Identification of Vaccine Candidates Against Serogroup B Meningococcus by Whole-Genome Sequencing

Mariagrazia Pizza, *et al.*

Science **287**, 1816 (2000);
DOI: 10.1126/science.287.5459.1816

Identification of Vaccine Candidates Against Serogroup B Meningococcus by Whole-Genome Sequencing

Mariagrazia Pizza,[1]* Vincenzo Scarlato,[1]* Vega Masignani,[1] Marzia Monica Giuliani,[1] Beatrice Aricò,[1] Maurizio Comanducci,[1] Gary T. Jennings,[1] Lucia Baldi,[1] Erika Bartolini,[1] Barbara Capecchi,[1] Cesira L. Galeotti,[1] Enrico Luzzi,[1] Roberto Manetti,[1] Elisa Marchetti,[1] Marirosa Mora,[1] Sandra Nuti,[1] Giulio Ratti,[1] Laura Santini,[1] Silvana Savino,[1] Maria Scarselli,[1] Elisa Storni,[1] Peijun Zuo,[1] Michael Broeker,[2] Erika Hundt,[2] Bernard Knapp,[2] Eric Blair,[3] Tanya Mason,[3] Hervé Tettelin,[3] Derek W. Hood,[4] Alex C. Jeffries,[4] Nigel J. Saunders,[4] Dan M. Granoff,[5] J. Craig Venter,[3] E. Richard Moxon,[4] Guido Grandi,[1] Rino Rappuoli[1]†

Neisseria meningitidis is a major cause of bacterial septicemia and meningitis. Sequence variation of surface-exposed proteins and cross-reactivity of the serogroup B capsular polysaccharide with human tissues have hampered efforts to develop a successful vaccine. To overcome these obstacles, the entire genome sequence of a virulent serogroup B strain (MC58) was used to identify vaccine candidates. A total of 350 candidate antigens were expressed in *Escherichia coli*, purified, and used to immunize mice. The sera allowed the identification of proteins that are surface exposed, that are conserved in sequence across a range of strains, and that induce a bactericidal antibody response, a property known to correlate with vaccine efficacy in humans.

Meningococcal meningitis and sepsis are devastating diseases that can kill children and young adults within hours despite the availability of effective antibiotics. The diseases are caused by *Neisseria meningitidis*, a Gram-negative, capsulated bacterium that has been classified into five major pathogenic serogroups (A, B, C, Y, and W135) on the basis of the chemical composition of distinctive capsular polysaccharides (*1*). The reported annual incidence of meningococcal disease varies from 0.5 to 10 per 100,000 persons (*2, 3*). However, during epidemics the incidence can rise above 400 per 100,000 (*2, 3*). The case fatality rate ranges from 5% to 15%, and up to 25% of survivors are left with neurological sequelae (*4*). Recently, 2600 and 5606 cases were reported annually in the United States and Europe, respectively (*4*).

In the 1960s, vaccines consisting of purified polysaccharide antigens were developed against four (A, C, Y, and W135) of the five pathogenic serogroups (*1*). These vaccines are highly effective in adults but are not efficacious in infants and young children, the age groups mostly exposed to disease. Second-generation glycoconjugate vaccines that are also effective in infants and children are now in the later phases of development (*2, 5*).

Currently, there are no vaccines available for prevention of serogroup B *N. meningitidis* (MenB) disease, which is responsible for 32% of all meningococcal disease in the United States and for 45% to >80% of the cases in Europe (*6*). The use of capsular polysaccharide as the basis of a vaccine for prevention of MenB diseases has been problematic. The MenB capsular polysaccharide is identical to a widely distributed human carbohydrate [α(2→8)*N*-acetyl neuraminic acid or polysialic acid], which, being a self-antigen, is a poor immunogen in humans. Furthermore, use of this polysaccharide in a vaccine may elicit autoantibodies (*7*). An alternative approach to vaccine development is based on surface-exposed proteins contained in outer membrane vesicles (OMVs). These vaccines have been shown both to elicit serum bactericidal antibody responses and to protect against developing meningococcal disease in clinical trials (*8*). However, the limitation of OMV vaccines is that the major protein antigens show sequence and antigenic variability and, although they induce protective antibodies against the homologous strain, they fail to induce protection against heterologous strains (*9*). With the recent exception of NspA (*10*), all the surface-exposed proteins described during the past three decades have in common the drawback of antigenic variability (*2*).

To identify potential vaccine candidates, we determined the genome sequence of the virulent strain MC58 [see (*11*)]. While the sequencing project was in progress, unassembled DNA fragments were analyzed to identify open reading frames (ORFs) that potentially encoded novel surface-exposed or exported proteins (*12*).

We identified 570 such ORFs and, by means of the polymerase chain reaction (PCR), we amplified and cloned the DNA sequences of these hypothetical genes in *Escherichia coli* to express each polypeptide as either His-tagged or glutathione S-transferase (GST) fusion proteins (*13*). We obtained successful expression with 350 ORFs (61%). More specifically, 70 predicted lipoproteins, 96 predicted periplasmic proteins, 87 predicted inner membrane proteins, and 45 predicted outer membrane proteins; there were 52 proteins with uncertain prediction. Proteins with more than one hydrophobic trans-membrane domain had the highest rate of expression failure. The recombinant proteins were purified and used to immunize mice (*13*). Immune sera were then tested in enzyme-linked immunosorbent assay (ELISA) and fluorescence-activated cell sorter (FACS) analyses to detect proteins that were present on the surface of a set of MenB strains selected to represent the diversity of invasive strains within the natural population of this species (*14*). In addition, we tested the immune sera for bactericidal activity (*14*) because this assay correlates with protection in humans (*15*). Of the 85 proteins found to be strongly positive in at least one of the above assays, we selected for further studies seven representative proteins (genome-derived *Neisseria* antigens; GNA) that were positive in all three assays (Table 1) and whose genes were not predicted to be phase variable (*11*). Each of the proteins raised an immune response that induced complement-mediated bactericidal activity. In the case of proteins GNA33 and GNA2132, the resulting bactericidal titers were similar in magnitude to that induced by OMV, which is known to confer protection in humans against homologous strains (*9*).

To test the suitability of these proteins as candidate antigens for conferring protection against different MenB strains and not just against the homologous strain, we used a collection of strains isolated worldwide and

[1]IRIS, Chiron S.p.A., Via Fiorentina 1, 53100 Siena, Italy. [2]Chiron Behring, Post Office Box 1630, D-35006 Marburg, Germany. [3]The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA. [4]Institute of Molecular Medicine, Oxford University, Department of Paediatrics, John Radcliffe Hospital, Headington, Oxford OX3 9DS, UK. [5]Children's Hospital Oakland Research Institute, 5700 Martin Luther King Jr. Way, Oakland, CA 94609, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: Rino_Rappuoli@biocine.it

over many years to investigate whether the new candidate molecules were conserved and accessible to antibodies. Our aim was to select strains representative of the diversity found in natural populations of MenB. We used a phylogenetic tree from 107 strains constructed by multilocus enzyme electrophoresis (MLEE) and validated by multilocus sequence typing (MLST) to select 22 representative, disease-associated MenB strains (*16*–*19*) (Fig. 1A). In addition, in the analysis we included three strains of *N. meningitidis* serogroup A; two strains of serogroup C; one strain each of serogroups Y, X, Z, and W135; three strains of *Neisseria gonorrhoeae*; and one strain each of *Neisseria cinerea* and *Neisseria lactamica* (*19*). PCR products from the seven genes were detected in each of the 31 strains of *N. meningitidis* and most were also found in the *N. gonorrhoeae* isolates. In *N. lactamica* or *N. cinerea,* PCR was often negative, but we detected similar sequences by Southern blotting (Table 2). We determined the nucleotide sequence of the seven genes and *porA* in all the *N. meningitidis* and *N. gonorrhoeae* strains (Fig. 2) (GenBank accession numbers: AF226325–AF226572, *N. meningitidis*; AF235143–AF235159, *N. gonorrhoeae*). PorA was included in the analysis because it is a well-characterized membrane-associated protein that displays sufficient sequence diversity (*2*, *20*) to render it unsuitable as an effective vaccine. As expected, PorA showed regions of substantial amino acid sequence divergence in MenB as well as in the other meningococcal serogroups (Fig. 2) (*19*). Among the antigens, GNA992 and GNA2132 had many hypervariable regions, located mostly in the amino-terminal half of the molecules, which suggests that these proteins, like PorA, might induce strain-specific immunity. In marked contrast, GNA33, GNA1162, GNA1220, GNA1946, and GNA2001 showed 99.2% ± 0.7%, 99.7% ± 0.7%, 99.7% ± 0.3%, 99.4% ± 0.3%, and 99.7% ± 0.3% average amino acid identity to the MC58 sequence, respectively, within the 31 *N. meningitidis* strains analyzed. The results suggest that these proteins may induce immunity against most strains of MenB and, possibly, against the other pathogenic strains of *N. meningitidis.* GNA33, GNA1162, GNA1220, and GNA1946 showed 95.8% ± 0.2%, 96.5% ± 0.4%, 99% ± 0.0%, and 97.6% ± 0.0% identity, respectively, to *N. gonorrhoeae*; therefore, they are also candidate vaccine antigens against this pathogen.

To further characterize these target antigens as vaccine candidates, we studied surface expression by immunoblotting outer membrane preparations and we studied accessibility of antibodies with encapsulated bacteria in a whole-cell ELISA (*14*). These data provided evidence that each of the seven proteins was expressed and available for an-

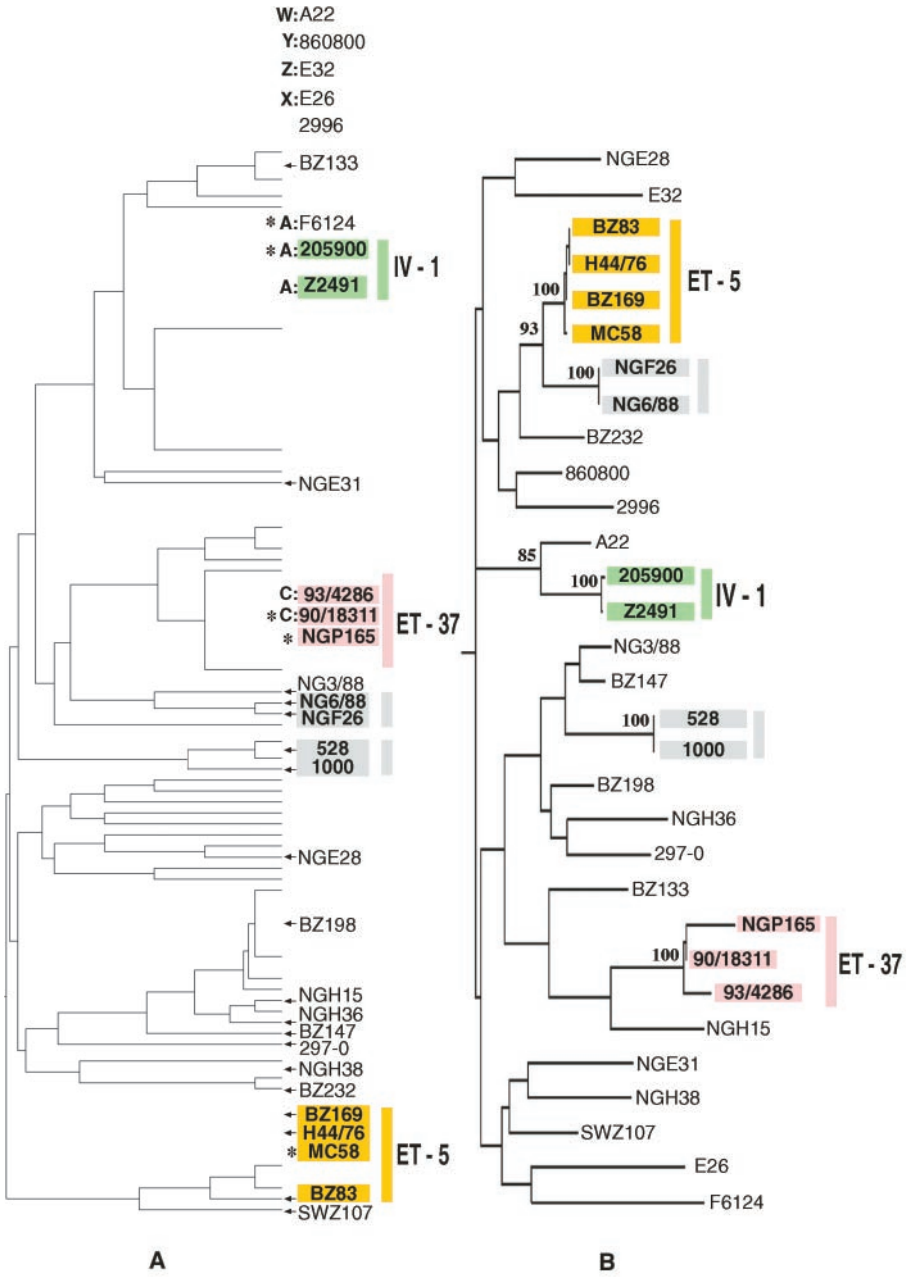

Fig. 1. (**A**) Dendrogram showing genetic relationship among 107 *N. meningitidis* strains based on MLST analysis of six gene fragments [adapted from Maiden *et al.* (*16*)]. The dendrogram was used to select strains representative of serogroup B meningococcus population (arrows). Five additional strains, for which genetic assignment to hypervirulent lineages was independently determined by Wang *et al.* (*18*), Seiler *et al.* (*17*), and Virji *et al.* (*19*), are superimposed in the dendrogram and indicated by asterisks. In addition to the 22 strains of MenB, three strains of MenA, two strains of MenC, and one strain each of Men Y, X, Z, and W135 were used. These are indicated with a boldface letter before the name. When phylogenetic data were not available, the strains were reported above the figure outside the tree. The hypervirulent clusters ET-5, ET-37, and IV-1 are indicated by colored vertical bars. (**B**) Dendrogram of *N. meningitidis* strains obtained from the conserved genes reported in Table 2. Phylogenetic analysis based on the new genes provided a dendrogram that clusters the hypervirulent strains, in agreement with the results of MLEE and MLST reported in (A). Colored and gray bars indicate strains that cluster with 100% bootstrap support in agreement with MLST analysis. Numbers at the base of each node are bootstrap scores (only those >80% are reported). Gene sequences from different strains were aligned with the program PileUp of the GCG package. Phylogenetic analysis was done with the neighbor-joining algorithm (*24*) as implemented in the program NEIGHBOR of the PHYLIP package. Pairwise distances were calculated by using the Kimura-two parameter (*25*) on the 31 *N. meningitidis* strains. We excluded the NH_2-terminal region of GNA992, the entire GNA2132, and the tandemly repeated regions of GNA2001 from the analysis. We allowed a total of 1000 bootstrap replicates to evaluate the level of support. We confirmed clustering of the hypervirulent strains by maximum parsimony analysis.

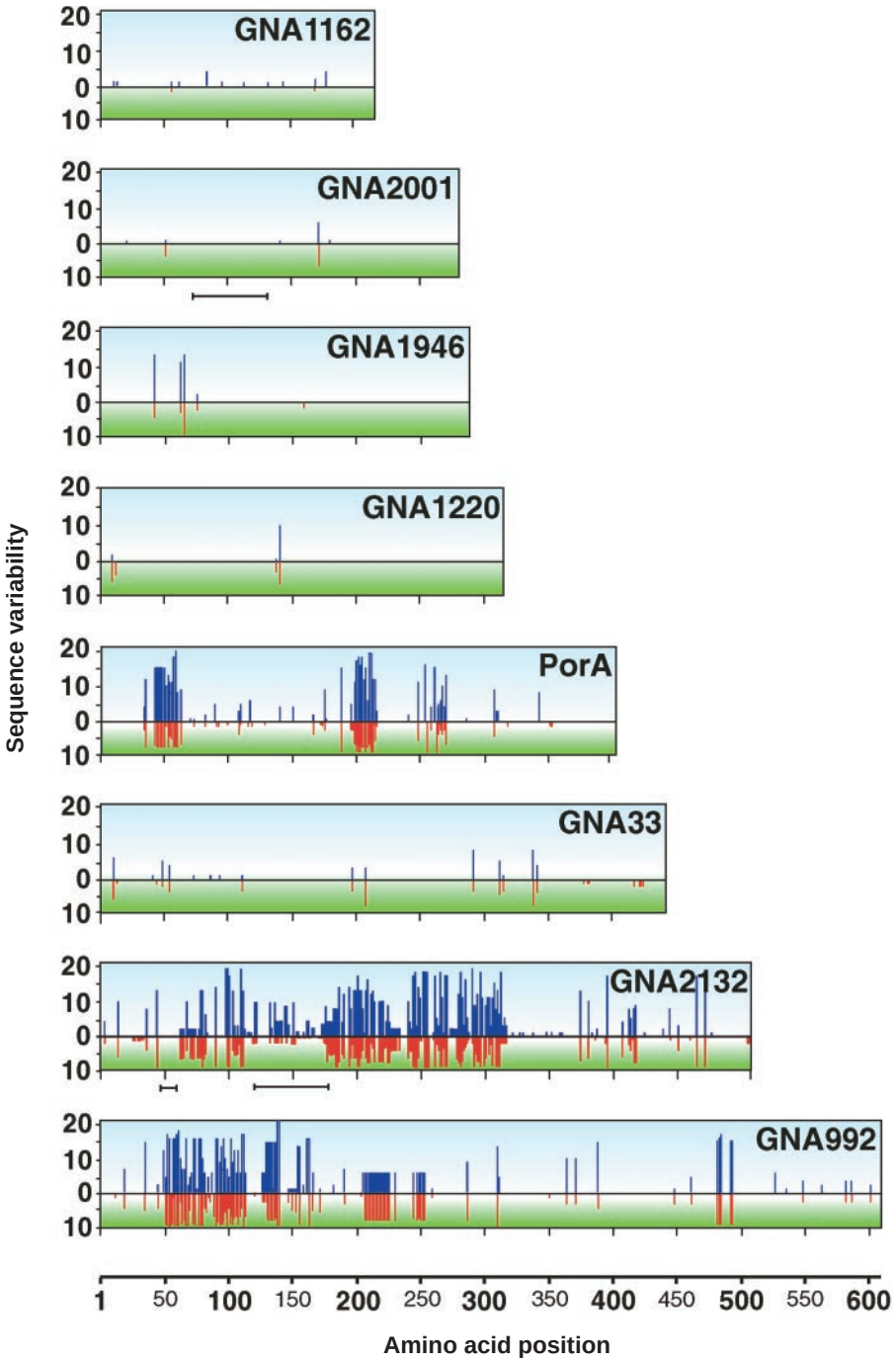

Fig. 2. Schematic representation of amino acid sequence variability within *N. meningitidis* of the seven antigens reported in Table 1 and of PorA. Abscissa, amino acid position; ordinate, number of strains analyzed. Line 0 represents sequence of the MC58 reference strain. Amino acid differences from the sequence of MC58 within the 22 strains of MenB are indicated by blue lines above the 0. Amino acid differences within the nine *N. meningitidis* strains from serogroups A, C, Y, X, Z, and W135 are indicated by red lines below the 0. Height of blue and red lines represents the number of strains with amino acid changes. Variable regions appear as blue and red peaks. Bars below GNA2001 and GNA2132 represent segments that are missing from some strains.

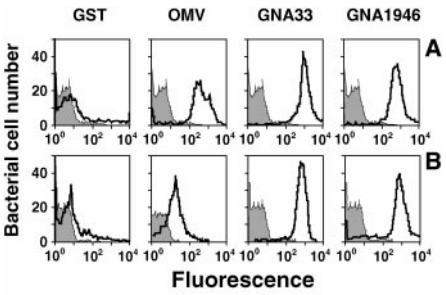

Fig. 3. FACS analysis showing binding of polyclonal OMV, GNA33, and GNA1946 antisera to the ethanol-treated homologous 2996 (**A**) and heterologous BZ232 (**B**) strains. Gray profiles show binding of preimmune sera; white profiles show binding of immune sera. Negative controls include sera of mice immunized with GST.

tibody binding in the presence of the polysaccharide capsule coating the bacteria. Surface exposure also was determined by FACS analysis, measuring binding of antibodies to bacterial strains whose capsule had been permeabilized by treatment with ethanol (*14*). We found the seven proteins to be expressed and surface exposed in all 31 *N. meningitidis* strains tested. The hyperimmune sera prepared against the conserved proteins GNA33 and GNA1946 recognized equally the homologous MenB strain 2996 (*13*) and the heterologous MenB strain BZ232 (Fig. 3) (*17*). Conversely, antiserum prepared against the OMV from strain 2996, which is known to confer strain-specific protection (*9*), bound strongly to the homologous strain but weakly to the heterologous strain. As expected, antisera against GNA33 and GNA1946 showed complement-mediated bacteriolysis against strains 2996 and BZ232. Both sera also had bactericidal activity against three unrelated, heterologous strains of MenB, for which a common source of suitable human complement was available (*14*).

The finding of several conserved surface-exposed antigens in *N. meningitidis* is surprising, because in three decades of studies, with one exception (*10*), only antigenically variable proteins had been described. To gain more information about the genetic diversity of the newly discovered conserved proteins, we evaluated the frequency of recombination of their genes by means of the homoplasy test (*19, 21*). We obtained an average value of 0.11, a result indicating a low level of recombination, similar to that previously obtained for *Neisseria* housekeeping genes (*19*). The result was substantiated by a phylogenetic analysis with the neighbor-joining algorithm. The outcome of this analysis was a dendrogram clustering the hypervirulent strains of complexes ET-37, ET-5, and subgroup IV-1 (Fig. 1B), in agreement with the dendrogram generated by MLST and MLEE through housekeeping genes (Fig. 1A) (*16–18*).

Thus, the newly identified surface proteins, apart from being surface exposed and accessible to antibodies, are surprisingly conserved, as previously observed in housekeeping proteins. In the case of protein GNA33, the presence of a predicted enzymatic activity of importance in peptidoglycan metabolism may explain the conservation of this gene (*22*). Alternatively, because the new proteins are less abundant than major outer membrane proteins (*23*), they may be poorly immunogenic during infection and therefore subject to lower selection pressures from the host's immune system. Overall, these data confirm that the newly discovered proteins behave differently from the other surface antigens known so far and are likely to confer

Table 1. Properties of the proteins.

Antigen (length, amino acids)	Remarks/similarities*	FACS	ELISA	Serum bactericidal activity (SBA)
GNA33 (441)	Lipoprotein/similar to *E. coli* membrane-bound lytic transglycosylase A (MltA) of *E. coli* and of *Synechocystis* sp. (*22*)	++++†	13,000	1/16,000‡
GNA992 (591)	Outer membrane protein/similar to Hsf and Hia of *Haemophilus influenzae* and FhaB of *Bordetella pertussis* (*26*)	+++	2,750	1/256
GNA1162 (215)	Lipoprotein/no significant similarities	++	1,270	1/4
GNA1220 (315)	Membrane protein/contains a stomatin-like domain	+++	1,000	1/256
GNA1946 (287)	Lipoprotein/similar to HlpA of *H. influenzae*, belongs to the NlpA family of lipoproteins (*27*)	+++	13,100	1/32
GNA2001 (251)	Outer membrane protein/similar to P60 invasion-associated extracellular proteins (*28*)	++	500	1/512
GNA2132 (488)	Lipoprotein/low similarity to transferrin binding proteins	++	1,700	1/16,000
GST§	—	−	<50	<1/4
OMV§	Mixture of proteins containing mainly PorA	++++	260,000	1/32,000

*Homology searches have been performed against nonredundant protein databases by means of the Φ-BLAST algorithm (National Center for Biotechnology Information). Hits with an *E* score of $<10^{-20}$ and an assigned biological role are reported. †Titers −, +, ++, +++, and ++++ indicate no difference with preimmune serum or differences in fluorescence of 0.5, 1, 1.5, and 2 or more orders of magnitude, respectively. ‡Serum bactericidal activity (SBA) was evaluated with pooled baby rabbit serum as complement source. Titers were expressed as the reciprocal of the serum dilution yielding ≥50% bacterial killing. §Sera against GST and OMV were used as negative and positive controls, respectively.

Table 2. Presence of genes in *Neisseria*.

Gene	*N. meningitidis* B (22 strains)	*N. meningitidis* A,C,Y,X,Z,W135 (9 strains)	*N. lactamica* (1 strain)	*N. cinerea* (1 strain)	*N. gonorrhoeae* (3 strains)
*gna*33	+	+	+	+	+
*gna*992	+*	+	+/−†	+/−	−
*gna*1162	+	+	+	+	+
*gna*1220	+‡	+	+/−	+/−	+
*gna*1946	+	+	+	+/−	+
*gna*2001	+	+	+	+/−	+
*gna*2132	+	+	+	−	+
*por*A	+	+	+	−	+

*In strains NG6/88 and NGF26 the start codon is 222 bases downstream from the starting codon in the other strains. †+/− indicates a negative PCR but positive Southern blotting. ‡In strain BZ133 a deletion of 31 nucleotides causes a frameshift in this gene.

protection against both homologous and heterologous MenB strains.

Genomic studies of bacterial pathogens have greatly increased our knowledge. However, they have not yet led to advances in therapeutic or preventive measures. Here we have described how availability of the genome sequence of an important bacterial pathogen allowed us to succeed in identifying conserved surface-exposed proteins from MenB. Furthermore, we have shown, by comparing sequences of the candidate gene in a selection of strains representing the known diversity of the species, that a single index sequence can be used as a reference to address potential antigenic variability very early in vaccine development.

In addition to proving the potential of the genomic approach, by identifying highly conserved proteins that induce bactericidal antibodies, we have provided candidates that will be the basis for clinical development of a vaccine against an important pathogen. This vaccine is likely to elicit cross protection not only against group B *N. meningitidis* but also against other serogroups and species of pathogenic *Neisseria*.

References and Notes

1. E. C. Gotschlich, T. Y. Liu, M. S. Artenstein, *J. Exp. Med.* **129**, 1349 (1969); E. C. Gotschlich, I. Goldschneider, M. S. Artenstein, *J. Exp. Med.* **129**, 1367 (1969).
2. W. D. Zollinger, in *New Generation Vaccines*, M. M. Levine, G. C. Woodrow, J. B. Kaper, G. S. Cobon, Eds. (Dekker, New York, 1997), p. 469; H. Peltola, *Drugs* **55** (no. 3), 347 (1998).
3. A. Schuchat *et al.*, *N. Engl. J. Med.* **337**, 970 (1997).
4. A. Naess *et al.*, *Acta Neurol. Scand.* **89**, 139 (1994); H. S. Jafari, B. A. Perkins, J. D. Wenger, *Morb. Mortal. Wkly Rep.* **46** (no. RR-5), 1 (1997); M. Connolly and N. Noah, *Surveillance of Bacterial Meningitis in Europe 1996* (King's College, London, 1997), p. 6.
5. D. M. Granoff, B. Forrest, R. Rappuoli, *Int. J. Infect. Dis.* **1**, 152 (1997).
6. *Morb. Mortal. Wkly Rep.* **46** (no. 54), 45; R. J. Scholten *et al.*, *Clin. Infect. Dis.* **16**, 237 (1993).
7. J. Hayrinen *et al.*, *J. Infect. Dis.* **171**, 1481 (1995); J. Finne, D. Bitter-Suermann, C. Goridis, U. Finne, *J. Immunol.* **138**, 4402 (1987).
8. G. Bjune *et al.*, *Lancet* **338**, 1093 (1991); J. C. de Moraes *et al.*, *Lancet* **340**, 1074 (1992); J. W. Tappero *et al.*, *JAMA* **281**, 1520 (1999).
9. J. T. Poolman, *Infect. Agents Dis.* **4**, 13 (1995).
10. D. Martin, N. Cadieux, J. Hamel, B. R. Brodeur, *J. Exp. Med.* **185**, 1173 (1997).
11. H. Tettelin *et al.*, *Science* **287**, 1809 (2000).
12. Sequence analysis: A primary screening for coding capacity of the *N. meningitidis* genome sequence (*11*) was carried out on DNA segments or contigs with databases and computer programs included in the Wisconsin package version 10.0, Genetics Computer Group (GCG), Madison, WI. First we used BLASTX to identify two classes of DNA segments with potential coding regions. The class of ORFs coding for known cytoplasmic functions was not further investigated, whereas the other classes of coding regions were selected for further analyses. A second screening step aimed at identifying putative proteins with a cellular localization spanning from the inner membrane to outside the bacterium was applied. Φ-BLAST, FASTA, MOTIFS, FINDPATTERNS, and PSORT (http://psort.nibb.ac.jp), as well as the ProDom, Pfam, and Blocks databases, were used to predict features typical of surface-associated proteins such as transmembrane domains, leader peptides, homologies to known surface proteins, lipoprotein signature, outer membrane anchoring motives, and host cell binding domains such as RGD. The screening yielded 570 ORFs with the desired features, which were screened for conservation against *N. meningitidis* serogroup A and *N. gonorrhoeae* partial genome sequences available at The Sanger Centre (http://www.sanger.ac.uk/Projects/N_meningitidis/) and the Advanced Center for Genome Technology at The University of Oklahoma (http://www.genome.ou.edu/gono.html), respectively. This analysis indicated that 98.8% of the selected *N. meningitidis* ORFs are also found and conserved in serogroup A and 95.3% are found and conserved in *N. gonorrhoeae*.
13. ORFs were amplified by PCR on chromosomal DNA from strain 2996 [P. van der Ley and J. T. Poolman, *Infect. Immun.* **60**, 3156 (1992)], with synthetic oligonucleotides used as primers. The amplified DNA fragments were cloned into pET-21b+ vector (Novagen) and pGEX-KG vector [K. L. Guan and J. E. Dixon, *Anal. Biochem.* **192**, 262 (1991)] to express the proteins as COOH-terminus His-tagged and NH_2-terminus GST fusions, respectively. We evaluated expression of recombinant proteins according to the appearance of bands in SDS–polyacrylamide gel electrophoreses. We purified fusion proteins by affinity chromatography on Ni^{2+}-conjugated chelating fast-flow Sepharose (Pharmacia) or glutathione–Sepharose 4B resin (Pharmacia). Proteins expressed as insoluble inclusion bodies were solubilized with guanidinium and/or urea and renatured after purification [J. E. Coligan, B. M. Dunn, H. L. Ploegh, D. W. Speiches, P. T. Wingfield, Eds., *Current Protocols in Protein Science* (Wiley, New York, 1997), pp. 6.0.1–6.7.14, 9.4.1–9.4.16]. We prepared OMVs by Sarkosyl extraction as described [R. L. Davies, R. A. Wall, S. P. Borriello, *J. Immunol. Methods* **134**, 215 (1991)]. We then mixed 20 μg of each purified protein or 2996 OMV with Freund's adjuvant and used it to immunize CD1 mice at days 1, 21, and 35. Blood samples were taken on day 34 and 49.
14. Sera analysis: Antibody titers were determined by ELISA on microtiter plates coated with the capsulated strains inactivated by treatment with 0.025% para-

formaldehyde or with the noncapsulated M7 strain [D. S. Stephens, J. S. Swartley, S. Kathariou, S. A. Morse, *Infect. Immun.* **59**, 4097 (1991)]. Immunoglobulin titers were expressed as the reciprocal of serum dilution that gave an A_{490} value of 0.4 above the preimmune sera. We prepared immunoblots on purified proteins, OMVs, or total cell extracts with 15% polyacrylamide gels and 1:200 sera dilutions. We evaluated serum bactericidal activity against strains 2996 and BZ232, as described [C. C. A. C. Peeters *et al.*, *Vaccine* **17**, 2702 (1999)], with pooled baby rabbit serum used as complement source (Cedar Lane). The serum bactericidal titers were reported as the reciprocal of the serum dilution yielding ≥50% bacterial killing. We tested serum bactericidal activity with human complement, as described [R. E. Mandrell, F. H. Azmi, D. M. Granoff, *J. Infect. Dis.* **172**, 1279 (1995)], against strains 8047 (obtained from W. Zollinger), M986 [D. A. Caugant *et al.*, *J. Bacteriol.* **169**, 2781 (1987)], and NGP165 (*17*)—the strains in our collection that could be tested with our human complement source. The noncapsulated M7 strain or capsulated strains whose capsule had been permeabilized by treatment with 70% ethanol at −20°C for 1 hour were analyzed for cell-bound fluorescence with a FACScan flow cytometer. We used R-phycoerythrin–conjugated goat $F(ab)_2$ antibody to mouse immunoglobulin (Jackson ImmunoResearch) to detect antibody binding.

15. I. Goldschneider, E. C. Gotschlich, M. S. Artenstrain, *J. Exp. Med.* **129**, 1307 (1969).
16. M. C. J. Maiden *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3140 (1998).
17. A. Seiler, R. Reinhardt, J. Sarkari, D. A. Caugant, M. Achtman, *Mol. Microbiol.* **19**, 841 (1996).
18. J. Wang, D. A. Caugant, G. Morelli, B. Koumare, M. Achtman, *J. Infect. Dis.* **167**, 1320 (1993).
19. Gene variability: *Neisseria* strains. Twenty-two *N. meningitidis* serogroup B strains have been used for gene variability analyses: NG6/88, BZ198, NG3/88, 297-0, 1000, BZ147, BZ169, 528, NGP165, BZ133, NGE31, NGF26, NGE28, NGH38, SWZ107, NGH15, NGH36, BZ232, BZ83, H44/76 (*17*), MC58 [B. T. McGuinness *et al.*, *Lancet* **337**, 514 (1991) (*11*)], and 2996 (*13*); three serogroup A strains: 205900 and F6124 (our collection) and Z2491 (*16*) (sequenced at The Sanger Centre); two serogroup C strains: 90/18311 (*18*) and 93/4286 (*16*); one strain of each serogroup W, X, Z, and Y—A22, E26, 860800, and E32, respectively (*16*); one strain of *N. cinerea* (our collection); one strain of *N. lactamica* (our collection); and three strains of *N. gonorrhoeae*—Ng F62 and Ng SN4 (our collection) and FA1090 [J. A. Dempsey, W. Litaker, A. Madhure, T. L. Snodgrass, J. G. Cannon, *J. Bacteriol.* **173**, 5476 (1991); sequenced at The Oklahoma University]. We amplified selected genes by PCR with the following primers mapping about 70 base pairs upstream and downstream from the coding regions. *gna*33, TCGCGCCCTGTGTTAAAATCCCCT (Forward), GGTATCGCAAAACTTCGCCTTAATGCG (Reverse); *gna*992, GTTGGGGGAATTTATCAGAAAACCC (Forward), GGTCTCAGGCGGCAAATCGC (Reverse); *gna*1162, AGAGAAAAGGCTGTTTCCCG (Forward), CTTGCAAGTATCAAGATTCGC (Reverse); *gna*1220, TTGAACCAGGAACGCGCGCCC (Forward), TATTTGAAGCGGAATACAACCTTGCCC (Reverse), TATTTGAAGCGGAATACAACCCTGTTCG (Reverse for gonococcus); *gna*1946, CGAATCCGGACGGCAGGACTC (Forward), GGCAGGGAATGGCGGATTAAAG (Reverse); *gna*2001, CAATCAACAAGATATTTTCGACTG (Forward), TTTGACCTTTTCGGTACAGG (Reverse); *gna*2132, GGCGTTCAGACGGCATATTTTTAC (Forward), GGTTTATCAACTGATGCGGACTTGA (Reverse), TTGGGATGCCGCCTTTTTCGG (Reverse for gonococcus); *por*A, TATCGGGTGTTTGCCCGATGTTTTTAGG (Forward), TGGGGCTGAAGCAGATTGGCAGTCAG (Reverse). We used about 10 ng of chromosomal DNA as template. PCR products were purified and sequenced by primer walking on both strands. To study the frequency of recombination among closely related nucleotide sequences, we used the homoplasy test (ftp://novell-del-valle.molgen.mpg.de), which was designed to estimate the significance of convergent mutations in a phylogenetic tree by comparing the expected and observed value of homoplasies, being 0 for a clonal population and 1 for complete linkage equilibrium. We applied the test to the seven selected genes (or selected conserved regions), using in each case only different allele sequences. The data reported were obtained as described (*21*), using no outgroup and considering a medium level of expression for all the genes. The values of the H ratio obtained were, respectively, 0.005 for *gna*1946, 0.065 for *gna*2001, 0.133 for *gna*1162, 0.251 for *gna*1220, 0.302 for the 3′-end of *gna*992, and 0.524 for *gna*33, which give a geometric mean of 0.11. This is lower than the value previously reported for conserved meningococcal housekeeping genes [S. Suerbaum *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12619 (1998)]. *gna*2132, which, even in the conserved 3′-end, gave an H ratio of 0.707, indicating high levels of recombination, was not included in the analysis.

20. A. Bart, J. Dankert, A. van der Ende, *Infect. Immun.* **67**, 3842 (1999).
21. J. Maynard Smith and N. H. Smith, *Mol. Biol. Evol.* **15**, 590 (1998).
22. J. Lommatzsch, M. F. Templin, A. R. Kraft, W. Vollmer, J.-V. Holtje, *J. Bacteriol.* **179**, 5465 (1997); J. V. Holtje, *Microbiol. Mol. Biol. Rev.* **62**, 181 (1998).
23. M. Pizza *et al.*, data not shown.
24. N. Saitou and M. Nei, *Mol. Biol. Evol.* **4**, 406, 1987.
25. M. Kimura, *J. Mol. Evol.* **16**, 111, 1980.
26. J. St. Geme III, D. Cutter, S. J. Barenkamp, *J. Bacteriol.* **178**, 6281 (1996); D. A. Relman, M. Domenighini, E. Toumanen, R. Rappuoli, S. Falkow, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2637 (1989).
27. M. Chanyangam, A. L. Smith, S. L. Moseley, M. Kuehn, P. Jenny, *Infect. Immun.* **59**, 600 (1991).
28. M. D. Wuenscher, S. Kohler, A. Bubert, U. Gerike, W. Goebel, *J. Bacteriol.* **175**, 3491 (1993).
29. Author contributions were as follows: experimental, M. Pizza; computer, V. Scarlato; coordination, G. Grandi; computer analysis, V. Masignani, M. Scarselli, M. Mora, C. L. Galeotti, and G. Ratti; gene expression, B. Aricò, M. Comanducci, B. Capecchi, L. Baldi, E. Storni, M. Broeker, B. Knapp, and E. Hundt; gene variability, M. Comanducci and B. Capecchi; sera analysis, M. M. Giuliani, L. Santini, E. Luzzi, and E. Bartolini; help with FACS analysis, S. Nuti; protein purification, G. T. Jennings, S. Savino, E. Marchetti, P. Zuo, and R. Manetti; nucleotide sequencing, E. Blair, T. Mason, and H. Tettelin; help with project planning and sequence annotation, E. R. Moxon, N. J. Saunders, D. Hood, and A. C. Jeffries; help with project planning and bactericidal assays, D. M. Granoff. We thank M. Achtman and F. Frati for useful discussions; B. Brunelli, D. Serruto, and D. Veggi for technical help; and J. Adu-Bobie for sharing unpublished observations. We also acknowledge the Gonococcal Genome Sequencing Project (B. A. Roe, S. P. Lin, L. Song, X. Yuan, S. Clifton, T. Ducey, L. Lewis, and D. W. Dyer) at The University of Oklahoma and the *N. meningitidis* Sequencing Group at the Sanger Centre. Finally, we thank C. Mallia for editing, G. Corsi for artwork, and D. Kingsbury for support and advice.

25 October 1999; accepted 27 January 2000

gridlock, an HLH Gene Required for Assembly of the Aorta in Zebrafish

Tao P. Zhong,* Michael Rosenberg,*† Manzoor-Ali P. K. Mohideen,‡ Brant Weinstein,§ Mark C. Fishman‖

The first artery and vein of the vertebrate embryo assemble in the trunk by migration and coalescence of angioblasts to form endothelial tubes. The *gridlock* (*grl*) mutation in zebrafish selectively perturbs assembly of the artery (the aorta). Here it is shown that *grl* encodes a basic helix-loop-helix (bHLH) protein belonging to the Hairy/Enhancer of the split family of bHLH proteins. The *grl* gene is expressed in lateral plate mesoderm before vessel formation, and thereafter in the aorta and not in the vein. These results suggest that the arterial endothelial identity is established even before the onset of blood flow and implicate the *grl* gene in assignment of vessel-specific cell fate.

Arteries and veins are morphologically and functionally very distinct. For example, arteries deliver oxygenated blood at high pressure from the heart, whereas veins serve as capacitance vessels for blood return. Some of the morphological differences may be imposed after, and depend upon, onset of function. However, a complete vascular loop, composed of the trunk dorsal aorta and posterior cardinal vein, is needed to accommodate the output of the first heartbeat. These simple tubes of endothelium form by local aggregation of angioblasts, a process termed vasculogenesis (*1*). Neither mutations nor molecular markers have revealed whether there are arterial-venous distinctions between angioblast progenitors. In the mouse, ephrinB2 is selectively expressed on the arteries and EphB3 and EphB4 on the veins, but this occurs after vasculogenesis (*2*). Furthermore, mutation of ephrinB2 does not affect vasculogenesis, although it does disrupt later vessel formation and remodeling, a process termed angiogenesis (*2*).

The gridlock mutation (grl^{m145}) was originally isolated in a large-scale chemical mu-

Cardiovascular Research Center, Massachusetts General Hospital–Harvard Medical School, 149 13th Street, 4th floor, Charlestown, MA 02129, USA.

*These authors contributed equally to this work.
†Present address: Aventis Pharma, 26 Landsdowne Street, Cambridge, MA 02139, USA.
‡Present address: Jake Gittlen Cancer Research Institute, Pennsylvania State University College of Medicine, 500 University Drive, Hershey, PA 17033, USA.
§Present address: Unit on Vertebrate Organogenesis, Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, Building 6B, 6 Center Drive, Bethesda, MD 20892, USA.
‖To whom correspondence should be addressed. E-mail: fishman@cvrc.mgh.harvard.edu