Rank	Clustering	Function	Orthology	Alternative orthology
1	236 (10.4)	245 (6.35)	236 (6.79)	245 (3.49)
2	238 (9.54)	238 (5.04)	232 (6.50)	236 (3.44)
3	239 (8.40)	232 (3.94)	234 (6.42)	246 (3.10)
4	246 (7.59)	236 (3.64)	238 (6.32)	256 (2.88)
5	234 (7.09)	246 (3.23)	242 (5.37)	234 (2.80)
6	232 (6.68)	234 (2.94)	228 (4.43)	242 (2.73)
7	235 (6.28)	239 (2.92)	241 (4.07)	240 (2.71)
8	242 (6.24)	247 (2.90)	233 (3.56)	232 (2.67)
9	253 (4.87)		247 (3.01)	
10	245 (4.15)			
11	228 (3.89)			
12	247 (3.35)			
13	249 (3.09)			
14	231 (3.09)			

Supplementary Table 1. Basic leucine zipper results with althernative orthology results. Our method is labeled “Clustering”, the functional grouping method is labeled “Function,” the KOG-based orthology method is labeled “Orthology,” and the alternative orthology method based on protein names is labeled “Alternative orthology.” Residues that are experimentally known to play a role in DNA half site specificity are underlined. Z-scores for the different methods are represented in parenthesis. The first eight or the total number of residues with z-score above 3.0 are shown for each method.

Rank	Clustering	Function	Orthology	Alternative orthology	ET
1	462 (15.9)	513 (4.89)	465 (4.89)	456 (8.47)	458
2	459 (13.4)	459 (4.33)	506 (3.16)	465 (8.14)	459
3	493 (13.3)	467 (4.23)	468 (3.06)	467 (7.85)	462?
4	490 (11.7)	493 (3.75)	449 (3.02)	497 (7.17)	465
5	465 (10.3)	465 (3.11)	442 (2.88)	513 (5.73)	490
6	452 (9.86)	450 (2.96)	456 (2.84)	491 (5.42)	493
7	458 (8.68)	458 (2.73)	459 (2.60)	462 (5.32)	511
8	467 (8.31)	462 (2.61)	472 (2.54)	490 (5.27)	513
9	491 (8.16)	507 (1.75)	441 (2.42)	493 (5.00)	
10	469 (7.97)	469 (1.65)	450 (2.11)	509 (4.88)	
11	446 (7.84)			511 (4.86)	
12	456 (7.26)			468 (4.84)	
13	450 (5.78)			442 (4.70)	
14	477 (5.30)			480 (4.70)	
15	470 (5.21)			501 (4.70)	
16	448 (4.94)			508 (4.66)	
17	468 (4.93)			450 (4.55)	
18	455 (4.78)			477 (4.51)	
19	510 (4.44)			498 (4.37)	
20	473 (4.20)			454 (4.29)	
21	445 (4.00)			507 (4.27)	

22	513 (3.92)			470 (4.24)	
23	506 (3.91)			488 (4.20)	
24	488 (3.73)			459 (4.02)	
25	447 (3.59)			441 (3.89)	
26	449 (3.23)			506 (3.85)	
27	511 (3.15)			484 (3.85)	
28	508 (3.08)			478 (3.77)	
29				455 (3.61)	
30				447 (3.51)	
31				469 (3.42)	
32				487 (3.38)	

Supplementary Table 2. Nuclear receptor results. . Our method is labeled “Clustering”, the functional grouping method is labeled “Function,” the KOG-based orthology method is labeled “Orthology,” and the alternative orthology method based on protein names is labeled “Alternative orthology.” “ET” stands for evolutionary trace; these are the predictions presented unranked by Lichtarge, et. al.(3) and are listed sequentially here. The question mark for evolutionary trace predicted positions 462 is included because this position was not originally found by their algorithm but argued for in their paper because of improper pruning. Residues that are experimentally known to play a role in DNA half site specificity are underlined. The first ten or the total number of residues with z-score above 3.0 are shown for each of the four groupings used in this work.