Fusobacterium nucleatum is a prominent member of the oral microbiota and is a common cause of human infection. *F. nucleatum* includes five subspecies: *polymorphum*, *nucleatum*, *vincentii*, *fusiforme*, and *animalis*. *F. nucleatum* subsp. *polymorphum* has been well characterized phenotypically and, in contrast to previously sequenced strains, is amenable to gene transfer. We sequenced and annotated the 2,429,698 bp genome of *F. nucleatum* subsp. *polymorphum* ATCC 10953. Plasmid pFN3 from the strain was also sequenced and analyzed. When compared to the other two available fusobacterial genomes (*F. nucleatum* subsp. *nucleatum*, and *F. nucleatum* subsp. *vincentii*) 627 open reading frames unique to *F. nucleatum* subsp. *polymorphum* ATCC 10953 were identified. A large percentage of these mapped within one of 28 regions or islands containing five or more genes. Seventeen percent of the clustered proteins that demonstrated similarity were most similar to proteins from the clostridia, with others being most similar to proteins from other gram-positive organisms such as Bacillus and Streptococcus. A ten kilobase region homologous to the Salmonella typhimurium propanediol utilization locus was identified, as was a prophage and integrated conjugal plasmid. The genome contains five composite ribozyme/transposons, similar to the CdISt IStrons described in Clostridium difficile. IStrons are not present in the other fusobacterial genomes. These findings indicate that *F. nucleatum* subsp. *polymorphum* is proficient at horizontal gene transfer and that exchange with the Firmicutes, particularly the Clostridia, is common.
and host immune cells [30–32], its ability to modulate host immune cell function including the induction of apoptosis [30,33–35], its ability to enhance the survival of strict anaerobes in biofilm and planktonic multispecies cultures [3,36], and its ability act synergistically with other oral pathogens to enhance virulence in animal model systems [37,38]. In addition, ATCC 10953 harbors a single native plasmid, pFN3 [39], and has been shown to be amenable to gene transfer [40].

The distinct taxonomic status, relatively extensive phenotypic analyses, and genetic transformability of \textit{F. nucleatum} ssp \textit{polymerum} ATCC 10953 suggested that genomic analysis would prove valuable to subsequent studies of this species. Thus, we determined, analyzed, and present the genomic DNA sequence of \textit{F. nucleatum} ssp \textit{polymerum} strain ATCC 10953. Analysis of the FNP genome revealed that 25% of the genes identified are not represented in the previously sequenced fusobacterial genomes, and that horizontal gene transfer (HGT) has contributed to the evolution of this strain.

**RESULTS AND DISCUSSION**

**Genome anatomy**

The FNP genome consists of a single circular chromosome containing 2,429,698 bp (Figure 1, accession number CM000440) and a single circular plasmid of 11,934 bp (Figure 2, accession number CP111710). The FNP genome is larger than the FNN genome (2,174,499 bp), and the unfinished FNV genome (2,118,259 bp). The GC content of the chromosome is 26.84%, similar to FNN and FNV (Table 1) and the GC content for the plasmid is 24.53%. There are 2,433 predicted ORFs, 42 similar to FNN and FNV (Table 1) and the GC content for the (2,118,259 bp). The GC content of the chromosome is 26.84%, similar to FNN and FNV (Table 1) and the GC content for the plasmid is 24.53%. There are 2,433 predicted ORFs, 42 pseudogenes, 45 tRNA, 15 rRNA, and 11 ncRNA genes in the FNP genome.

Forty-two of the tRNAs lie in one of seven clusters, each containing from two to fourteen tRNAs. Identical clusters, in both content and internal gene order, are found in FNN. However, the relative position of the clusters in the genome is different. In addition, FNN has two additional asparagine tRNAs, which are associated with tRNA gene regions that have not been fully sequenced in ATCC 10953 so it is possible that these tRNAs are present in FNP. The tRNAs represent all twenty amino acids. In five cases (Ala, Cys, His, Trp, Tyr), there is only a single tRNA. Many of the tRNAs are duplicated (Asn, Glu, Gly, Leu, Pro, Lys, Ser, Thr, Val) or triplicated (Asp, Gln, Ala), the clusters include duplicates as well as singleton tRNAs with no clear pattern associated with the distribution of singletons, duplicates and triplicates. Some tRNAs have multiple copies only one of which is unique, e.g. Gly, Lys, Ser, Val. All four Arg tRNAs, the three Met tRNAs (one is identified as the likely initiator) and the two Phe tRNAs have unique sequences.

A 131 bp repeat sequence occurs seventeen times within the genome (Figure 3). In several cases, the repeat regions have the potential to encode small hypothetical peptides, which we believe were incorrectly annotated as ORFs in FNN and FNV. All of the 131 nt. repeats from FNP were aligned and used to generate a consensus sequence for additional BLASTN searches. Five nearly complete repeats (≥90% of the element’s length) and eight repeats with gaps were observed. Sequences corresponding to the 3’ half of the element were also found in six locations. In most cases, the repeat sequence was found in intergenic regions and not within coding sequence. One complete copy of the repeat and numerous subsequences were identified in FNN and twenty-eight complete copies plus numerous subsequences were identified in FNV. As in FNP, these fell within intergenic regions. Because the long repeats occur within intergenic regions, it is possible that the sequence is involved in gene regulation, though no particular regulatory motifs were found within the sequence.

Many examples of conservation of gene order, operon structure and gene clustering were observed in the genome. Most of the ribosomal protein genes are organized into operons similar to the L11, L10, S10, spc and x operons in \textit{Escherichia coli}; each encodes between two and eleven ribosomal proteins [41]. Several non-r-protein clusters, similar to conserved non-r-protein gene clusters described in \textit{E.coli} and other bacteria [42] were also identified in the FNP genome. Table 2 shows the conservation of these gene regions between FNP, FN, \textit{Clostridium difficile} (NC_009089), \textit{Bacillus anthracis} str. Ames (NC_003997), \textit{E.coli} MG1655 (NC_000913). The conserved gene clusters were categorized in five groups according to protein function. Clusters 1–11 in group I are primarily RNA and protein constituents of the ribosome. Group II encodes subunits of the F-type two-sector ATPase. Proteins encoded by the group III clusters are involved in RNA synthesis, modification, transcription, and translation. Cluster 17 in group IV, encodes a spermidine/putrescine ATP binding cassette (ABC) superfamily transporter, while group V contains cluster 18, which codes for the molecular chaperones GroEL and GroES.
Genome Comparisons

Slightly more than 62% of the FNP genes (1514) are found in both of the previously sequenced fusobacterial genomes. Thus, nearly 38% of the genome is either wholly unique to FNP or is shared by FNP and only one of the other two genomes. In terms of coding potential, these 919 genes account for the differences between FNP, FNN, and FNV and thus serve to distinguish FNP. Three comparative lists have been generated: ORFs unique to FNP with respect to FNN and FNV; ORFs common to FNP and FNV, but absent in FNN; and ORFs common to FNP and FNN, but absent in FNV (Table S1).

627 FNP ORFs have no ortholog in either FNN or FNV (Table S1a), including 106 conserved hypothetical proteins, 287 hypothetical proteins and 9 pseudogenes. Thirty-eight ORFs functioning in transport, including transporters of amino acids, oligopeptides, a siderophore, and divalent metals (Hg\textsuperscript{2+}, Cu\textsuperscript{2+}, Ni\textsuperscript{2+}) are

![Figure 2. Plasmid pFN3.](image)

**Figure 2. Plasmid pFN3.** a) Map of plasmid pFN3. Replication and recombination ORFs are shown in blue and hypotheticals are colored green. b) Alignments of fusobacterial relaxase protein domains to mobilization class consensus motifs [66]. Consensus sequence abbreviations: uppercase letters, conserved; lowercase letters, present in 50% of sites; U or u, bulky hydrophobic residues (I, L, V, M, F, Y and W); *, no consensus at this site; Y, putative active site tyrosine residue. Asterisks (*) above residues indicate identity with consensus sequence. Alignments were performed using Clustal W [102] and then adjusted to best fit the consensus.

doi:10.1371/journal.pone.0000659.g002

**Table 1. General genome statistics for FNP, FNN and FNV.**

| Genome | Length (bp) | %GC content | % Coding | ORFs | Proteins | rRNA | tRNA | ncRNA |
|--------|-------------|-------------|----------|------|----------|------|------|-------|
| FNP\textsuperscript{a} | 2,429,698 | 26.84 | 94.28 | 2510 | 2391 | ND | 45 | 15 |
| FNN\textsuperscript{b} | 2,174,499 | 27.15 | 89.10 | 2129 | 2067 | 15 | 47 | ND |
| FNV\textsuperscript{c} | 2,118,259 | 27.56 | ND | 2277 | >2212 | 12 | 43 | ND |

\textsuperscript{a}F. nucleatum subsp. polymorphum ATCC 10953, this project, accession number CM000440.

\textsuperscript{b}F. nucleatum subsp. nucleatum ATCC 25586, accession number NC_003454.

\textsuperscript{c}F. nucleatum subsp. vincentii ATCC 49256, accession number NZ_AABF00000000.

ND, not determined.

doi:10.1371/journal.pone.0000659.t001

F. nucleatum
Figure 3. Intergenic repeats. The repeats were aligned using ClustalW [102] and conserved bases were shaded using BOXSHADE (www.ch.embnet.org/software/BOX_form.html). Coordinates are shown in the left column. doi:10.1371/journal.pone.0000659.g003
also unique. Seven additional membrane proteins, two phosphotransferases, and two beta-lactamases are also present only in FNP. Twenty-seven ORFs related to transcriptional regulation are unique to FNP, including a LuxS autoinducer ortholog and two sensor histidine kinases. Additionally, FNP encodes several unique proteins related to DNA modification. These include such functions as methylation, histone acetylation, recombination, integration, topoisomerase, and type I restriction and modification. FNP also contains numerous prophage and transpose genes related to DNA modification. These include such proteins as methylation, histone acetylation, recombination, integration, topoisomerase, and type I restriction and modification.

| Group | Cluster | Conserved gene cluster | FNN | C. difficile | B. anthracis | E. coli |
|-------|---------|------------------------|-----|-------------|-------------|--------|
| I     | 16S, 23S, 5S rRNAs | ND | + | + | + |
| I     | 175 | ND | + | + | + |
| I     | 200 operon (rpoA, rpoB, rpoC) | ND | + | + | + |
| I     | 235 | ND | + | + | + |
| II    | 290 | ND | + | + | + |
| II    | 345 | ND | + | + | + |
| III   | 390 | ND | + | + | + |
| III   | 445 | ND | + | + | + |
| IV    | 490 | ND | + | + | + |
| IV    | 545 | ND | + | + | + |
| V     | 590 | ND | + | + | + |

*The rplX gene is duplicated in this operon in the FNN genome.*

*The L34 operon is present in the C. difficile, B. anthracis and E. coli genomes.*

*The ABC1 family protein gene is missing.*

*An L8A gene maps between nusA and infB.*

*ND, not determined.*

DOI:10.1371/journal.pone.0000659.t002

**Table 2. Conserved gene clusters/operons in FNP.**

**Horizontal Gene Transfer**

HGT can be detected by several parametric methods based on deviant nucleotide composition [45,46], dinucleotide frequencies [47], codon usage biases [48–50] or patterns inferred by Markov chain analysis [51]. Phylogenetic methods determine a gene’s unusual similarity or distribution among organisms by comparing phylogenetic trees of different genes from the genome and assessing the significance of any resulting incongruities. Alternative phylogenetic methods exist that do not reconstruct phylogenetic trees like Clarke’s phylogenetic discordance test [52] and Lawrence’s [53] rank correlation test. Another reliable inference of recent HGT events is the anomalous phylogenetic distribution method wherein a gene is present in one genome but not found in several closely related genomes [54]. This is the approach used examine genes that had no top BLAST hit to either of the two sequenced fusobacterial genomes, FNN and FNV.
Based on BLASTP similarity searches, a total of 1235 ORFs, composed of the 621 FNP ORFs and 9 pseudogenes, was identified without top hits to FNN or FNV (Table S1a) and 608 hypothetical or conserved hypothetical proteins, were graphically plotted to identify clusters that could represent regions of HGT (Figure 4). About 21% of these, or 255 ORFs, mapped within gene clusters. There were 28 specific regions or islands of interest with clusters of 5 or more genes. Top BLAST hits for each cluster (Table S2) were examined to determine a consensus genus and species.

One hundred forty of the ORFs (out of 255), or 55% were putative proteins with no matches to other bacterial proteins. Of the remaining 115, 20 ORFs or 17% had top hits to the Clostridia. The most common top hits in this class were to Clostridium tetani, Clostridium thermodonum, Clostridium perfringens, and Desulfitobacterium hafniense. Other ORFs had top hits to other Firmicutes including Bacillus, Streptococcus, Listeria and Enterococcus species. Hits to the archaea Methanosarcina mazei, Methanoanoicoids burtonii, and Methanothermobacter species, as well as to cyanobacteria Nostoc punctiforme, Trichodesmium erythraeum and Synechocystis sp., were also observed.

A 10 kilobase (kb) region of the FNP genome from nt. 27349 to 37954 (FNP_2111–FNP_2124) appears to have arisen via HGT since this region is not found in the other published fusobacterial genomes and since its GC content (30.4%) is higher than that of the other published fusobacterial genomes including an additional copy of ilvA (FNP_1302), which is not in the other genomes, and two copies of ilvE (branched-chain-amino-acid transaminase), one that is unique to F. nucleatum ATCC 10953 (FNP_1952) and one that is also found in the other two genomes (FNP_1165).

A prophage genome was identified immediately downstream of an arginine tRNA gene between coordinates 2053849 and 2053649 (28.9% GC) (Figures 1 and 5). The ORFs are not found in FNN or FNV (Table S1a). Forty-two open reading frames (FNP_1662–1703) were predicted in this region spanning four clusters of unique genes (XXI–XXIV), including genes encoding integrase, DNA polymerase, anti-repressor, helicase, and terminase proteins. Several genes encoding bacteriophage structural components, such as capsid and tail proteins, were also identified in the region, though 20 of the open reading frames encode hypothetical proteins. The predicted tail proteins were most similar to those in a potential prophage genome in C. tetani E88 [59], while the terminase and packaging proteins were most similar to those in the C. perfringens phage phi3626 [60]. High scoring matches to the nonstructural proteins were found in other gram-positive genomes, such as Bacillus halodurans, Streptococcus mitis, and C. thermodonum. Homologs of 10 of the phage proteins were found in FNV, though only one (a helicase, FNP_1671) was found in FNN. An additional block of phage-like genes map between coordinates 1775962 and 1786249 (FNP_1415–1432) (Figure 1 and Table S2, Cluster XVIII). Only one of the 19 proteins encoded in this region had homologs in the other fusobacteria and only two of the proteins (a replication protein and integrase) matched to other bacteriophage sequences.

The region between nts. 2174023 and 2218775 in the FNP genome (FNP_1820–1879, Figure 1), containing 59 genes that include Clusters XXV–XXVII (25.4% GC), is predicted to contain a large conjugal plasmid (Figure 1, Table S1a, Table S2). Fifty-five unique genes, not found in FNN or FNP, encode a primase/helicase that could function as a replication initiation protein, topoisomerase, integrase, recombinase, a plasmid partitioning protein, and pseudogenes of a mobilization protein and
a plasmid addiction system. This region also carries genes encoding homologs of seven Type IV secretion system (T4SS) proteins. T4SS can translocate DNA and proteins out of the bacterial cell to recipient cells; bacterial conjugation systems are a subset of this family [61]. Full-length copies genes encoding the T4SS proteins VirB4, VirB8, VirB10, VirB11 and VirD4 (FNP_1868–1871, 1873, 1875) are found within the conjugal plasmid region, as are truncated versions of VirB6 and VirB9. These proteins are most similar to orthologs in *Ralstonia*, *Pseudomonas*, *Caulobacter* and *Rhodopseudomonas*. The T4SS proteins identified in FNP could constitute the inner membrane proteins identified in FNP could constitute the inner membrane proteome of the transporter, but genes encoding components for biogenesis of the T-pilus (VirB1, VirB2, VirB3, VirB5 and VirB7) are missing. A different set of proposed Type IV pili genes are present in FNP. A cluster of eleven genes (FNP_2389–2399) plus an unlinked pilT gene may encode the pilus, as suggested by Desvaux, et al. [62].

Five composite ribozyme/transposons, similar to the CdIS/ IStrons described in *C. difficile* [63] were identified in the genome (Figure 1). The consensus IStron in FNP is 1811 nt. long and contains a 477 nt. intron followed by an open reading frame encoding the transposase-like protein, TlpB. The FNP IStron is 31% identical to CdIS-C34 from *C. difficile* and contains four conserved RNA sequences that form the catalytic core of group I introns [64]. All five IStrons in FNP are inserted directly downstream of the pentanucleotide TTCCGAT, which is the conserved site of insertion of the IS8301 family of transposons [65]. The IStrons have an average G+C content of 29%, consistent with that of both fusobacteria and clostridia. In *C. difficile*, CdIS1I self-splices to remove itself from the mRNA into which it is inserted. As a result, the insertion does not disrupt expression of the gene. In FNP, we predict that only one copy of the IStron (1361160 to 1362978) is a fully functional element with self-splicing and transposition activities, since the other copies have mutations in either the ribozyme or *tlpB* regions of the element. Three additional sequences with homology to portions of the ribozyme were identified in the FNP genome and ten additional copies of *tlpB*-like genes occur in the genome. Homologs of this element were not found in any other organism, including the two strains of *Fusobacterium* that have been previously sequenced, though TlpB sequences are found in a variety of organisms, including cyanobacteria, *Bacillus cereus*, *Entenococcus*, *Denococcus* and *Exiguobacterium*. Thus, it appears that a unique exchange between *C. difficile* and FNP has occurred.

ATCC 10953 harbors a single plasmid, pFN3 [40] (Figure 2a), which is 11,934 bp in length and has a GC content of 24.53%. Eleven pFN3 ORFs were identified: two possible replication protein genes, a possible resolvase/recombinase gene, a DNA relaxation protein gene and seven hypothetical protein genes. The two replication protein genes (FNP_pFN3g01 and FNP_pFN3g03) have predicted protein sequences with 20–22% identity and 27–32% similarity to the putative replication protein of the *F. nucleatum* native plasmid pFN1 [40]. The sequence upstream of the pFN3 replication protein gene at 1315 has a sequence (1007 to 1136) characterized by clusters of two overlapping 18 bp repeats (repeat 1: TAATAGTACAAATTTCCG; repeat 2: TAGTACAAATTTCCGAT). Several of the repeats are spaced at 22 bp intervals, suggesting that they may represent replication protein binding sites that are characteristic of the replication origin of irt-r regulated plasmids. The resolvase (FNP_pFN3g09) was identified based on the presence of a N-terminal resolvase domain (psam02796). The DNA relaxation protein (relaxase) (FNP_pFN3g07) has a relaxase domain and contains the conserved consensus motifs defined for relaxase proteins [66] (Figure 2b). The pFN3 resolvase and relaxase genes both have potential significance for HGT. Resolvases are important in DNA recombination events, including excision and integration of mobile DNA elements. Relaxase proteins mediate the initiation of conjugal transfer of plasmid DNA. Plasmids that encode relaxases, which are not conjugative themselves, may be mobilized with the additional conjugative functions provided in *trans*. Two other native *F. nucleatum* plasmids, pFN1 (AF139249) and pPAS2 (AF022647), which are 98% identical, also carry relaxase genes [40]. The occurrence of the relaxase genes suggests the possibility that these plasmids were introduced into *F. nucleatum* by conjugal processes. Consistent with this mechanism of HGT is the finding that plasmids or DNA sequences related to pPAS2 have been detected in 18% of *F. nucleatum* strains examined [67].

**Virulence**

We identified 132 predicted proteins that may play a role in fusobacterial virulence (Table 3). Most of these are found in FNN and FNV, though there are a few notable exceptions. As in the two previously sequenced fusobacterial genomes [20,21], we identified a VacJ homolog (FNP_0314). This protein has been shown to play a role in the intracellular spread of *Shigella flexneri* [68]. Although its mechanism of action has not been examined, VacJ may play a similar role in FNP since recent evidence suggests that *F. nucleatum* can invade epithelial cells [32], which may allow dissemination throughout the host to cause infections at non-oral sites [69]. Other previously known virulence factors were identified in FNP including the porin FomA (FNP_0971) and MviN (FNP_1360), which plays a role in virulence in *Salmonella typhimurium* [72], TraT (FNP_1881) which provides resistance to complement [73], and VacB (FNP_1921), a ribonuclease involved in virulence gene expression in *S. flexneri* [74,75]. The strain also carries genes for butyrate fermentation (FNP_0790, 0791, 0969, 0970, 0971, 1762 and either 1467 or 2146). The production of butyrate has been associated with mouth odor and gingival inflammation [76]. We also include FipA (described above) as a virulence factor because of its immunosup-
### Table 3. Potential virulence factor genes.

| Locus_Tag | Start  | Stop   | Gene | Definition                                      |
|-----------|--------|--------|------|------------------------------------------------|
|           |        |        |      | Miscellaneous virulence factors                |
| FNP_0314* | 693051 | 692284 | vacJ | VacJ family lipoprotein                         |
| FNP_0972* | 133039 | 132928 | fomA | porin FomA                                       |
| FNP_1118  | 145957 | 145875 | bocA | undecaprenol kinase                              |
| FNP_1264  | 161929 | 161829 | fipA | fibronectin-binding protein A                    |
| FNP_1337  | 169232 | 169395 | fbpA | fibronectin-binding protein A                    |
| FNP_1360  | 172053 | 172305 | MviN | family protein                                   |
| FNP_1391  | 175065 | 175258 | vacB | ribonuclease R                                   |
| FNP_1446  | 179736 | 179813 | Bvg | family transcriptional regulator                |
| FNP_1762* | 210632 | 210753 | FipA | acetyl-CoA acetyltransferase/immunosuppressive  |
| FNP_1880  | 221936 | 221978 |       | von Willebrand factor domain protein            |
| FNP_1888  | 222088 | 222854 | traT | complement resistance protein TraT               |
| FNP_1921  | 226332 | 226542 | vacB | ribonuclease R                                   |
|           |        |        |      | Butyrate fermentation                            |
| FNP_2146  | 62092  | 60956  |      | butyryl-CoA dehydrogenase                        |
| FNP_0790  | 115836 | 115913 |      | 3-hydroxybutyryl-CoA dehydratase                |
| FNP_0791  | 115918 | 115987 | fadB | 3-hydroxybutyryl-CoA dehydrogenase              |
| FNP_0969  | 132681 | 132651 | atoA | butyrate-acetoacetate-CoA-transferase, beta subunit |
| FNP_0970  | 132748 | 132683 | atoD | butyrate-acetoacetate-CoA-transferase, alpha subunit  |
| FNP_0971  | 132902 | 132764 | atoE | MFS superfamily major facilitator short chain fatty acids symporter |
| FNP_1467  | 181768 | 181881 |      | butyryl-CoA dehydrogenase                        |
| FNP_1762* | 210632 | 210753 | FipA | acetyl-CoA acetyltransferase                     |
|           |        |        |      | Iron acquisition                                 |
| FNP_2267  | 200871 | 200998 | hmuV | heme ATP binding cassette transporter, ABC protein |
| FNP_2268  | 201893 | 200868 | hmuU | heme ATP binding cassette transporter, membrane protein HmuU |
| FNP_2269  | 202768 | 201896 | hmuT | heme ATP binding cassette transporter, binding protein HmuT |
| FNP_2270  | 205012 | 203039 |      | possible TonB-dependent iron (Fe) receptor       |
| FNP_2353  | 300274 | 299846 | fur  | ferric uptake regulator protein                  |
| FNP_0006  | 386027 | 386275 |      | possible alpha-hemolysin                         |
| FNP_0155* | 520499 | 522103 | tpsB | probable TPS family two-partner secretion family |
| FNP_0156* | 522114 | 530369 | tpsA | probable TPS family two-partner secretion family |
| FNP_0159  | 531328 | 532326 |      | possible hemolysin                               |
| FNP_0338  | 716299 | 716574 |      | cobalamin/Iron ATP binding cassette transporter, ABC protein |
| FNP_0339  | 716593 | 717216 |      | cobalamin/Iron ATP binding cassette transporter, ABC protein |
| FNP_0340  | 717349 | 718404 |      | cobalamin/Iron ATP binding cassette transporter, binding protein |
| FNP_0341  | 718423 | 720105 |      | cobalamin/Iron ATP binding cassette transporter, membrane protein |
| FNP_0428  | 795363 | 796424 |      | Iron ABC superfamily ATP binding cassette transporter, binding protein |
| FNP_0429  | 796439 | 797554 |      | Iron ABC superfamily ATP binding cassette transporter, ABC protein |
| FNP_0430  | 797544 | 799190 |      | Iron ABC superfamily ATP binding cassette transporter, membrane protein |
| FNP_0531  | 896606 | 895221 |      | OfeT family oxidase-dependent iron transporter   |
| FNP_0999  | 135039 | 135104 |      | possible hemolysin III                           |
| FNP_1246* | 1599267| 1588666| tpsA | probable TPS family two-partner secretion exoprotein |
| FNP_1247* | 1601074| 1599281| tpsB | pseudogene of TPS family two-partner secretion |
| FNP_1451  | 1801085| 1801975|      | Iron ABC superfamily ATP binding cassette transporter, binding protein |
| FNP_1452  | 1802030| 1804210|      | Iron ABC superfamily ATP binding cassette transporter, binding protein |
| FNP_1453  | 1805190| 1805966|      | Iron ABC superfamily ATP binding cassette transporter, membrane protein |
| FNP_1454  | 1805980| 1805966|      | Iron ABC superfamily ATP binding cassette transporter, ABC protein |
| FNP_1660  | 2022933| 2021746|      | probable Nramp family metal ion transporter      |
| FNP_1765  | 2114622| 2112373|      | possible TonB-dependent iron (Fe) receptor       |
| Locus_Tag | Start | Stop  | Gene | Definition                                      |
|-----------|-------|-------|------|------------------------------------------------|
| **Drug transporters** |       |       |      |                                                 |
| FNP_0174  | 546224| 547564|       | MOP/MATE family multidrug-resistance efflux pump |
| FNP_0388  | 760374| 759283| perM | probable MFS superfamily major facilitator transporter, macrolide symporter |
| FNP_0507  | 872530| 869465| RND  | superfamily resistance-nodulation-cell division antiporter |
| FNP_0508  | 873639| 872533| RND  | superfamily resistance-nodulation-cell division antiporter |
| FNP_0622  | 986693| 987907|       | probable MFS superfamily major facilitator transporter, multidrug symporter |
| FNP_0640  | 1009181| 1010527| norM1| MOP/MATE superfamily multidrug-resistance efflux pump NorM |
| FNP_0769  | 1137188| 1136280|       | DMT superfamily drug/metabolite transporter |
| FNP_0890  | 1260254| 1261621| norM2| MOP/MATE family multidrug-resistance efflux pump NorM |
| FNP_1162  | 1513737| 1512364| norM3| MOP/MATE family multidrug-resistance efflux pump NorM |
| FNP_1207  | 1557921| 1559300|       | MOP/MATE family multidrug-resistance efflux pump |
| FNP_1299  | 1653389| 1652052| norM4| MOP/MATE family multidrug-resistance efflux pump NorM |
| FNP_1503  | 1864175| 1865314|       | possible MFP membrane fusion protein family transporter |
| FNP_1504  | 1865341| 1866003|       | antimicrobial peptide ATP binding cassette transporter, ABC protein |
| FNP_1505  | 1866000| 1867226|       | antimicrobial peptide ATP binding cassette transporter, membrane protein |
| FNP_1524  | 1881745| 1882185|       | possible DMT superfamily drug/metabolite transporter |
| FNP_1596  | 1955589| 1956983|       | MOP/MATE family multidrug-resistance efflux pump |
| **Beta-lactamas** |       |       |      |                                                 |
| FNP_2175  | 93177 | 92392 |       | beta-lactamase                                   |
| FNP_0581  | 941979| 943874|       | beta-lactamase superfamily zinc-dependent hydrolase |
| FNP_0627  | 993090| 992767|       | probable beta-lactamase superfamily zinc-dependent hydrolase |
| FNP_0629  | 995488| 994865|       | beta-lactamase superfamily zinc-dependent hydrolase |
| **Outer membrane proteins** |       |       |      |                                                 |
| FNP_2182  | 98598 | 97345 |       | probable lipoprotein                            |
| FNP_2196  | 112459| 110366|       | outer membrane protein                          |
| FNP_2270  | 205012| 203039|       | possible TonB-dependent outer membrane receptor  |
| FNP_2283* | 226990| 219536|       | AT family autotransporter                       |
| FNP_2284  | 227596| 227006|       | outer membrane protein                          |
| FNP_2361* | 315873| 308506|       | AT family autotransporter                       |
| FNP_2362  | 316452| 315883|       | OmpA family outer membrane protein              |
| FNP_0032  | 413132| 404055|       | fusobacterial outer membrane protein            |
| FNP_0217  | 583789| 591255|       | fusobacterial outer membrane protein            |
| FNP_0314* | 693051| 692824| vacJ  | VAcJ family lipoprotein                        |
| FNP_0378  | 751249| 751797|       | outer membrane protein OmpF                     |
| FNP_0436  | 804352| 805092|       | outer membrane protein                          |
| FNP_0509  | 874940| 873669|       | TolC family outer membrane protein              |
| FNP_0517  | 882518| 883156|       | probable outer membrane protein                 |
| FNP_0668  | 1046781| 1047941|       | OmpA family outer membrane protein              |
| FNP_0820  | 1189564| 1191015|       | possible outer membrane protein P1              |
| FNP_0972* | 1330393| 1329281| famA | porin FomA                                      |
| FNP_1046* | 1391420| 1402057|       | AT family autotransporter                       |
| FNP_1248  | 1601352| 1601074|       | OmpW                                           |
| FNP_1784  | 2135004| 2136413|       | outer membrane protein ToIC                    |
| FNP_1877  | 2216275| 2217012|       | possible outer membrane protein                 |
| FNP_1891  | 2231145| 2236019|       | probable outer membrane protein                 |
| FNP_1996  | 2339205| 2338783|       | probable lipoprotein                           |
| **Type IV secretion proteins** |       |       |      |                                                 |
| FNP_2389  | 332089| 340542| pulD  | general secretion pathway protein D             |
| FNP_2396  | 346381| 345881| A24   | family prepilin peptidase                       |
| FNP_2397  | 346854| 346266| pulG  | general secretion protein G                    |
| Locus_Tag | Start  | Stop   | Gene | Definition                        |
|-----------|--------|--------|------|-----------------------------------|
| FNP_2398  | 348121 | 347081 | pulG | general secretion protein F       |
| FNP_2399  | 349362 | 348118 | pulE | general secretion protein E       |
| FNP_1034  | 1383491| 1382541| pilT | Tfp pilus assembly protein PilT    |
| FNP_1868  | 2206819| 2207523| trbF | probable conjugal transfer protein TrbF/VirB8 |
| FNP_1869  | 2207535| 2208371| trbG | probable conjugal transfer protein TrbG/VirB9 |
| FNP_1870  | 2208381| 2209583| trbI | probable conjugal transfer protein TrbI/VirB10 |
| FNP_1871  | 2209586| 2211616| traG | probable conjugal transfer protein TraG/VirD4 |
| FNP_1873  | 2212045| 2213010| trbB | probable conjugal transfer protein TrbB/VirB11 |
| FNP_1875  | 2213293| 2216031| trbE | probable conjugal transfer protein TrbE/VirB4 |

**Type V secretion proteins**

| Locus_Tag | Start  | Stop   | Gene | Definition                        |
|-----------|--------|--------|------|-----------------------------------|
| FNP_2152  | 69523  | 69237  |      | frameshift of AT family transporter |
| FNP_2283* | 226990 | 219536 | AT   | family autotransporter            |
| FNP_2361* | 315873 | 308506 | AT   | family autotransporter            |
| FNP_0035  | 417960 | 415012 |      | possible autotransporter           |
| FNP_1046* | 1402057| 1391420| AT   | family autotransporter            |
| FNP_1391  | 1750650| 1752584|      | possible autotransporter adhesin   |
| FNP_1637  | 1986896| 1996964| AT   | family autotransporter            |
| FNP_2077  | 2423459| 2420256| AT   | family autotransporter            |
| FNP_0155* | 520499 | 522103 | TpsB | probable TPS family two-partner secretion family protein TpsB |
| FNP_0156* | 522114 | 530369 | TpsA | probable TPS family two-partner secretion family exoprotein TpsA |
| FNP_1246* | 1599267| 1599281| TpsB | pseudogene of TPS family two-partner secretion protein TpsB |

**Proteases**

| Locus_Tag | Start  | Stop   | Gene | Definition                        |
|-----------|--------|--------|------|-----------------------------------|
| FNP_0461  | 825985 | 827004 |      | M50A family metalloprotease       |
| FNP_0897  | 1266394| 1267038|      | O-sialoglycoprotein endopeptidase  |
| FNP_1813  | 2167021| 2168046|      | O-sialoglycoprotein endopeptidase  |

**Lipopolysaccharide biosynthesis**

| Locus_Tag | Start  | Stop   | Gene | Definition                        |
|-----------|--------|--------|------|-----------------------------------|
| FNP_2334  | 282271 | 280451 |      | O-antigen acetylase               |
| FNP_0533  | 898226 | 899365 |      | possible ADP-heptose:LPS heptosyltransferase |
| FNP_0534  | 899377 | 900162 | licD | lipopolysaccharide cholinephosphotransferase |
| FNP_0537  | 902005 | 903150 | waaG | glycosyltransferase               |
| FNP_0538  | 903150 | 903902 |      | possible polysaccharide deacetylase |
| FNP_0539  | 904782 | 905864 |      | possible glycosyltransferase      |
| FNP_0540  | 905879 | 906610 |      | probable glycosyltransferase      |
| FNP_0541  | 908004 | 909200 | waaL | possible O-antigen ligase         |
| FNP_0830  | 1201261| 1200185| waaF1| heptosyltransferase II (inner core) |
| FNP_1103  | 1443645| 1442839|      | possible glycosyltransferase      |
| FNP_1104  | 1444893| 1443745|      | UDP-N-acetylgalactosamine 2-epimerase |
| FNP_1105  | 1446154| 1444898| neuA | CMP-N-acetylneuraminic cytidylyltransferase |
| FNP_1106  | 1447200| 1446157| neuB | possible N-acetyl neuramic acid synthetase |
| FNP_1107  | 1447822| 1447205|      | N-acetylneuraminic synthase       |
| FNP_1108  | 1449226| 1447934|      | oligosaccharidyl-lipid/polysaccharide flippase |
| FNP_1109  | 1450280| 1449300|      | possible lipooligosaccharide sialyltransferase |
| FNP_1205  | 1556498| 1557478| waaE | possible ADP-heptose synthase     |
| FNP_1807  | 2161658| 2162707| waaF2| LPS heptosyltransferase II        |
| FNP_1808  | 2162704| 2163732| waaF3| LPS heptosyltransferase II        |
| FNP_1809  | 2163725| 2164327| waaY | possible lipopolysaccharide core biosynthesis protein WaaY |
| FNP_1810  | 2164329| 2165336| waaQ | lipopolysaccharide core biosynthesis glycosyl transferase WaaQ |
| FNP_1907  | 2251044| 2251837| lpxC | UDP-3-O-acyl-N-acetylgalactosamine deacetylase |
pressive properties [43,77], though it is most similar to an acetyl-CoA transferase of the butyrate fermentation system. The *fipA* gene is not present in FNV.

The acquisition of iron from the host environment is an important function of most bacterial pathogens [78]. We have identified 26 predicted proteins involved in iron uptake in FNP. Three proteins, HmuV (FNP_2267), HmuU (FNP_2266), and HmuI (FNP_2269), form a heme ABC superfamily ATP binding cassette transporter while two additional proteins (FNP_2270 and FNP_1765) are probable TonB-dependent heme receptors. There are also two additional iron ATP binding cassette transporters (FNP_428–430 and FNP_1451–1454) and a cobalamin/iron ATP binding cassette transporter (FNP_0398–0341). An Nrsmp family iron transporter (FNP_1660), and an OftT family oxidase-dependent iron transporter (FNP_0531) were also annotated. Three hemolysin genes were identified (FNP_0006, FNP_0159, and FNP_0999); two of these (FNP_0159 and FNP_0999) have associated TPS family two-partner secretion proteins (FNP_0155–0156 and FNP_1246–1247, respectively). This is similar to what is seen in FNV, but different from FNN, which has three such pairs [21]. Several of the iron transporters (FNP_0339, FNP_0428, FNP_0531, and FNP_0769) are present in FNN but are missing in FNV; thus FNV may have a diminished requirement for iron or may occupy a different niche. As mentioned previously, a homolog of the ferric uptake regulator, Fur (FNP_2353), was identified in the FNP genome, though it is not present in FNV.

Sixteen possible drug transporters were annotated. These included 7 MOP/MATE family multidrug efflux pumps (FNP_0174, FNP_0640, FNP_0890, FNP_1162, FNP_1207, FNP_1299, and FNP_1596), 2 DMT superfamily drug/metabolite transporters (FNP_0388 and FNP_0622) and 2 RND family antiporters (FNP_0507 and FNP_0508). Our annotation did not permit us to predict the substrates of these transporters but it is likely that many of them are antibiotic transporters. With respect to antibiotic resistance, we also annotated 4 genes predicted to encode beta-lactamasases. One of these (FNP_0627) is unique to FNP.

We identified all but one (FN0387) of the 14 outer membrane protein genes described by Kapatral, et al [20] (two, FNP_1046 and FNP_2283, have been re-annotated as AT family autotransporters) and we discovered a gene encoding OmpW (FNP_1248), which is not found in FNN or FNV. Four potential adhesion proteins were identified including a fibronectin-binding protein homolog (FNP_1337), a possible autotransporter adhesion (FNP_1391), and two proteins (FNP_1880 and FNP_1888) containing von Willebrand (vWF) type A domains. In addition to the Type IV secretion system discussed previously, FNP also carries genes that belong to the Type V secretion system. These include ten autotransporter genes (8 class 1, Type Va; 2 class 2, Type Vc) and the Tps secretion genes *tpsA* and *tpsB* (Type Vb) (Table 3). These are a subset of the genes found in FNN and FNV [62].

Twenty-five ORFs predicted to be involved in the biosynthesis of LPS were identified. This is of interest because *F. nucleatum* has been shown to have endotoxin activity [23,79]. Unlike FNN, however, FNP does not possess the *lic* operon, which is predicted to attach choline residues to the LPS [20]; only the *licC* and *licD* genes, encoding phosphocholine cytidylyltransferase and a phosphotransferase, respectively, are present. In contrast, FNP does contain genes (FNP_1105–1107) that encode N-acetylneuraminic cytidylyltransferase, N-acetyl neuraminic synthase, N-acetylneuraminic synthase and a possible lipooligosaccharide sialyltransferase (FNP_1109) that might incorporate sialic acid into LPS, like FNV (note, however, that FNP_1109 is not found in FNN or FNV). This may facilitate evasion from the host immune response.

### Signal Transduction

Six potential two-component signal transduction systems were revealed in FNP and three are of particular interest. An EnvR-related response regulator maps immediately upstream of a possible sensor histidine kinase gene. While the response regulator (FNP_2108) has homologs in both previously sequenced *F. nucleatum* genomes [20,21] (FN1261, FN1053), the sensor protein (FNP_2107) is not present in either of these two genomes. Also unique to this region is an ORF (FNP_2109) immediately downstream of the response regulator that is not found in the other sequenced fusobacterial genomes. This protein appears to contain both sensor histidine kinase and response regulator domains and may represent a fusion of two two-component domains. Another system is a possible ethanolamine two-component regulatory system (FNP_0128 and FNP_0129). This system is present in the FNN genome but is not found in the FNV genome, though it may be located in one of the unfinished regions of that genome. Both FNN and FNV have genes encoding the YesM/YesN two-component regulatory system. In FNP the response regulator, YesN (FNP_0212) has been interrupted by the insertion of an IStron. The IStron at this locus is predicted to be inactive so *yesN* should be non-functional. In addition to the six possible two-component systems identified in FNP, there is unmatched response regulator and one unmatched sensor histidine kinase.

### Communication with other bacteria

*F. nucleatum* plays an important role in the formation of oral biofilms, or plaque. *F. nucleatum* is believed to act as a bridging organism between the Gram-positive early colonizers and the Gram-negative late colonizers [23,36,70]. It has been proposed [80] that some bacteria use a compound known as AI-2 (autoinducer-2) for intra-species communication. AI-2 signaling is also involved with biofilm formation [51]. We have identified the protein responsible for AI-2 synthesis, LuxS (FNP_1558). LuxS activity has been previously reported in ATCC 10953 and in three additional *F. nucleatum* strains [82]. There is no homolog of *luxS* in either of the two previously annotated *F. nucleatum* genomes.

### Oxidative stress

Oral anaerobes must cope with oxidative stress to survive and contribute to pathogenesis. Established cultures of FNP are aerotolerent, indicating the presence of mechanisms to detoxify oxygen or oxygen radicals [83]. The ability of *F. nucleatum* to...
manage oxidative stress and to maintain reduced conditions is thought to facilitate the survival of other anaerobic pathogens. This facet of F. nucleatum ecology may explain why less aerotolerant organisms such as Porphyromonas gingivalis are increased in number in the presence of F. nucleatum [36]. Aeroprotection of bystander organisms may also explain the synergistic increase in virulence of F. nucleatum when combined with Porphyromonas gingivalis, as compared to either species alone [84,85]. Physiological studies demonstrate that in response to oxidative stress, FNP maintains a reduced environment [5], with increases in NADH oxidase and superoxide dismutase activities [83]. Our analysis of the FNP genome revealed an NADH oxidase (FNP_1794), most closely related to treponemal and streptococcal homologs. An ORF encoding a superoxide dismutase was not identified, though, a rubrerythrin protein (FNP_1721), which confers superoxide dismutase-like activity [86] and has homology to a C. perfringens rubrerythrin, is present in FNP. Other FNP proteins of potential importance in oxidative stress include a glutathione peroxidase (FNP_2310) [87,88], a thioredoxin (FNP_0146) [89], a glutaredoxin (FNP_1273) [90], and an alkyl hydroperoxide (FNP_2238) [89,91]. Orthologs of these genes are present in both the FNN and FNV genomes.

Conclusion
Analysis of the genome sequence of F. nucleatum subsp. polymorphum revealed that this microorganism has obtained numerous genes from the Firmicute phylum of bacteria. In particular, many of the regions of FNP that were unique to the Fusobacterium encoded proteins with top BLAST hits to the Clostridia. This is perhaps not unexpected as Firmicutes were the largest phylotype isolated in a gingival bacterial diversity study, and the Clostridia represented the largest class within the Firmicutes that were isolated [1]. It appears that FNP was the recipient of DNA from bacteria that are common to the subgingival niche in HGT events. A contrasting interpretation, presented by Mira, et al. [92], is that Fusobacterium represents a genus with a Clostridial metabolic apparatus that has obtained a gram-negative envelope from the Proteobacteria. If this is the case, then it appears that FNN and FNV have lost a complement of genes that are present in FNP. As was revealed in the comparison of FNN to FNV, the fusobacterial genomes are mosaic in structure.

MATERIALS AND METHODS
Bacterial strain, growth conditions, and DNA isolation
F. nucleatum subsp. polymorphum ATCC 10953 is a human strain originally isolated from the inflamed gingiva of an adult male [93]. The strain was acquired as a lyophil from the American Type Culture Collection and recovered on Columbia agar with 5% sheep blood under anaerobic growth conditions. Genomic DNA was prepared from an anaerobically cultivated Columbia broth culture, as previously described [71].

Library construction and sequencing
Genomic DNA of F. nucleatum ATCC 10953 was sheared to 2–6 kb in size using a nebulizer (CIS-US, Inc., Bedford, Mass.), purified from an agarose gel, cloned into a pUC18 derivative sequencing vector, and sequenced as previously described [94].

Sequence assembly
PHRED [95,96] was used to determine the sequence and quality of each base of sequencing reads. Atlas genome assembly tools [97] were used to process reads, to remove repetitive reads, and to bin overlapping reads before assembling the contigs with the program PHRAP [95]. An initial assembly of 54,946 reads gave 146 contigs. Determination of linkages among contigs and gap closing was carried out as previously described [94] except that repetitive reads were put back into the contigs by separated bin assemblies. The rRNAs were not completely sequenced and there are 14 short gaps in the sequence.

The coordinate system for the genome was selected to be similar to other sequenced fusobacteria. In this system, the origin of replication is tentatively localized to the region upstream of an ORF with low identity to DnaA. revF and gsvB are found downstream of this putative dnaA gene. dnaE and repB, which are often clustered with dnaA around the origin of replication, are found elsewhere in the FNP genome, as they are in the other two Fusobacterium genomes.

Gene identification and annotation
Gene prediction and manual annotation were performed as previously described [94]. For GeneMark [90] gene predictions, Borrelia burgdorferi was used as model. Predicted proteins with unknown functions were classified into three categories: "hypothetical protein" for proteins that do not have matches in the GenBank NR database, "conserved hypothetical protein" for proteins that have matches to proteins of unknown functions from organisms outside the genus Fusobacterium, and "fusobacterial conserved hypothetical protein" for proteins that have matches only to other fusobacterial proteins of unknown functions.

Transfer RNAs were identified using rRNascan-SE [99]. Non-coding RNA (ncRNA) elements where identified by BLASTing the known ncRNA sequences of FNN, available at the Rfam database (http://www.sanger.ac.uk/Software/Rfam/), against the FNP genome [100]. The FNP regions containing potential ncRNAs were then analyzed by the Rfam database to identify the exact coordinates of each RNA element.

Genome analysis and identification of possible HGT regions
We performed automatic and manual comparison of best hits to the GenBank NR database and BLAST results to the FNN and FNV genomes to find possible genomic islands and recent horizontally transferred genes in the FNP genome. Genes and regions of possible HGT were examined by GC content, cumulative GC profile, and codon usage bias analysis [101] but this did not add significantly to the conclusions. GlustalW was also used to align DNA sequences and to identify repeat sequences in the genome [102].

Analysis of intergenic regions
After completing annotation of all ORFs called by either GeneMark or Glimmer we decided to confirm that the intergenic regions (IGR) contained no ORFs missed by the ORF calling software. Thus, the entire nucleotide sequence of each IGR was compared to Genbank using BLASTX and those results that had an e value of less than 1 x 10^-5 were marked for further analysis by annotators. The annotators defined the exact coordinates of each hit and then followed the same process used for the called ORFs to assign the annotation for each region, resulting in the addition of 50 annotation entries. While this process did result in the identification of a few missed ORFs (5–10), the vast majority of IGR annotations were pseudogenes. Of note concerning this analysis is that it allowed us to extend proteins called as small hypothetical proteins into full-length pseudogenes.
Acknowledgments

We thank Richard Gibbons, Donna Muzny, Christie Kovar-Smith, Lynne Nazareth, Erica Sodergren, David Parker, Aleks Milosavljevic, and the rest of the staff at the Human Genome Sequencing Center for their support during this project. Part of the DNA sequencing was performed atSeqsprint, Inc., Houston, TX.

Author Contributions

Conceived and designed the experiments: QX SK GW SHa SHi XQ. Performed the experiments: SK SHi. Analyzed the data: JP QX SK HJ SY YL JG SF GHa SHi XQ. Contributed reagents/materials/analysis tools: QX HJ SY GF GW SHa SHi. Wrote the paper: SK GW SHa.

References

1. Paster BJ, Boches SK, Galvin JT, Ericson RE, Lau CN, et al. (2001) Bacterial diversity in human subgingival plaque. J Bacteriol 183: 3770–3783.
2. Ximeñez-Fyvie LA, Haffajee AD, Socransky SS (2000) Comparison of the microflora of suprag- and subgingival plaque in health and periodontitis. J Clin Periodontal 27: 649–657.
3. Kononen E (2000) Development of oral bacterial flora in young children. Ann Med 32: 107–112.
4. Kolenbrander PE, London J (1993) Adhere today, here tomorrow: Oral bacterial adherence. J Bacteriol 175: 3247–3252.
5. Diaz PI, Zilm PS, Rogers AH (2002) Fusobacterium nucleatum supports the growth of Porphyromonas gingivalis in oxygenated and carbon-dioxide-depleted environments. Microb Ecol 44: 467–472.
6. Moore WE, Moore LV (1994) The bacteria of periodontal diseases. Periodontol 2000 0: 56–77.
7. Brown LJ, Lee H (2000) Prevalence, extent and severity of periodontal disease. Periodontol 2000 0: 57–71.
8. Brook I, Frazier EH (1998) Microbiology of liver and spleen abscesses. J Med Microbiol 47: 1075–1080.
9. Chadflyy F, Dhawan B, Laxmi BVJ, Mehta VS (1998) The microbial spectrum of brain abscess with special reference to anaerobic bacteria. Br J Neurosurg 12: 127–130.
10. Chrysagis AM, Brusselmann CB, Rombouts JJ (2001) Septic arthritis of the hip due to Fusobacterium nucleatum. Clin Rheumatol 20: 229–231.
11. Civera R, Joussin-Sommer H, Marina M, Borestein L, Shah H, et al. (1995) A retrospective review of cases of anaerobic empyema and update of bacteriology. Clin Infect Dis 20 Suppl 2: S224–229.
12. Goldstein EJ, Summanen PH, Citron DM, Fingold SM (1995) Fatal sepsis due to a beta-lactamase-producing strain of Fusobacterium nucleatum subspecies polymorphum. Clin Infect Dis 20: 797–800.
13. Holst E, Goffeng AR, Andersch B (1994) Bacterial vaginosis and vaginal microorganisms in idiopathic premature labor and pregnancy outcome. J Clin Microbiol 32: 176–186.
14. Jousimies-Somer H, Savolainen S, Makite A, Ylikoski J (1993) Bacteriological findings in perintonellar abscesses in young adults. Clin Infect Dis 16 Suppl 4: S292–300.
15. Cahill RJ, Tan S, Dougan G, O’Gaora P, Pickard D, et al. (2005) Universal DNA primers amplify bacterial DNA from human fetal membranes and link Fusobacterium nucleatum with prolonged preterm membrane rupture. Mol Hum Reprod 11: 761–766.
16. Bearfield C, Davernport ES, Sirapahasundaram V, Allaker RP (2002) Possible association between amniotic fluid micro-organism infection and microflora in the mouth. BJOG 109: 527–533.
17. Hill GB (1998) Preterm birth: associations with genital and possibly oral microflora. Ann Periodontol 3: 222–232.
18. Mikamo H, Kawaike K, Sato Y, Imai A, Tamaya T (1998) Preterm labor and bacterial intraamniotic infection: arachidonic acid liberation by phospholipase A2 of Fusobacterium nucleatum. Am J Obstet Gynecol 179: 1579–1582.
19. Mikamo H, Kawaike K, Sato Y, Tamaya T (1999) Elastase activity of anaerobes isolated from amniotic fluid with preterm rupture of membranes. Am J Obstet Gynecol 180: 378–380.
20. Kapatal V, Anderson I, Ivanova N, Reznik G, Los T, et al. (2002) Genome sequence analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586. J Bacteriol 184: 2005–2018.
21. Kapatal V, Ivanova N, Anderson I, Reznik G, Bhattacharyya A, et al. (2003) Genome analysis of F. nucleatum sub sp nucleatum and its comparison with the genome of F. nucleatum sub sp mitis. Genome Res 13: 1180–1189.
22. Morris ML, Andrews RH, Rogers AH (1996) The use of allozyme electrophoresis to assess genetic heterogeneity among previously subspecies isolates of Fusobacterium nucleatum. Oral Microbiol Immunol 11: 15–21.
23. Bolstad AL, Jensen HB, Bakken V (1996) Taxonomy, biology, and periodontal aspects of Fusobacterium nucleatum. Clin Microbiol Rev 9: 55–71.
24. Connah G, Claro MG, Citron DM, Tyrell KL, Merriam V, et al. (2002) 16S rDNA internal transcribed spacer sequences for analysis of the phylogenetic relationships among species of the genus Fusobacterium. Int J Syst Evol Microbiol 52: 493–499.
25. Gnur R, Munson RA, Warde WG (2006) Genotypic and phenotypic characterization of fusobacteria from Chinese and European patients with inflammatory periodontal disease. Syst App Microbiol 29: 120–130.
26. Robins AH, Oliver G, Thompson J (1987) Amino acid-dependent transport of sugars by Fusobacterium nucleatum ATCC 10953. J Bacteriol 169: 3891–3897.
27. Robins SA, Thompson J (1990) Regulation of fructose metabolism and polymer synthesis by Fusobacterium nucleatum ATCC 10953. J Bacteriol 172: 5714–5723.
28. Rogers AH, Gully NJ, Pleinig AL, Zilm PS (1992) The breakdown and utilization of peptides by strains of Fusobacterium nucleatum. Oral Microbiol Immunol 7: 299–303.
29. Rogers AH, Zilm PS, Gully NJ, Pleinig AL, Marsh PD (1991) Aspects of the growth and metabolism of Fusobacterium nucleatum ATCC 10953 in continuous culture. Oral Microbiol Immunol 6: 230–235.
30. Tuttle RS, Strebel NA, Mourad J, Mangan DF (1992) A non-lectin-like mechanism by which Fusobacterium nucleatum 10953 adheres to and activates human lymphocytes. Oral Microbiol Immunol 7: 78–83.
31. Ozaki M, Miyake Y, Shirakawa M, Takemoto T, Okamoto H, et al. (1999) Blunting specificity of fusobacteirium nucleatum to human erythrocytes, polymorphonuclear leukocytes, fibroblasts, and HeLa cells. J Periodontal Res 25: 129–134.
32. Han VW, Shi W, Huang GT, Kinder Haake S, Park NH, et al. (2000) Interactions between periodontal bacteria and human oral epithelial cells: Fusobacterium nucleatum adheres to and invades epithelial cells. Infect Immun 68: 3140–3146.
33. Ribeiro-Sobrinho AP, Rabelo FL, Albuquerque CB, Alvarez-Leite JL, Nicoli JR, et al. (2005) Bacteria recovered from dental pulp induce apoptosis of lymph node cells. J Med Microbiol 54: 413–418.
34. Jewett A, Hume WR, Le H, Huynh TN, Han YW, et al. (2006) Induction of apoptotic cell death in peripheral blood mononuclear and polymorphonuclear cells by an oral bacterium, Fusobacterium nucleatum. Infect Immun 74: 1093–1098.
35. Takada H, Ogawa T, Yoshimura F, Otsuka K, Koseguchi S, et al. (1988) Immunobiological activities of a porin fraction isolated from Fusobacterium nucleatum ATCC 10953. Infect Immun 56: 853–863.
36. Bradshaw DJ, Marsh PD, Watson GK, Allison C (1998) Role of Fusobacterium nucleatum and coaggregation in anaerobe survival in planktonic and biofilm oral microbial communities during aeration. Infect Immun 66: 4729–4732.
37. Kuriyama T, Nakagawa K, Kawashiri S, Yamamoto E, Nakamura S, et al. (2000) The virulence of mixed infection with Streptococcus constellatus and Fusobacterium nucleatum in a murine orofacial infection model. Microbes Infect 2: 1425–1430.
38. Takemoto T, Kurihara H, Dahlen G (1997) Characterization of Bacteroides forsythus inclusions. J Clin Microbiol 35: 1378–1381.
39. Kinder Haake S, Yoder SC, Attarian G, Podkaminer K (2000) Native plasmds of Fusobacterium nucleatum: Characterization and use in development of genetic systems. J Bacteriol 182: 1176–1180.
40. Haake SK, Yoder SC, Attarian G, Podkaminer K (2000) Native plasmds of Fusobacterium nucleatum: characterization and use in development of genetic systems. J Bacteriol 182: 1176–1180.
41. Lindahl I, Zengel JM (1986) Ribosomal genes in Escherichia coli. Ann Rev Genet 20: 297–326.
42. Siefert JL, Martin KA, Abdi F, Wijgert WR, Fox GE (1997) Conserved gene clusters in bacterial genomes provide further support for the primacy of RNA. J Mol Evol 43: 467–472.

43. Demuth DR, Savarey R, Golub E, Shenker BJ (1996) Identification and analysis of fipA, a Fusobacterium nucleatum immunosuppressive factor gene. Infect Immun 64: 1333–1341.

44. Hunt Gerardo S, Yoder SC, Citron DM, Goldstein EJC, Kinder Haake S (2002) Sequence conservation and distribution of the fusobacterial immunosuppressive protein gene, fipA. Oral Microbiology and Immunology 17: 313–320.

45. Lawrence JG, Ochman H (1997) Amplification of bacterial genomes: rates of change and exchange. J Mol Evol 44: 303–397.

46. Lawrence JG, Ochman H (1998) Molecular archaeology of the Escherichia coli genome. Proc Natl Acad Sci U S A 95: 9413–9417.

47. Karlin S (1998) Global dimodulate signatures and analysis of genomic heterogeneity. Curr Opin Microbiol 1: 596–610.

48. Mrazek J, Karlin S (1999) Detecting alien genes in bacterial genomes. Ann NY Acad Sci 879: 318–329.

49. Medigue C, Rouxel T, Viger P, Henaut A, Danchin A (1994) Evidence for horizontal gene transfer in Escherichia coli speciation. J Mol Biol 222: 851–856.

50. Moxey I, Roca EP, Danchin A (1999) Codon usage and lateral gene transfer in Bacillus subtilis. Curr Opin Microbiol 2: 524–528.

51. Hayes WS, Borodovsky M (1998) How to interpret an anonymous bacterial genome: Machine learning approach to genome identification. Genome Res 8: 1154–1171.

52. Ragan MA (2001) On surrogate methods for detecting lateral gene transfer. Curr Opin Microbiol 4: 524–528.

53. Philippe H, Douady CJ (2003) Horizontal gene transfer and phylogenetics. Curr Opin Microbiol 6: 498–505.

54. Ragan MA (2001) On surrogate methods for detecting lateral gene transfer. Curr Opin Microbiol 4: 524–528.

55. Lawrence JG, Ochman H (1997) Amplification of bacterial genomes: rates of change and exchange. J Mol Evol 44: 303–397.

56. Lawrence JG, Ochman H (1998) Molecular archaeology of the Escherichia coli genome. Proc Natl Acad Sci U S A 95: 9413–9417.

57. Karlin S (1998) Global dimodulate signatures and analysis of genomic heterogeneity. Curr Opin Microbiol 1: 596–610.

58. Mrazek J, Karlin S (1999) Detecting alien genes in bacterial genomes. Ann NY Acad Sci 879: 318–329.

59. Moxey I, Roca EP, Danchin A (1999) Codon usage and lateral gene transfer in Bacillus subtilis. Curr Opin Microbiol 2: 524–528.

60. Zimmer M, Scherer S, Loessner MJ (2002) Genomic analysis of Bacillus subtilis chromosome I. Bioinform 20: 612–622.

61. Christie PJ (2004) Type IV secretion: the fipA gene in Plasmodium falciparum. Bioinform 20: 612–622.

62. Knorr 1922: 281–286. Fusobacterium nucleatum subsp. nucleatum. J Bacteriol 182: 5290–5299.

63. Diaz PI, Zilm PS, Rogers AH (2000) The response to oxidative stress of Bacillus subtilis. J Mol Biol 298: 351–359.

64. Bruggemann H, Baumer S, Fricke WF, Wiezer A, Liesegang H, et al. (2003) Propanediol utilization genes (pdu) of Salmonella typhimurium: identification of close relatives. J Bacteriol 185: 6633–6639.

65. Bobik TA, Hlavacek GD, Busch RJ, Williams DS, Aldrich HC (1999) The propanediol utilization (pdu) operon of Salmonella enterica serovar Typhimurium LT2 includes genes necessary for formation of polymeric organosilanes involved in coenzyme B12-dependent 1,2-propanediol degradation. J Bacteriol 181: 5967–5975.

66. Ilyina TV, Koonin EV (1992) Conserved sequence motifs in the initiator protein of the T18. Arch Oral Biol 37: 318–326.

67. Shigella flexneri. Plasmid 33: 15–20.

68. Yoshida A, Ansai T, Takehara T, Kuramitsu HK (2005) LuxS-based signaling affects Streptococcus mutisii biofilm formation. Appl Environ Microbiol 71: 2372–2380.

69. Friis O, Olle E, Almá N (2001) Periodontal pathogens produce quorum sensing signal molecules. Infect Immun 69: 3431–3434.

70. Kaufman J, DiRienzo JM (1989) Isolation of a corncob (coaggregation) protein from Fusobacterium nucleatum. Infect Immun 57: 131–137.

71. Kindzierska S, Wang X (1997) Cloning and expression of FomA, the major outer-membrane protein gene from Fusobacterium nucleatum T18. Arch Oral Biol 42: 19–24.

72. Caisnov M, Stocker B, Weinstein D, O’Brien A (1989) A Salmonella typhimurium virulence gene linked to fipA. Infect Immun 57: 3276–3280.

73. Sukopoulou S, O’Connor CD (1990) TcT Lipoepoxygenase, a plasmid-specified mediator of interactions between Gram-negative bacteria and their environment. Microbiol Rev 54: 331–341.

74. Tobe T, Saakawa C, Okada N, Honna Y, Yoshikawa M (1992) vacB, a novel chromosomal gene required for expression of virulence genes on the large plasmid of Shigella flexneri. J Bacteriol 174: 6539–6563.

75. Cheng Z-F, Zuo Y, Li Z, Rudi KE, Deutscher MP (1998) The Shigella flexneri vacB gene required for virulence in Shigella flexneri and Escherichia coli encodes the escherichia coli phosphotransferase. J Bacteriol 170: 408–413.

76. Niedermaier R, Zhang J, Kashket S (1997) Short-chain carboxylic-acid-stimulated, PNP-mediated gingival inflammation. Crit Rev Oral Biol Med 8: 269–290.

77. Shigekawa SJ, Dzink JL, Sheenan MT, Socransky SS (1990) Proposal of three subspecies of Porphyromonas gingivalis. J Bacteriol 168: 6008–6015.

78. Brenot A, King KY, Janowiak B, Griffith O, Caparon MG (2004) Contribution of glutathione peroxidase to the virulence of Staphylococcus pyogenes. Infect Immun 72: 408–413.

79. King KY, Hornstein JA, Caparon MG (2000) Acetolactate and peroxide resistance in perfringolysin and PerR mutants of Staphylococcus pyogenes. J Bacteriol 182: 5290–5299.

80. McLeod MP, Qin X, Karpathy SE, Gioia J, Highlander SK, et al. (2004) Complete genome sequence of Rickettsia typhi and comparison with sequences of other rickettsiae. J Bacteriol 186: 5842–5855.

81. Yoshida A, Ansai T, Takehara T, Kuramitsu HK (2005) LuxS-based signaling affects Streptococcus mutisii biofilm formation. Appl Environ Microbiol 71: 2372–2380.

82. Friis O, Olle E, Almá N (2001) Periodontal pathogens produce quorum sensing signal molecules. Infect Immun 69: 3431–3434.

83. Diaz PI, Zilm PS, Rogers AH (2000) The response to oxidative stress of Bacillus subtilis. J Mol Biol 298: 351–359.

84. Bruggemann H, Baumer S, Fricke WF, Wiezer A, Liesegang H, et al. (2003) Propanediol utilization genes (pdu) of Salmonella typhimurium: identification of close relatives. J Bacteriol 185: 6633–6639.

85. Bobik TA, Hlavacek GD, Busch RJ, Williams DS, Aldrich HC (1999) The propanediol utilization (pdu) operon of Salmonella enterica serovar Typhimurium LT2 includes genes necessary for formation of polymeric organosilanes involved in coenzyme B12-dependent 1,2-propanediol degradation. J Bacteriol 181: 5967–5975.

86. Braun V, Mehlig M, Moos M, Rupnik M, Kalt B, et al. (2000) A chimeric protein for rolling circle DNA replication encoded by diverse replicons from transposable elements of Helicobacter pylori. Biochim Biophys Acta 1713: 92–1125.

87. Vijayakumar LW, Bapat BM, Medzhitov R, Tauber MG, Caparon MG (1999) Role of perfringolysin in the pathogenesis of their superoxide dismutase function. J Bacteriol 178: 7152–7158.

88. Lawrence JG, Ochman H (1997) Amplification of bacterial genomes: rates of change and exchange. J Mol Evol 44: 303–397.

89. Lawrence JG, Ochman H (1998) Molecular archaeology of the Escherichia coli genome. Proc Natl Acad Sci U S A 95: 9413–9417.