



DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

The Complete Genome Sequence of *Proteus mirabilis* Strain BB2000 Reveals Differences from the *P. mirabilis* Reference Strain

The Harvard community has made this article openly available.
[Please share](#) how this access benefits you. Your story matters.

Citation	Sullivan, N. L., A. N. Septer, A. T. Fields, L. M. Wenren, and K. A. Gibbs. 2013. "The Complete Genome Sequence of <i>Proteus Mirabilis</i> Strain BB2000 Reveals Differences from the <i>P. Mirabilis</i> Reference Strain." <i>Genome Announcements</i> 1 (5) (September 5): e00024–13–e00024–13. doi:10.1128/genomea.00024-13. http://dx.doi.org/10.1128/genomeA.00024-13 .
Published Version	doi:10.1128/genomeA.00024-13
Accessed	July 23, 2018 12:15:36 AM EDT
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:11857755
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

(Article begins on next page)

1 The complete genome sequence of strain BB2000 reveals differences from the *Proteus*
2 *mirabilis* reference strain

3

4

5

6 Nora L. Sullivan¹, Alecia N. Septer, Andrew T. Fields, Larissa M. Wenren, and Karine
7 A. Gibbs*

8

9 Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue,
10 Cambridge, MA 02138

11 ¹current address: W.M. Keck Science Department, 925 N. Mills Ave, Claremont
12 McKenna, Pitzer and Scripps Colleges, Claremont, CA 91711

13

14 *Corresponding author:

15 Karine A. Gibbs, Ph.D.

16 Department of Molecular and Cellular Biology

17 Harvard University

18 16 Divinity Avenue

19 Cambridge, MA 02138

20 Em: kagibbs@mcb.harvard.edu

21

22

23 Running title: Complete genome sequence of *P. mirabilis* BB2000

24

24 Abstract

25 We announce the complete genome for *Proteus mirabilis* strain BB2000, a model
26 system for self recognition. This opportunistic pathogen contains a single, circular
27 chromosome (3,846,754 base pairs). Comparisons between this genome and that of strain
28 HI4320 reveal genetic variations corresponding to previously unknown physiological and
29 self-recognition differences.

30

30 The gut commensal bacterium *Proteus mirabilis* is the primary cause of urinary
31 tract infections in patients with long-term indwelling catheters (1-4). Interestingly,
32 migrating colonies of *P. mirabilis* cells can distinguish self from non-self: a visible
33 boundary forms at the interface between two genetically distinct colonies, while two
34 genetically identical populations merge together (5). The genetic determinants of this
35 self-recognition behavior, first identified in *P. mirabilis* strain BB2000, included self-
36 identity genes containing numerous inter-strain nucleotide polymorphisms and suggested
37 that additional genetic differences between strains are likely (6). To date, only the
38 genome of *P. mirabilis* strain HI4320 (NCBI NC_010554) has been completed (7). Here
39 we report a second closed genome, that of the genetically distinct strain, BB2000 (8).

40 BB2000 genomic DNA was isolated and sequenced using standard protocols.
41 Briefly, DNA was isolated from cells cultured in modified LB broth using
42 phenol/chloroform extraction and ethanol (9). Beckman Coulter Genomics (Danvers,
43 MA) performed initial library preparation and sequencing using the Roche 454 platform.
44 Illumina sequencing was used to confirm the 454 data and resolve stretches of unknown
45 nucleotides; genomic DNA libraries were prepared according to the Illumina
46 Multiplexing Sample Preparation protocol and sequenced by Harvard FAS Systems
47 Biology Core using an Illumina HiSeq 2000. Illumina reads were assembled onto the 454
48 genomic data using Galaxy software (10). Genome closure was accomplished by
49 amplifying across gaps using polymerase chain reactions followed by Sanger sequencing
50 performed by Genewiz Corporation (South Plainfield, NJ).

51 The *P. mirabilis* BB2000 genome consists of a single chromosome (3,846,754
52 base pairs) with 38.6% G+C content. Potential coding sequences (CDSs) were identified

53 using the xBase annotation service, which predicted CDS regions using Glimmer (11),
54 and assigned predicted protein products based on a direct comparison to the *P. mirabilis*
55 HI4320 genome (12-16). CDSs absent in the HI4320 genome were assigned
56 “hypothetical protein” as the predicted product. Twenty-eight genes related to self-
57 recognition (6, 17) were annotated manually using blastx (12) and the HMMER web
58 interface (18). Sequence assembly and annotation were completed using Artemis
59 software (19). The BB2000 genome encodes 3,457 potential CDSs, of which 2,592 are
60 assigned a putative function; the remaining 865 CDSs are classified as hypothetical
61 proteins, with an additional 81 tRNA genes and 22 rRNA genes.

62 Comparison of the BB2000 genome to that of strain HI4320 (7) revealed 93%
63 similarity between the chromosomes. The CDSs unique to each genome include genes
64 related to phage, toxin elements, and self recognition. The HI4320 genome encodes iron
65 acquisition proteins that are absent in BB2000. Strain HI4320 also contains a plasmid
66 (NCBI NC_010555.1) (7), and the HI4320 chromosome encodes a complete set of *tra*
67 genes for conjugative transfer. No plasmid was identified in BB2000, nor does its
68 genome encode *tra* genes or any HI4320 plasmid-encoded genes. Further analysis of
69 variations between *P. mirabilis* isolates will advance our understanding of the genetic
70 determinants of pathogenicity and self recognition.

71 **Nucleotide sequence accession number.** The *P. mirabilis* BB2000 genome
72 sequence has been deposited in GenBank under the accession number BankIt1590180
73 BB2000 CP004022.

74

75 ACKNOWLEDGMENTS

76 Harvard University provided funding for this research.

77

78 The authors thank Beckman Coulter Genomics, the Harvard FAS Systems Biology Core,
79 and the Harvard Research Computing Group for insightful advice during the genome
80 construction.

81

82 REFERENCES

- 83 1. **Mobley HLT, Belas R.** 1995. Swarming and pathogenicity of *Proteus mirabilis* in
84 the urinary tract. *Trends Microbiol.* **3**:280-284.
- 85 2. **Mathur S, Sabbuba NA, Suller MTE, Stickler DJ, Feneley RCL.** 2005.
86 Genotyping of urinary and fecal *Proteus mirabilis* isolates from individuals with
87 long-term urinary catheters. *Eur. J. Clin. Microbiol.* **24**:643-644.
- 88 3. **Nicolle LE.** 2005. Catheter-related urinary tract infection. *Drug Aging* **22**:627-639.
- 89 4. **Armbruster CE, Mobley HL.** 2012. Merging mythology and morphology: the
90 multifaceted lifestyle of *Proteus mirabilis*. *Nat. Rev. Microbiol.* **10**:743-754.
- 91 5. **Dienes L.** 1946. Reproductive processes in *Proteus* cultures. *J. Bacteriol.* **51**:24-24.
- 92 6. **Gibbs KA, Urbanowski ML, Greenberg EP.** 2008. Genetic determinants of self
93 identity and social recognition in bacteria. *Science* **321**:256-259.
- 94 7. **Pearson MM, Sebahia M, Churcher C, Quail MA, Seshasayee AS, Luscombe**
95 **NM, Abdellah Z, Arrosmith C, Atkin B, Chillingworth T, Hauser H, Jagels K,**
96 **Moule S, Mungall K, Norbertczak H, Rabbinowitsch E, Walker D, Whithead S,**
97 **Thomson NR, Rather PN, Parkhill J, Mobley HLT.** 2008. Complete genome

- 98 sequence of uropathogenic *Proteus mirabilis*, a master of both adherence and
99 motility. J. Bacteriol. **190**:4027-4037.
- 100 8. **Belas R, Erskine D, Flaherty D.** 1991. Transposon mutagenesis in *Proteus*
101 *mirabilis*. J. Bacteriol. **173**:6289-6293.
- 102 9. **Sambrook J, Russell DW.** 2001. Molecular cloning a laboratory manual, 3rd ed.
103 Cold Spring Harbor Laboratory Press, New York.
- 104 10. **Goecks J, Nekrutenko A, Taylor J.** 2010. Galaxy: a comprehensive approach for
105 supporting accessible, reproducible, and transparent computational research in the
106 life sciences. Genome Biol. **11**:R86.
- 107 11. **Delcher AL, Bratke KA, Powers EC, Salzberg SL.** 2007. Identifying bacterial
108 genes and endosymbiont DNA with Glimmer. Bioinformatics **23**:673-679.
- 109 12. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman**
110 **DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database
111 search programs. Nucleic Acids Res. **25**:3389-3402.
- 112 13. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: A program for improved detection of
113 transfer RNA genes in genomic sequence. Nucleic Acids Res. **25**:955-964.
- 114 14. **Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C,**
115 **Salzberg SL.** 2004. Versatile and open software for comparing large genomes.
116 Genome Biol. **5**:R12.
- 117 15. **Chaudhuri RR, Loman NJ, Snyder LAS, Bailey CM, Stekel DJ, Pallen MJ.**
118 2008. xBASE2: a comprehensive resource for comparative bacterial genomics.
119 Nucleic Acids Res. **36**:D543-D546.

- 120 16. **Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007.
121 RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids*
122 *Res.* **35**:3100-3108.
- 123 17. **Wenren LM, Sullivan NL, Cardarelli L, Septer AN, Gibbs KA.** 2013. Two
124 independent pathways for self-recognition in *Proteus mirabilis* are linked by type VI-
125 dependent export. *mBio* **4**:e00374-00313.
- 126 18. **Finn RD, Clements J, Eddy SR.** 2011. HMMER web server: interactive sequence
127 similarity searching. *Nucleic Acids Res.* **39**:W29-37.
- 128 19. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA,**
129 **Barrell B.** 2000. Artemis: sequence visualization and annotation. *Bioinformatics*
130 **16**:944-945.
- 131