INFECTION AND IMMUNITY, July 2003, p. 4217–4221 Vol. 71, No. 7
0019-9567/03/$08.00+0 DOI: 10.1128/IAI.71.7.4217–4221.2003

Asymmetrical Distribution of *Neisseria* Miniature Insertion Sequence DNA Repeats among Pathogenic and Nonpathogenic *Neisseria* Strains

Eliana De Gregorio, Chiara Abrescia, M. Stella Carlomagno, and Pier Paolo Di Nocera*

Dipartimento di Biologia e Patologia Cellulare e Molecolare, "L. Califano," Università degli Studi di Napoli Federico II, 80131 Naples, Italy

Received 31 January 2003/Returned for modification 18 March 2003/Accepted 27 March 2003

***Neisseria* miniature insertion sequences (*nemis*) are miniature DNA insertion sequences found in *Neisseria* species. Out of 57 elements closely flanking cellular genes analyzed by PCR, most were conserved in *Neisseria meningitidis* but not in *N. lactamica* strains. Since mRNAs spanning *nemis* are processed by RNase III at hairpins formed by element termini, gene sets could selectively be regulated in meningococci at the posttranscriptional level.**

DNA repeats known as Correia (4) or *Neisseria* miniature insertion sequences (*nemis* [9]) represent about 2% of *Neisseria meningitidis* genomes (10, 12). These elements mostly differ in the presence and/or absence of a 50-bp long internal segment, contain terminal inverted repeats (TIRs) of variable length (Fig. 1A), and induce the specific duplication of the TA dinucleotide upon genomic integration (3, 8, 9). *nemis* have no coding capacity, and whether they are inactive remnants of larger mobile elements or can still be mobilized by other insertion sequences is unknown.

Intriguingly, most repeats are found inserted close to open reading frames (ORFs). Family members carry transcription initiation (2) and termination (6) signals, and full-length elements contain functional integration host factor sites (3). These observations suggest that *nemis* may impinge on gene expression at the transcriptional level. The finding that *N. meningitidis* mRNAs spanning *nemis* are processed by RNase III at hairpins that are formed by *nemis* TIRs (5, 9) allows one to hypothesize that *nemis* influence the level of expression of neighboring genes mostly by acting at the posttranscriptional level. *nemis* are (or have been) mobile elements, and their distribution in sequenced neisserial genomes is partly different (8, 9). Hence, before concluding on the base of whole-genome data (10, 12) that the expression of specific *N. meningitidis* genes could be regulated by *nemis*-mediated RNase III cleavage, we thought it important to verify the degree of conservation of *nemis* repeats in *N. meningitidis* populations. To this end, the position of a representative set of repeats spread throughout the genomes of the *N. meningitidis* MC58 (Fig. 1B) and Z2491 strains was monitored by PCR analyses in a variety of meningococci and in three strains of the apathogenic species *N. lactamica* (Table 1). The 57 elements selected are inserted close to either the start or the end of neisserial ORFs (Fig. 2). Ten nanograms of DNA from each strain was amplified by using the AmpliTaq DNA polymerase and 100 nanograms of 25- to 30-mers complementary to DNA segments flanking each repeat that were located 300 to 700 bp apart and were designed on the base of sequence conservation among fully sequenced *N. meningitidis* DNAs. Amplimers were resolved by electrophoresis on either 1.4% agarose or 6% polyacrylamide gels, and some were sequenced by the dideoxy chain termination method. In the FAM18 strain, whose sequence is available (http://www.sanger.ac.uk/Projects/N_meningitidis/seroC.shtml), the presence of *nemis* at sites of interest was monitored in silico by BLAST searches (1).

Data are summarized in Fig. 3. Size prediction of the PCR products allowed easy classification of most DNA regions as either "empty" (i.e., lacking *nemis*) or "filled" (i.e., containing *nemis*). Amplimers selected for sequence analysis differed essentially in the presence and/or absence of *nemis* DNA that was replaced in empty sites by TA, the target site duplicated at *nemis* termini. Two major types of variations emerge from our survey. At some sites, long and short *nemis* alternated among *N. meningitidis* strains (see repeats 5, 7, 28, 48, 49, 50, and 55 in Fig. 3). Such heterogeneity likely reflects recombination events that occurred in one strain or a few and eventually spread in neisserial populations by transformation-mediated DNA exchanges. Regions marked by the number sign in Fig. 3 matched neither empty nor filled sites in length and either contained or lacked *nemis* DNA, as shown by Southern and/or sequence analyses. Size identities exhibited by amplimers found in different strains (not shown) suggest that most of these alternative intergenic regions plausibly arose in one strain and were propagated to other clones by transformation.

On the whole, most of the tested repeats were fairly conserved among meningococci belonging to different serogroups and/or sequence types. Thirty-one of 57 elements were found at the same relative position in all the *N. meningitidis* strains analyzed; 11 of 57 were found in all but one or two strains. The degree of conservation of the remaining 15 repeats ranged from 70 to 30%. *nemis* were consistently more conserved in strains belonging to hypervirulent lineages than in other me-

* Corresponding author. Mailing address: Dipartimento di Biologia e Patologia Cellulare e Molecolare, "L. Califano," Universita degli Studi Napoli Federico II, Via S. Pansini 5, 80131 Naples, Italy. Phone: 0039-081-7462059. Fax: 0039-081-7703285. E-mail: dinocera@unina.it.

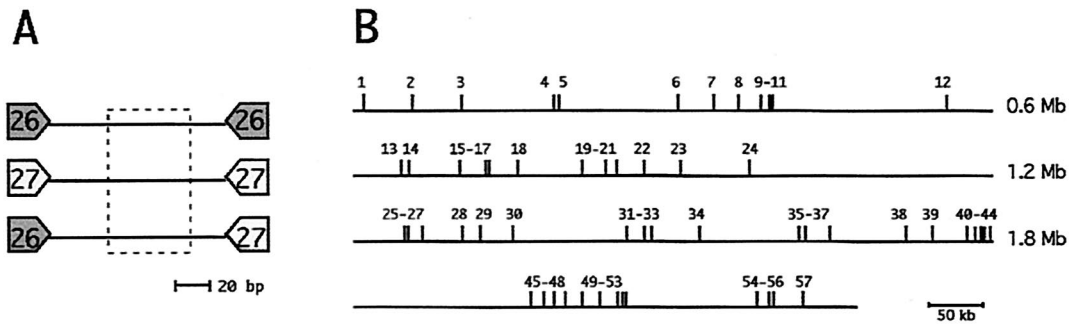

FIG. 1. (A) Organization of *nemis* repeats. *nemis* contain TIRs that are, including the TA dinucleotide target duplicated upon genomic insertion, either 26 or 27 bp. The 50-bp-long central region found only in long elements is boxed. (B) The relative chromosomal positions of *nemis* repeats 1 to 57 used in this study are shown.

ningococci (Fig. 3, bottom panel). Interestingly, the distribution of empty sites among strains is partly lineage specific. Thus, for example, *nemis* 19, 20, and 42 were not found in strains of the ET-5 complex, and *nemis* 19 was also absent in strains of the L1 cluster. *nemis* 55 was absent in lineage 4 strains; *nemis* 51 was absent in strains of both this lineage and the L1 cluster (Fig. 3).

The number of filled sites detected in *N. lactamica* genomes was surprisingly low. Only three repeats were found common to all the strains; 20 were conserved in one to two strains, but 34 were absent from all strains (Fig. 3). Data suggest that *nemis* may be approximately three times less abundant in *N. lactamica* than in *N. meningitidis.* According to in silico analyses, *nemis* are similarly underrepresented in *N. gonorrhoeae* strain F1090 (9), and it is intriguing that most *N. meningitidis nemis*-positive sites are *nemis*-negative sites in both *N. lactamica* and *N. gonorrhoeae* chromosomes (not shown). This would suggest that *nemis* arose in cells ancestral to the divergence of *Neisseriae* in pathogenic and apathogenic species and subsequently spread in a selective fashion in meningococci only.

Many *N. lactamica* regions, shown by Southern analyses to lack *nemis* DNA, are marked by the number sign. These regions not only differed in size from empty sites but varied also in length among strains (not shown) and represent either vestiges of *nemis*-positive intervals or never experienced the insertion of *nemis*. In either instance, it is intriguing that, while genes analyzed occupy the same position in *N. meningitidis* and *N. lactamica* and hence were detected by PCR, the corresponding intergenic regions evolved differently in the two species.

Taking into account that DNA exchanges between pathogenic and apathogenic *Neisseria* species are plausibly as fre-

TABLE 1. Strains used in this study

Species and strain	Serogroup	Epidemiological group	Origin	Source[a]
N. meningitidis				
BF2	B	ET-37 complex	Italy	a
93/4286	C	ET-37 complex	Norway	b
NGP165	B	ET-37 complex	Norway	b
FAM18	C	ET-37 complex	United States	World Wide Web
BZ169	B	ET-5 complex	The Netherlands	b
H44/76	B	ET-5 complex	Norway	b
MC58	B	ET-5 complex	Scotland	World Wide Web
205900	A	Subgroup IV-1	Italy	b
Z2491	A	Subgroup IV-1	The Gambia	World Wide Web
BL859	B	Lineage 3	Italy	c
BS845	B	Lineage 3	Italy	c
BL892	B	Lineage 3	France	d
BF9	B		Italy	a
B1940	B		Germany	e
BL947	B		France	d
NGF26	B		Norway	b
NGE31	B		Norway	b
NGH15	B		Norway	b
N. lactamica				
21			France	d
411			France	d
4627			France	d

[a] a, II Policlinico, Università di Napoli, Naples, Italy; b, IRIS, Chiron S.p.A, Siena, Italy; c, Istituto Superiore di Sanità, Rome, Italy; d, Institut Pasteur, Paris, France; and e, Bayerische Julius-Maximilians Universität, Würzburg, Germany.

ORF	gene function		*nemis*		gene function	ORF
11	murA, UDP-N-acetylgl. carboxyvinyltransf.	26 u	1	d 310	transmembrane transport protein	12
51	pilU, twitching motility protein	100 d	2	u 30	integral membrane protein	50
89	pykA, pyruvate kinase	173 d	3	u 34	outer membrane protein	88
189	integral membrane protein	33 u	4	u 222	conserved hypothetical protein	188
194	alanine symporter protein	135 d	5	u 35	gidA, regulatory protein	193
295	ffh, signal recognition particle protein	62 d	6	d 114	dsbA, thiol:disulphide interchange protein	294
329	pilF, type IV assembly protein	46 u	7	d 289	conserved hypothetical protein	328
353	hypothetical protein	67 u	8	d -4	hypothetical protein	352
380	anaerobic transcriptional regulator	35 u	9	u 122	hemN, coproporphyrinogen III oxidase	379
388	integral membrane transport protein	28 u	10	u 204	ABC transporter ATP-binding protein	387
390	mapA, maltose phosphorylase	45 u	11	d 35	galM, aldose 1-epimerase	389
534	membrane protein	45 u	12	u 390	transmembrane hexose transporter	535
614	putative oxidoreductase	13 d	13	d 76	amtB, probable ammonium transporter	615
619	conserved hypothetical protein	62 d	14	d 21	phosphoglycolate phosphatase	620
671	putative malate oxidoreductase	63 d	15	u 330	tetraacyldisaccharide 4'-kinase	672
698	unknown protein	40 d	16	u 40	tryptophan synthase, beta subunit	699
699	tryptophan synthase, beta subunit	14 d	17	d 59	IgA endopeptidase	700
723	rplT, 50S ribosomal protein L20	114 d	18	u 125	pheS, phenylalanyl-tRNA synthetase a chain	724
785	recB, exodeoxyribonuclease V beta chain	28 u	19	d 50	unknown protein	786
811	murB, UDP-N-acetylenolpyruvoylgl. reduct.	53 u	20	d 19	transmembrane efflux protein	812
823	adk, adenylate kinase	320 d	21	u 49	pyrF, orotidine 5'-phosphate decarboxylase	824
845	PhoH-related protein	39 u	22	d -57	LPS biosynthesis related protein	846
885	dnaB, putative replicative DNA helicase	126 d	23	u 26	FimT, fimbrial protein	886
956	sucB, dihydrolipoamide succinyltransf. E2 comp.	49 d	24	u 209	lpdA, dihydrolipoamide dehydrogen. E3 comp	957
1241	cca, tRNA nucleotidyltransferase	21 d	25	u 96	hypothetical protein	1242
1244	rpe, ribulose phosphate epimerase	23 u	26	d 184	hypothetical protein	1245
1257	site-specific DNA methylase, pseudogene	18 u	27	d 123	unknown protein	1258
1287	iron sulphur binding protein	17 u	28	d 101	nrdB, ribonucleoside-diphosphate reductase	1288
1302	hip, integration factor beta subunit	112 d	29	d 84	putative transcriptional regulator	1303
1331	uvrB, excinuclease ABC subunit B	124 u	30	d 99	prc, carboxy-terminal processing protease	1332
1421	nifR3 protein	150 u	31	u 62	ATP-dependent RNA helicase	1422
1433	putative lipoprotein	50 u	32	d 3	phospholipase D family protein	1434
1438	hypothetical iron-sulphur protein	57 d	33	d 19	purE, phosphoribosylaminoimidazole carboxyl.	1439
1474	possible tautomerase	90 u	34	u 99	possible periplasmic protein	1475
1558	dgk, diacylglycerol kinase	25 u	35	d 132	gshB, glutathione synthetase	1559
1563	transcriptional regulator [GntR-family]	102 u	36	d 315	hypothetical protein	1564
1584	unknown protein	88 d	37	u 34	transcriptional regulator [MarR-family]	1585
1650	transcriptional regulator [AsnC-family]	43 u	38	u 174	alr, alanine racemase	1651
1669	hemO, haem utilisation protein	62 u	39	u 66	integral membrane protein	1670
1695	hypothetical protein	-89 d	40	d 139	hypothetical protein	1699
1706	unknown protein	-5 d	41	u 27	integral membrane ion transporter	1707
1710	gdhA, glutamate dehydrogenase	96 d	42	d 38	transcriptional regulator [GntR family]	1711
1711	lrp, transcriptional regulator [GntR family]	42 u	43	u 234	integral membrane protein	1712
1716	mtrC, membrane fusion protein	57 u	44	u 192	transcriptional regulator [mtrR family]	1717
1861	prmA, ribosomal protein L11	18 u	45	d 62	accC, acetyl-CoA carboxylase	1862
1877	prolyl endopeptidase	28 u	46	d 65	argA, acetylglutamate synthase	1876
1883	hypothetical protein	101 d	47	d 30	ferric siderophore receptor protein	1882
1897	leuS, leucyl-tRNA synthetase	159 d	48	u 23	drg, type II restriction endonuclease	1896
1918	fabD, malonyl CoA-acyl c. p. transacylase	74 u	49	d -44	integral membrane protein	1917
1933	atpC, ATP synthase epsilon chain	146 d	50	u 103	glyQ, glycyl-tRNA synthetase alpha chain	1932
1953	sspA, stringent starvation protein A	42 u	51	d 86	hypothetical protein	1954
1956	rpmE, 50S ribosomal protein L31	81 d	52	u 66	cad, cadmium resistance protein	1955
1957	putative acetyltransferase	21 u	53	u 109	rpmE, 50S ribosomal protein L31	1956
2056	rpsI, 30S ribosomal protein S9	105 d	54	d 43	transcriptional regulator [metR family]	2055
2066	tldD, regulatory function	46 u	55	u 237	conserved hypothetical protein	2067
2071	thiG, thiamine biosynthesis protein	14 u	56	d 43	unknown protein	2070
2104	mafA, adhesin	45 u	57	d 61	pyrH, uridylate kinase	2103

FIG. 2. *nemis* were analyzed. Flanking ORFs are numbered as in the *N. meningitidis* MC58 strain (12). The distance in base pairs separating *nemis* from upstream (u) and downstream (d) ORFs is given.

quent as those occurring between meningococci (7), the asymmetry in the partition of *nemis*-positive and *nemis*-negative intergenic regions between *N. meningitidis* and *N. lactamica* strains is striking. This permits the hypothesis that the persistence of *nemis* DNA at specific chromosomal sites may be functional to meningococci.

Many *N. meningitidis* genes listed in Fig. 2 have a functional role. Some encode either transcriptional regulators (ORFs 380, 1585, 1650, and 1711) or regulatory proteins (ORFs 193, 1953, and 2066); others encode proteins known to be involved in pathogenesis (ORFs 329, 700, and 886) or shown to be essential for the development of bacteremia in the rat (ORFs 1422, 1558, and 1671) (12). Transcripts spanning the underlined ORFs are processed at *nemis* RNA hairpins (5, 9). The same holds for mRNAs spanning the additional ORFs listed in Fig. 2 (unpublished results). The hypothesis that *nemis*-mediated RNA processing may have relevance in the life of meningococci as pathogens is strengthened by the observation that RNase III, while dispensable for viability, is crucial for the survival of meningococci in the infected host (11).

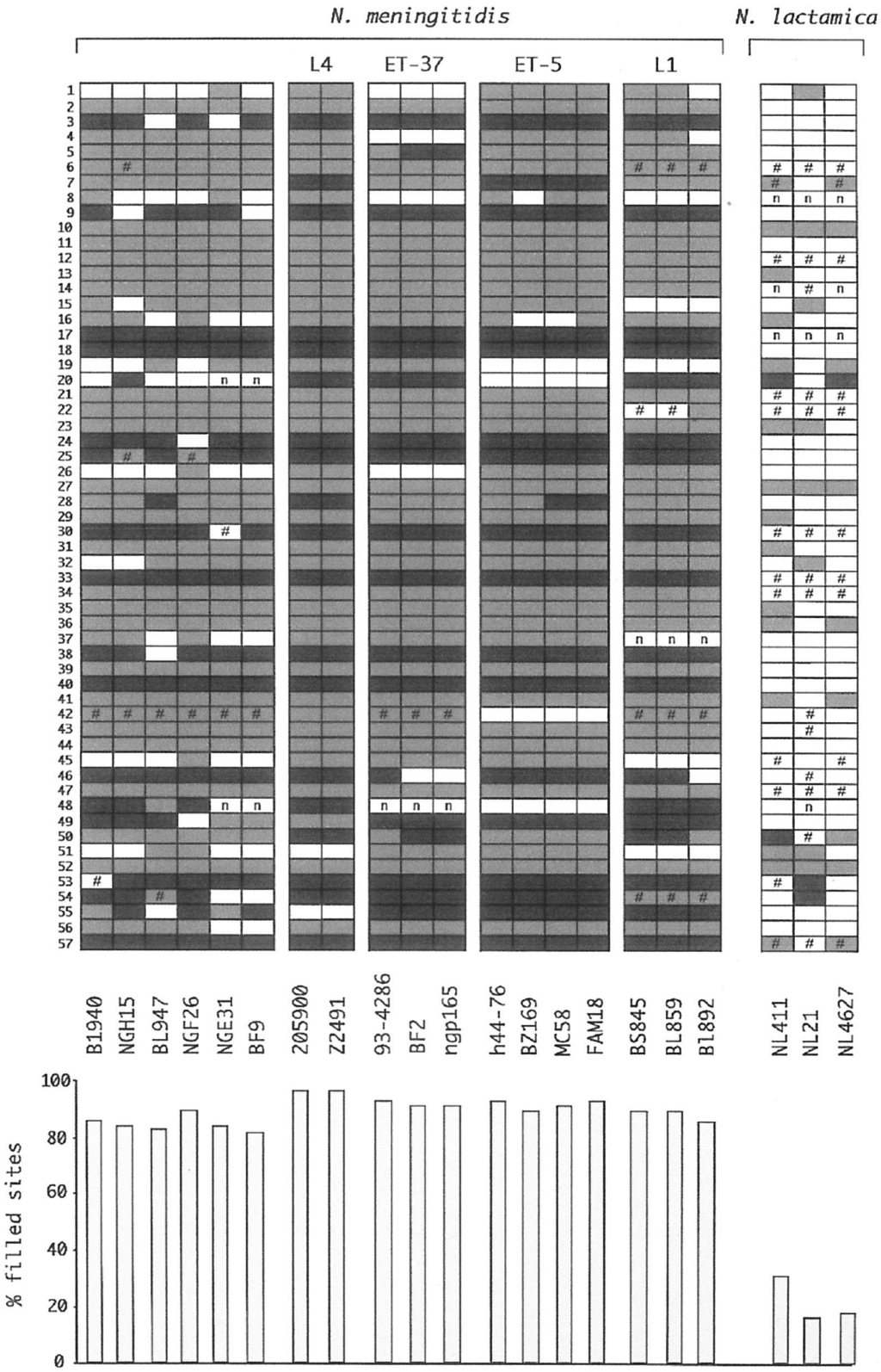

FIG. 3. Conservation of *nemis* repeats in neisserial chromosomes. The distribution of the 57 *nemis* elements listed in Fig. 2 is diagrammed as follows: empty and filled boxes represent intergenic chromosomal regions lacking and containing *nemis* elements, respectively. The presence of long and short *nemis* is marked by light and dark grey filling, respectively. The number sign represents regions differing in size from either filled or empty sites. Regions for which reliable PCR amplification signals could not be obtained, regardless of changes in either PCR settings or primer pairs, are labeled by n. The relative abundance of *nemis*-positive regions within each strain is highlighted in the histogram at the bottom.

We thank Caterina Pagliarulo for providing us with neisserial strains and Giustina Silvestro for help in biocomputing analyses.

This work was partly supported by a grant of program PRIN 2002 of MIUR to P.P.D.N.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Black, C. G., J. A. Fyfe, and J. K. Davies.** 1995. A promoter associated with the neisserial repeat can be used to transcribe the *uvrB* gene from *Neisseria gonorrhoeae*. J. Bacteriol. **177:**1952–1958.
3. **Buisine, N., C. M. Tang, and R. Chalmers.** 2002. Transposon-like Correia elements: structure, distribution and genetic exchange between pathogenic Neisseria species. FEBS Lett. **522:**52–58.
4. **Correia, F. F., S. Inouye, and M. Inouye.** 1988. A family of small repeated elements with some transposon-like properties in the genome of Neisseria gonorrhoeae. J. Biol. Chem. **263:**12194–12198.
5. **De Gregorio, E., C. Abrescia, M. S. Carlomagno, and P. P. Di Nocera.** 2002. The abundant class of *nemis* repeats provides RNA substrates for ribonuclease III in Neisseriae. Biochim. Biophys. Acta **1576:**39–44.
6. **Francis, F., S. Ramirez-Arcos, H. Salimnia, C. Victor, and J. R. Dillon.** 2000. Organization and transcription of the division cell wall (dcw) cluster in Neisseria gonorrhoeae. Gene **251:**141–151.
7. **Linz, B., M. Schenker, P. Zhu, and M. Achtman.** 2000. Frequent interspecific genetic exchange between commensal neisseriae and *Neisseria meningitidis*. Mol. Microbiol. **36:**1049–1058.
8. **Liu, S. V., N. J. Saunders, A. Jeffries, and R. F. Rest.** 2002. Genome analysis and strain comparison of Correia repeats and Correia repeat-enclosed elements in pathogenic *Neisseria*. J. Bacteriol. **184:**6163–6173.
9. **Mazzone, M., E. De Gregorio, A. Lavitola, C. Pagliarulo, P. Alifano, and P. P. Di Nocera.** 2001. Whole-genome organization and functional properties of miniature DNA insertion sequences conserved in pathogenic Neisseriae. Gene **278:**211–222.
10. **Parkhill, J., M. Achtman, K. D. James, S. D. Bentley, et al.** 2000. Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. Nature **404:**502–506.
11. **Sun, Y. H., S. Bakshi, R. Chalmers, and C. M. Tang.** 2000. Functional genomics of Neisseria meningitidis pathogenesis. Nat. Med. **6:**1269–1273.
12. **Tettelin, H., N. J. Saunders, J. Heidelberg, A. C. Jeffries, et al.** 2000. Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. Science **287:**1809–1815.

Editor: J. N. Weiser