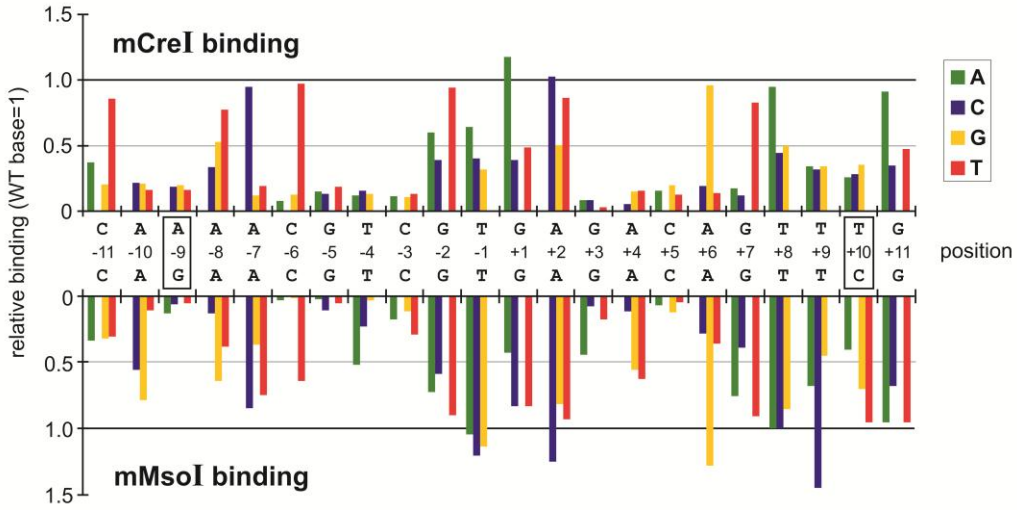
Supplementary Information for *Nucleic Acids Research* ms:

‘Comprehensive target site specificity profiling reveals homing endonuclease evolutionary constraints and enables genome engineering applications’

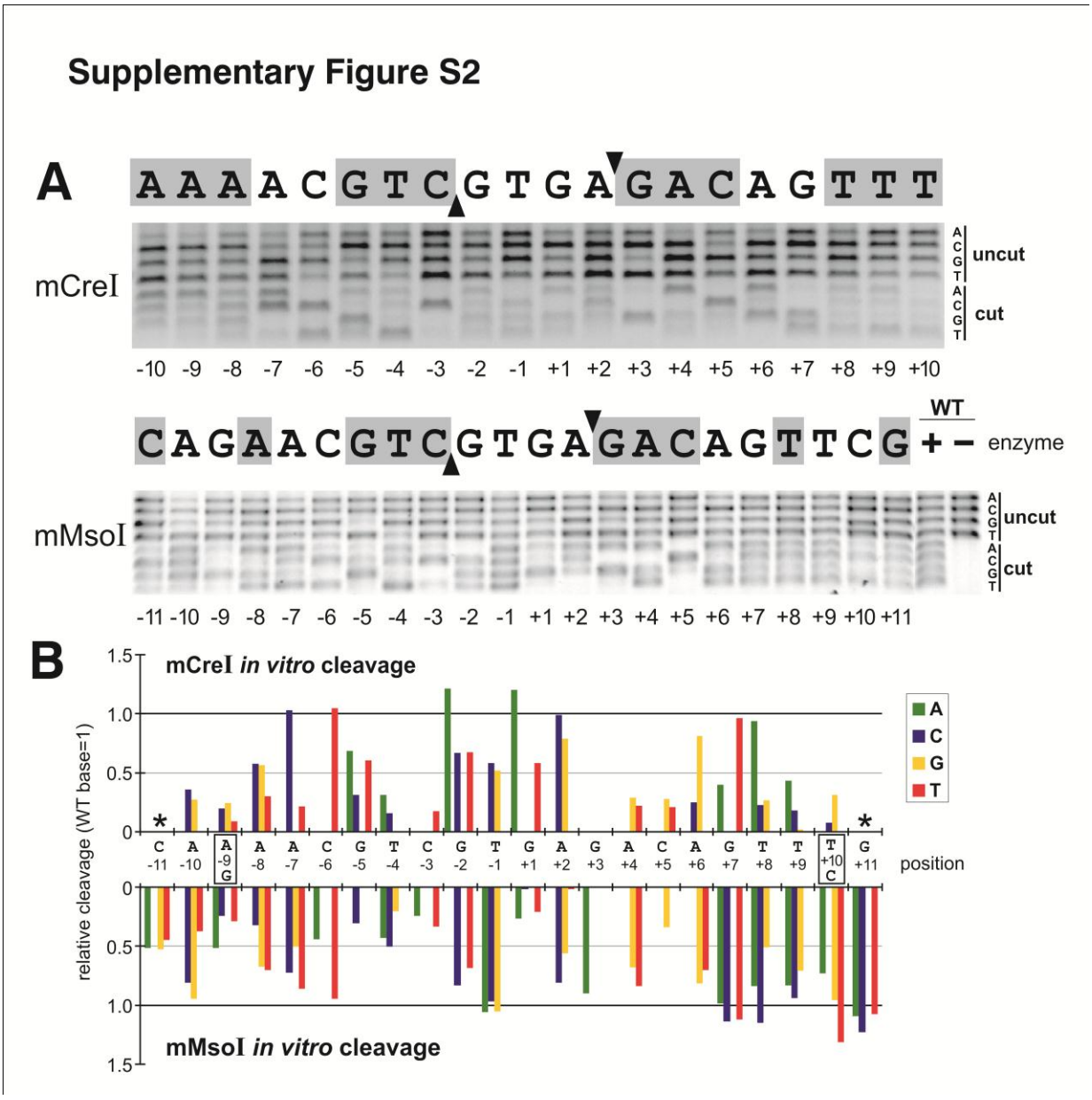
Authors: Hui Li, Umut Ulge, Blake T. Hovde, Lindsey Doyle and Raymond J. Monnat, Jr.

Note: This supplement contains 4 Supplementary Figures and Legends and 1 Supplementary Table

Supplementary Figure S1

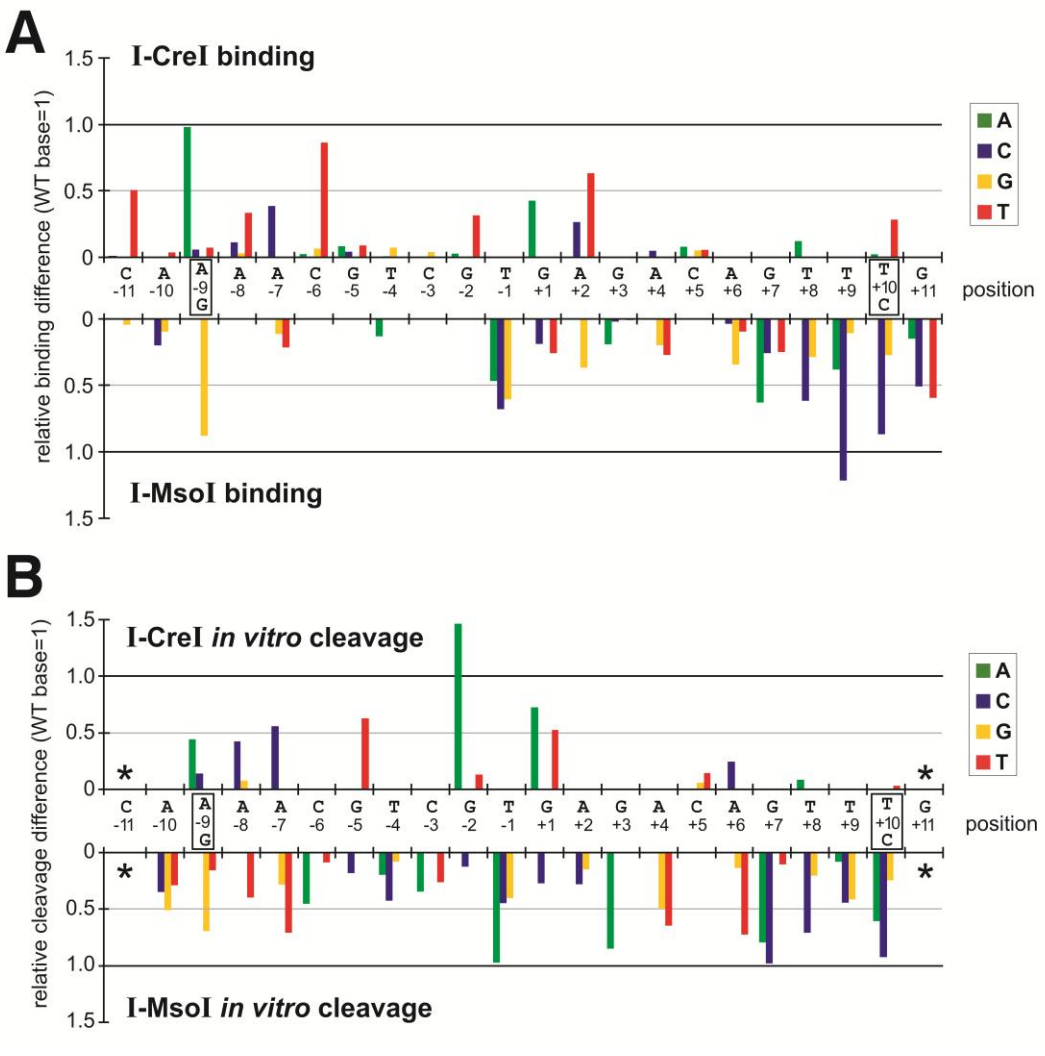


Supplementary Figure 1. Comparative binding specificity profiles for mCreI and mMsoI. The relative binding affinities of mCreI (top panel) and mMsoI (bottom panel) for all four base pairs at each target site position were determined by competitive binding (19). Target site positions that differ between the I-CreI and I-MsoI target sites are shown in boxes (-9 and +10). Bar heights indicate binding affinity for alternative base pairs relative to the native base pair, whose binding affinity has been arbitrarily set to 1.0. All results are the mean of three replications in which the standard deviation between experiments was ± 5%. Error bars have been omitted for graphical clarity.



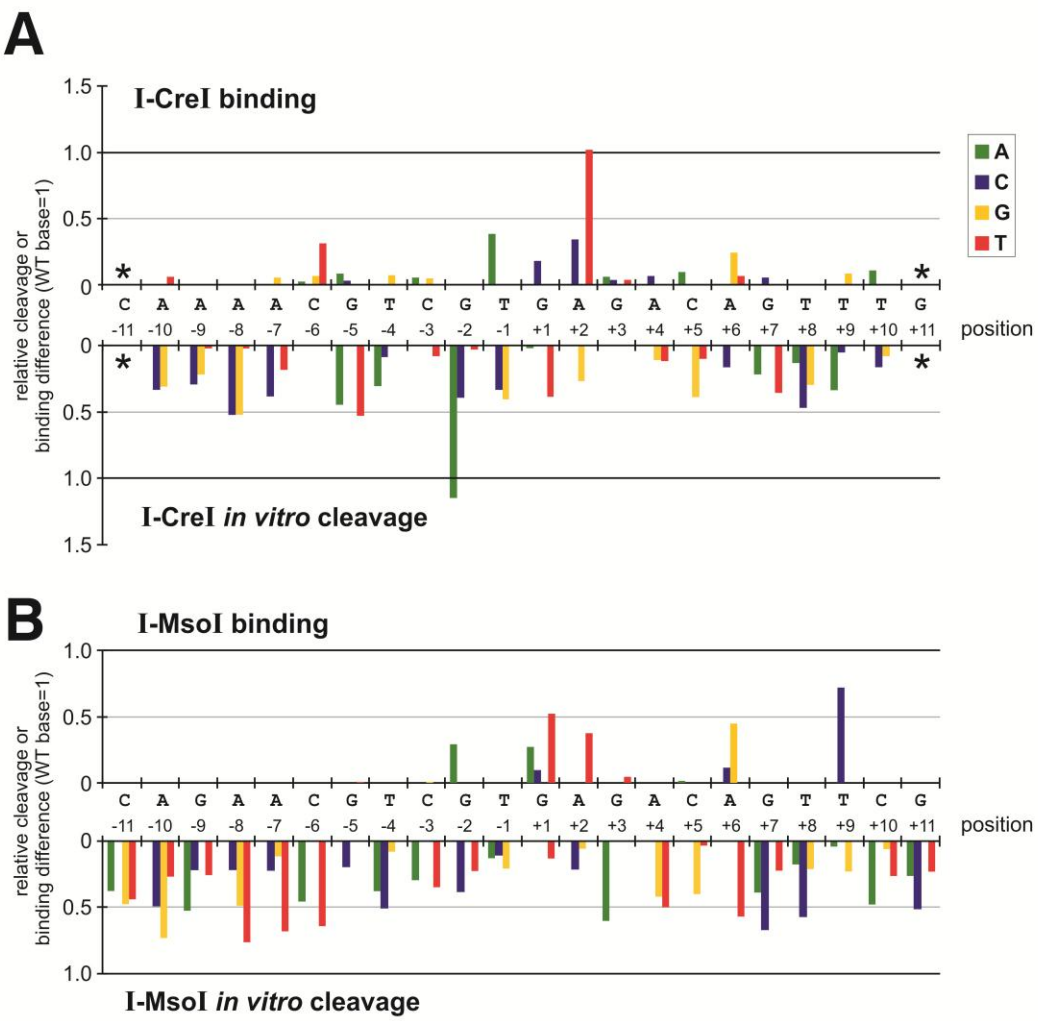
Supplementary Figure 2. Comparative cleavage specificity profiles for mCreI and mMsoI. (A) Agarose gels of *in vitro* cleavage specificity profiles for mCreI and mMsoI, determined as described in Figure 3 by a competitive 'bar code' cleavage assay. (B) The relative cleavage efficiencies of mCreI (top panel) and mMsoI (bottom panel) on target sites containing base substitutions relative to the cleavage efficiency of the native base pair plotted as in Figure 2. Asterisks indicate positions where we could not determine a difference measure due to an absence of mCreI data. All results are the mean of three replicates in which the standard deviations were ± 5%. Error bars have been omitted for graphical clarity.

Supplementary Figure S3



Supplementary Figure 3. Difference plot of I-CreI and I-MsoI *in vitro* binding and cleavage profiles. (A) Difference plot of binding of I-CreI (top panel) *versus* I-MsoI (bottom panel), where bar direction indicates the better binder (up = I-CreI>I-MsoI; down = I-MsoI>I-CreI) and the bar height indicates the difference in relative binding . Base pair positions that differ between the I-CreI and I-MsoI target sites are shown in boxes. (B) Difference plot of cleavage efficiencies of I-CreI (top panel) *versus* I-MsoI (bottom panel), where bar direction indicates the better cleaver (up = I-CreI>I-MsoI; down = I-MsoI>I-CreI) and the bar height indicates the difference in relative cleavage. Asterisks indicate positions where we lack cleavage data for I-CreI, and thus could not determine a difference measure.

Supplementary Figure S4



Supplementary Figure 4. Difference plot of I-CreI and I-MsoI *in vitro* binding *versus* cleavage specificity profiles. Difference plots of binding *versus* cleavage profiles for I-CreI (A) and I-MsoI (B) where bar direction indicates better binding (up) or cleavage (down) for each enzyme across all target site positions and bar heights the magnitude of the binding or cleavage differential. These plots readily reveal base pair substitutions that have a disproportionate effect on binding or cleavage. Asterisks indicate positions where we lack data for I-CreI cleavage, and could not determine a difference measure.

Supplementary Table 1. Blastn results of I-CreI/I-MsoI target sites with two symmetric base pair changes.

Two positions in I-CreI/I-MsoI target sites were changed symmetrically to maintain the ribosomal RNA secondary structure depicted in Figure 6A, and cleavage sensitivity of the resultant target site variants was predicted based on combined relative cleavage activities from Figure 3. For I-CreI/mCreI, the target site variant with -7C and +8G changes is predicted to be the most cleavage sensitive. For I-MsoI, the target site variants with -6T/+7A, -7T/+8A changes are predicted to be the most cleavage sensitive. For both HEs, the target site variant with -3G and +3C changes is predicted to be the most cleavage resistant. The perfect matches of these target site variants were searched using Blastn. Only 75 perfect matches are found for I-CreI target site variant with -7C and +8G changes, 55 of which are identified as ribosome large subunit (LSU) genes. No perfect match is found for the remaining high-scoring target site variants.

	Target site sequence	Changes	Sum of relative cleavage activity (WT=2)	Blastn results
I-CreI	AAAcCGTCGTGAGACAGgTT	-7C, +8G	1.68	75(55)
	AAAACGTgGTGAcACAGTTT	-3G, +3C	0	0
I-MsoI	CAGAAtGTCGTGAGACAaTTCG	-6T, +7A	2.04	0
	CAGAtCGTCGTGAGACAGaTCG	-7T, +8A	2.03	0
	CAGAACGTgGTGAcACAGTTCG	-3G, +3C	0	0