Network analysis reveals sex and antibiotic resistance-associated antivirulence targets in clinical uropathogens

Kaveri S. Parker[1,2,3], James D. Wilson[4], Jonas Marschall[2,3,6], Peter J. Mucha[5]*, Jeffrey P. Henderson[1,2,3]*

[1]Center for Women's Infectious Diseases Research, [2]Division of Infectious Diseases, [3]Department of Internal Medicine, Washington University in St. Louis, St. Louis, Missouri, United States of America, [4]Department of Mathematics and Statistics, the University of California, San Francisco, United States of America, [5]Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, North Carolina, United States of America, [6]Department of Infectious Diseases, Bern University Hospital and University of Bern, Bern, Switzerland
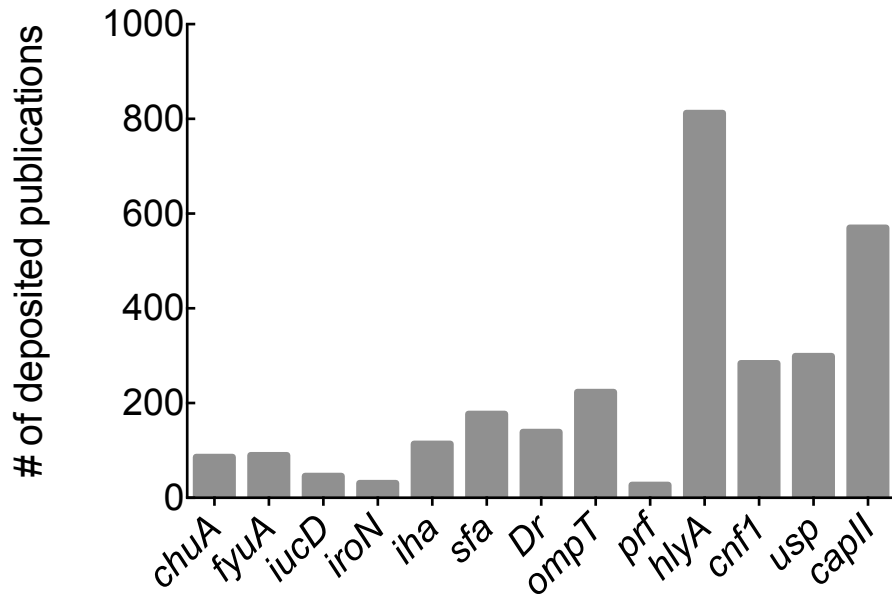
*Co-corresponding authors

In this file:

- **Supplementary figures**

*Jeffrey P. Henderson, M.D. Ph.D.	***Peter J. Mucha, Ph.D.**
Center for Women's Infectious Disease Research	Department of Mathematics
Box 8051 Washington University School of Medicine	120 E. Cameron Avenue
	CB#3250, UNC
660 S. Euclid Ave., St. Louis, MO 63110	Chapel Hill, NC 27599
E-mail: jhenderson@DOM.wustl.edu	E-mail: mucha@unc.edu

**Table S1. Distributions of stereotypical virulence groupings within clinical isolate phylogenetic groups.**

|	A (n = 10)	B1 (n = 41)	B2 (n = 232)	D (n = 54)
CI community				
C1	0	0	38	7
C2	0	2	109	7
C3	0	0	75	1
C4	10	39	10	39
bicluster				
BC1	0	0	76	8
BC2	0	1	0	0
BC3	0	0	0	1
BC4	0	2	3	24
BC1+2	0	0	89	0
BC1+3	0	0	49	0
BC2+4	0	2	0	8
BC1+2+3	0	0	12	0
none	10	36	3	13

Figure S1. Total publications addressing virulence factors used in this study exceed 2,700. The figure shows the distribution of papers recorded in PubMed using each virulence factor as a query [(virulence factor*) AND (UPEC* OR E.coli*)]. The search was conducted on 3 February 2014.
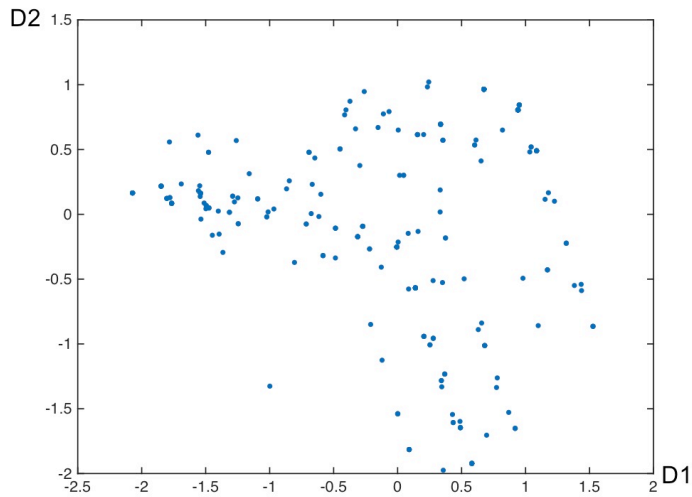
Figure S2. Principal component analysis (PCA) plot does not identify organized clusters. Principal Component Analysis (PCA) projects the original 16-dimensional point cloud data representing VF presence (1) and absence (0) in a specified CI into a lower-dimensional Euclidean subspace. While some of the outlying points of this projection hint at possible organization, there is no strong clustering indicated in this figure.
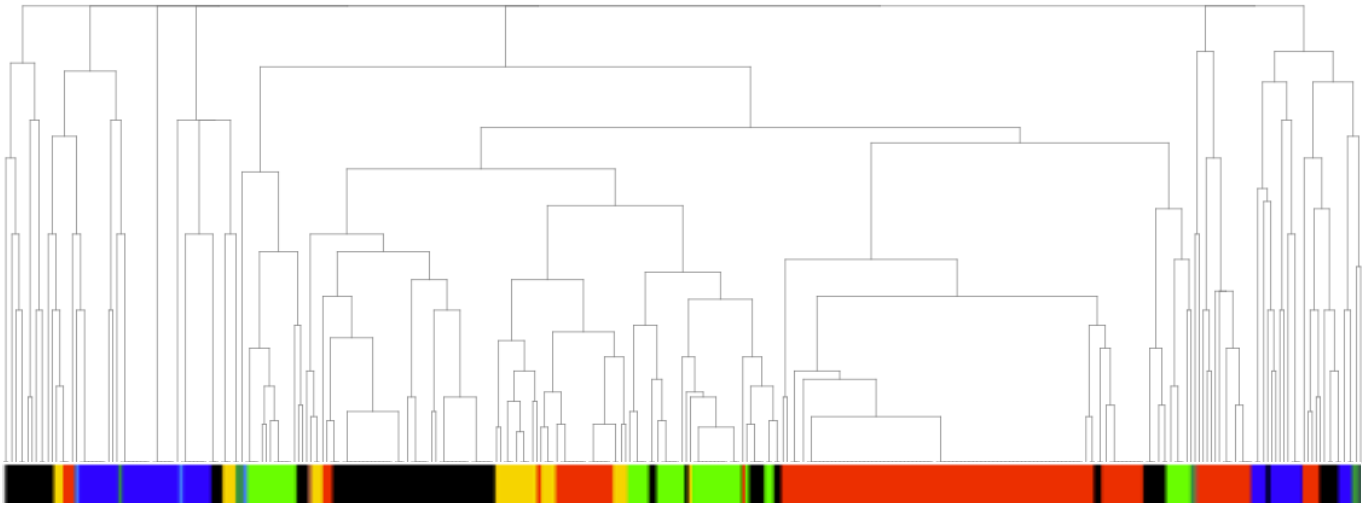
Figure S3. Naïve hierarchical analysis can yield complex solutions. Complete hierarchical clustering of the data yields seven clinical isolate groups. We chose the number of clusters, k, according to the clustering that had the largest average between cluster sum of squares (BCSS). To do this, we calculate the BCSS of the clusters identified by Hierarchical Clustering for k from 2 to 336.
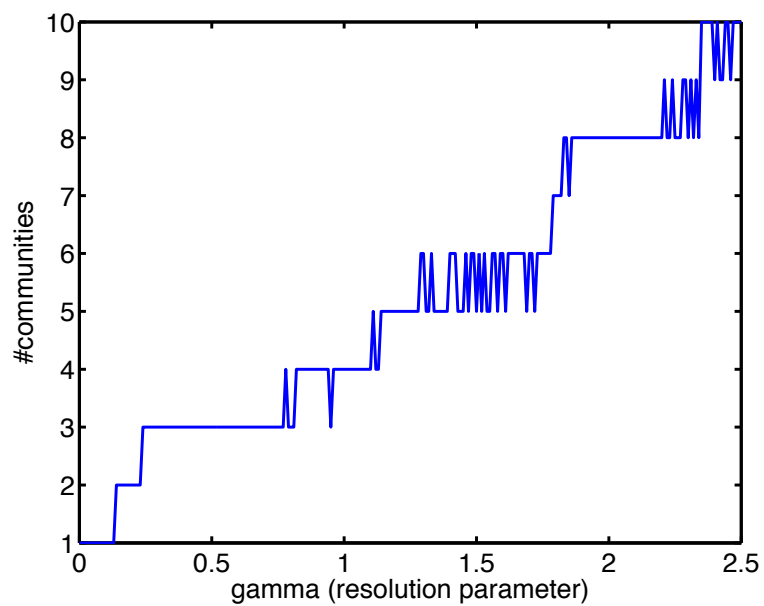
Figure S4. Gamma (γ) resolution parameter for community detection. Multiple plateaus of different numbers of CI communities are indicated. The plateau of four communities is indicated robustly for **γ** near its default value **(γ=1).**