. The Northwestern University Biostatistics Col- laboration Center assisted with the statistical analysis.

\subsection*{RT-PCR.}

Total RNA was purified from \textit{N. gonorrhoeae} cultures using the RNeasy kit (Qiagen) and treated with RQI DNase (Promega) to re- move genomic contamination. The LINEfor and LINErev oligonucleo- tidese were used to prime cDNA synthesis using Superscript III reverse transcriptase (Invitrogen) according to the manufacturer’s recommenda- tions. Internal nL1 primers L1for and L1rev were used to PCR amplify cDNA products.

\subsection*{MLST.}

Seven housekeeping genes; \textit{abcZ}, \textit{add}, \textit{aroE}, \textit{fumC}, \textit{gdh}, \textit{pdhC}, and \textit{pyk}, were used to establish the sequence types of \textit{N. gonorrhoeae} strains as described previously (18), with the exception of the oligonu- cleotide primers that were used for amplification and sequencing of MLST alleles. Allele numbers were assigned by querying the Neisseria MLST database (18) using existing genomic sequences when available. The MLST primers listed in Table S2 in the supplemental material were used to PCR amplify and sequence (both strands) all alleles from strain PID334 and any other allele not previously established in the database. The pro- files of all of the strains listed in Table 1 have been deposited in the Neis- seria MLST database.

\section*{ACKNOWLEDGMENTS}

We acknowledge M. Apicella, J. Dillard, K. Lee, R. Nicholas, P. Rice, W. Shafer, D. Trees, and W. Whittington for contributing \textit{Neisseria} strains; J. Bodily for providing preparations of human genomic DNA; K. Clark for performing the bioinformatics screen that identified the nL1 sequence; D. Ward and M. Giovanni for facilitating the genomic sequencing of \textit{N. gonorr- hoeae} strains; M. Fitzgerald for providing genomic assemblies; M. Maiden for providing access to \textit{N. meningitidis} genome sequences; K.-Y. Kim for assisting with the statistical analysis; and A. Chen and A. Hauser for critical reading of the manuscript.

This work was supported by NIH grants R01 AI044239 and R37 AI033493 to H.S.S. M.T.A. was partially supported by NIH fellowship F32AI080083.

\section*{SUPPLEMENTAL MATERIAL}

Supplemental material for this article may be found at http://mbio.asm.org

Table S1, PDF file, 0.189 MB.

Table S2, PDF file, 0.042 MB.

\section*{REFERENCES}


8. Fudyk, T. C., I. W. Maclean, J. N. Simonsen, E. N. Njagi, J. Kimani,


