Supplementary Figures and Tables: CSI-Tree: A regression tree approach for modeling binding properties of DNA binding molecules based on cognate site identification (CSI) data

Sündüz Keleş*

Department of Statistics

Department of Biostatistics and Medical Informatics

University of Wisconsin, Madison

Christopher L. Warren

Department of Biochemistry

University of Wisconsin, Madison

Clayton D. Carlson

Department of Biochemistry

University of Wisconsin, Madison

Aseem Z. Ansari

Department of Biochemistry

Genome Center

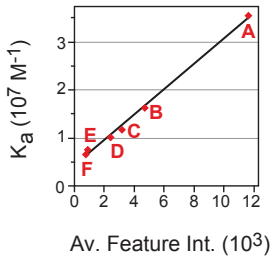
University of Wisconsin, Madison.

January 6, 2008

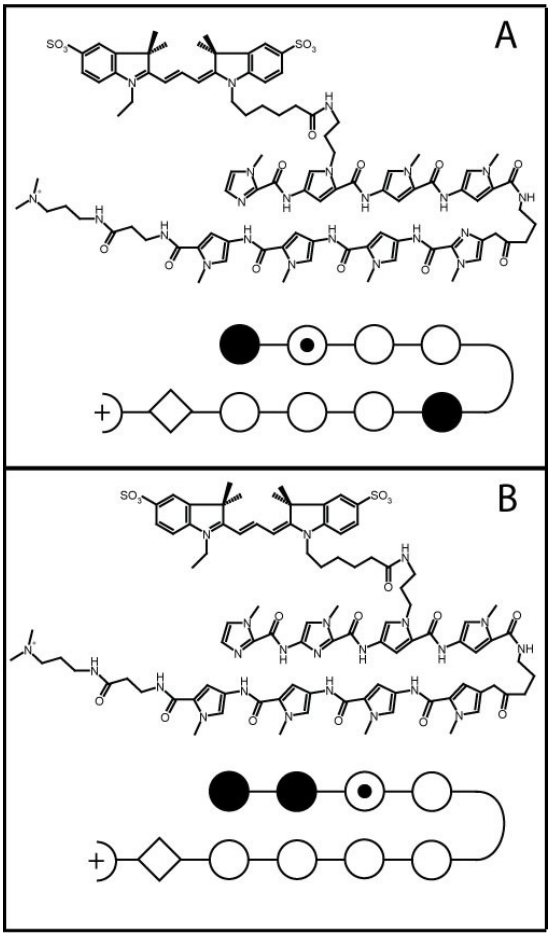
*Corresponding Author: keles@stat.wisc.edu

Data Point	Sequence	Ka (10^7 M^{-1})	Intensity (10^3)
A	AAGGTTA	3.59	11.6
B	AACGTTA	1.62	4.71
C	ATGGAAT	1.17	3.16
D	CAGGTTA	0.95	2.39
E	TACGATT	0.73	0.86
F	AAGGTCA	0.65	0.79

Supplementary Table 1. CSI fluorescent intensities and corresponding binding affinities Kas determined from nuclease protection assays for the six sequences spanning the range of the PA2 CSI data.



Supplementary Figure 1. CSI fluorescent intensities versus binding affinities Ka for six sequences spanning the range of the PA2 CSI data. Similar linear correlations are observed for PA1 [1] and Nkx-2.5 (data not shown).



Supplementary Figure 2. Chemical structures of polyamides 1 (PA1) and 2 (PA2).

References

[1] C.L. Warren, N. C. Kratochvil, K. E. Hauschild, S. Foister, M. L. Brezinski, P. B. Dervan, G. N. Phillips, and A.Z. Ansari. Defining the sequence-recognition profile of dna-binding molecules. *Proceedings of the National Academy of Sciences U.S.A*, 103(4):867–872, January 2006.