Supporting Information

Nakagawa et al. 10.1073/pnas.1310430110

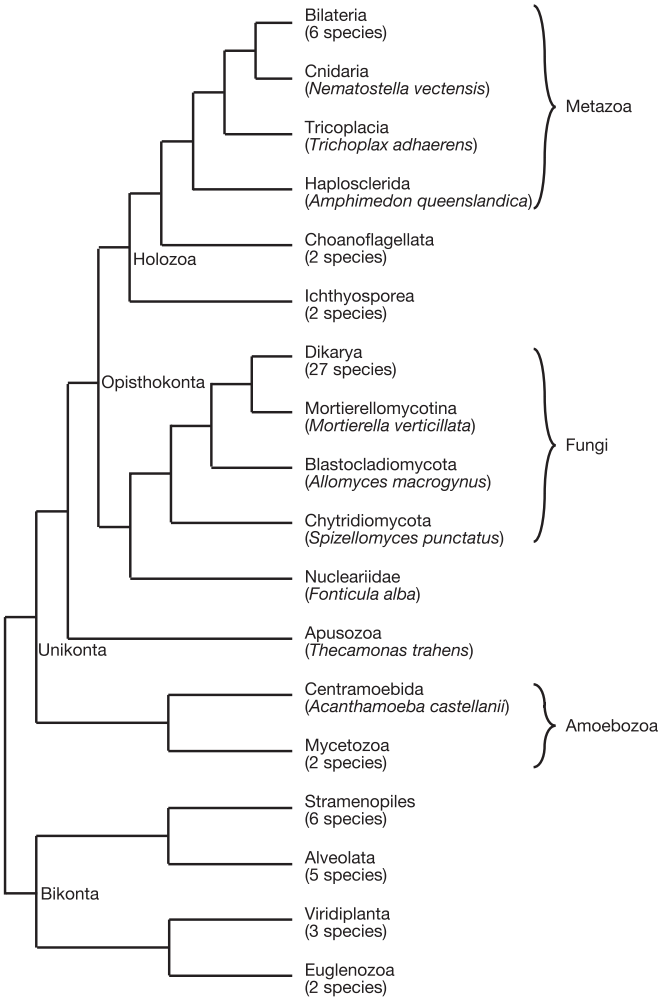


Fig. S1. Phylogeny of species used in this study.

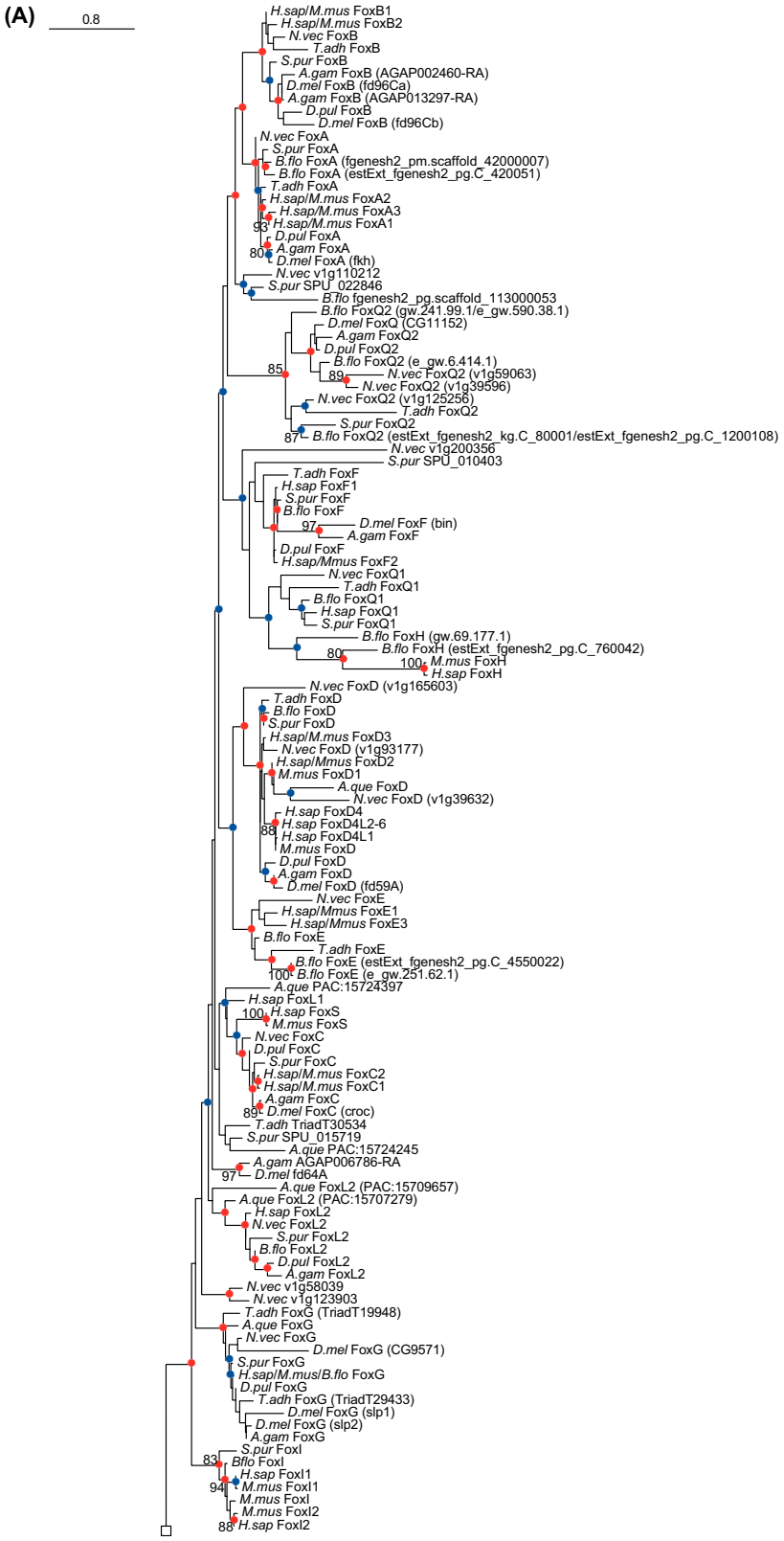


Fig. S2. (Continued)

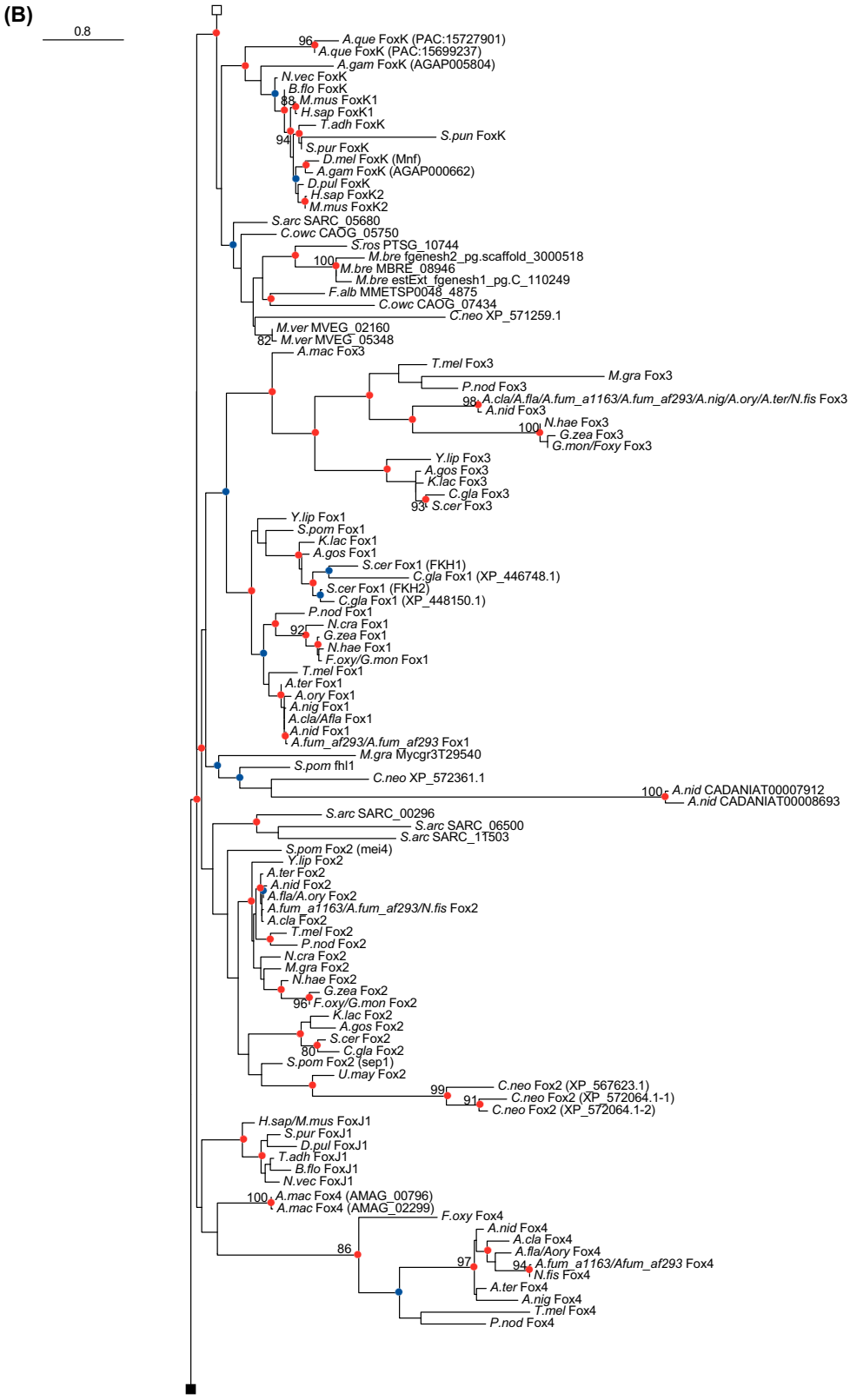


Fig. S2. (Continued)

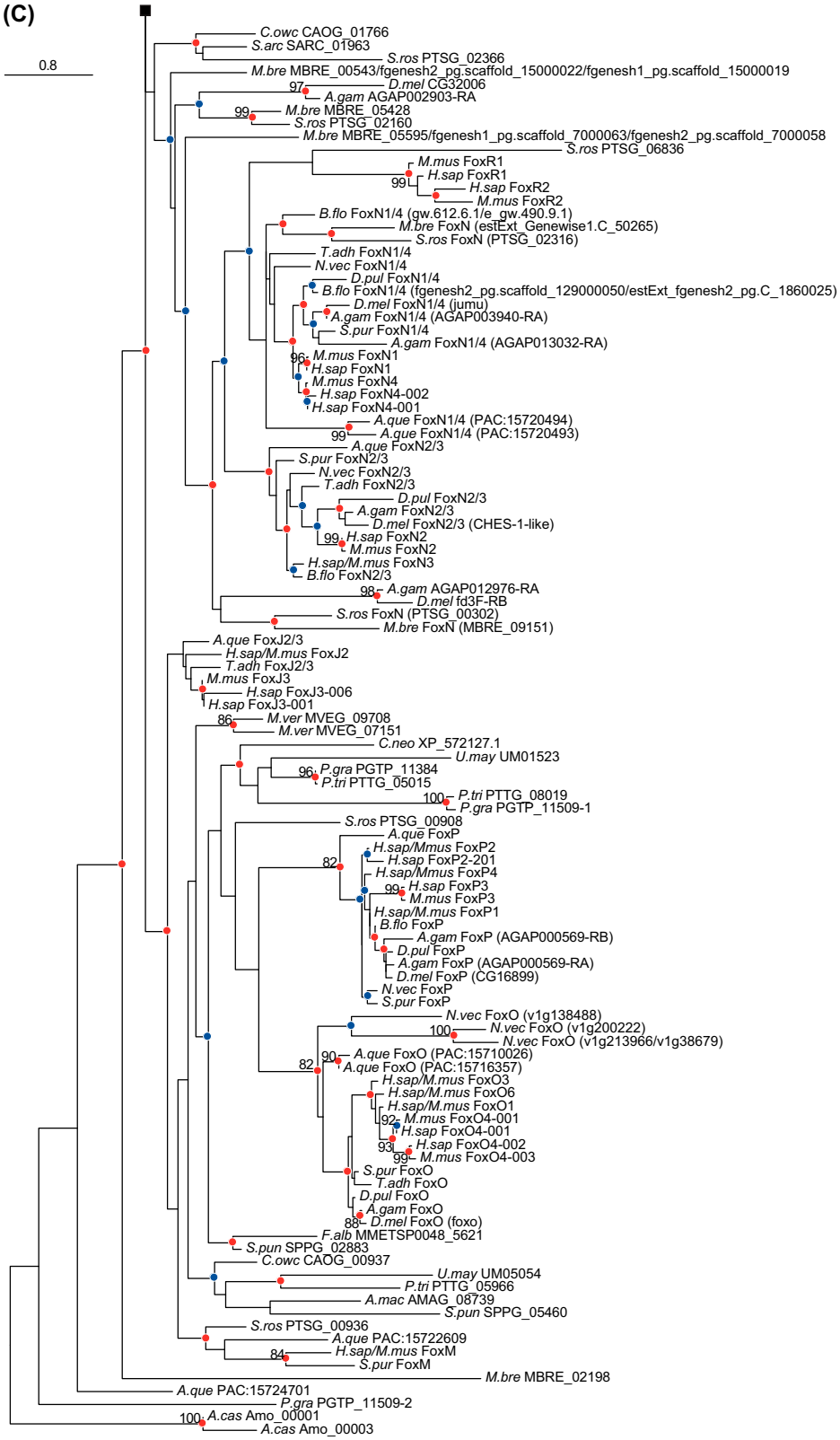


Fig. S2. Maximum likelihood (ML) phylogenetic tree of forkhead domains. The tree is split into three panels for presentation, due to its size. (*A*) The open square continues the branch of the tree represented by the open square in *B*. *A* corresponds to the metazoan-specific clade I. (*B*) The black square continues the branch of the tree represented by the black square in *C*. (*C*) The root and outgroup of the tree.

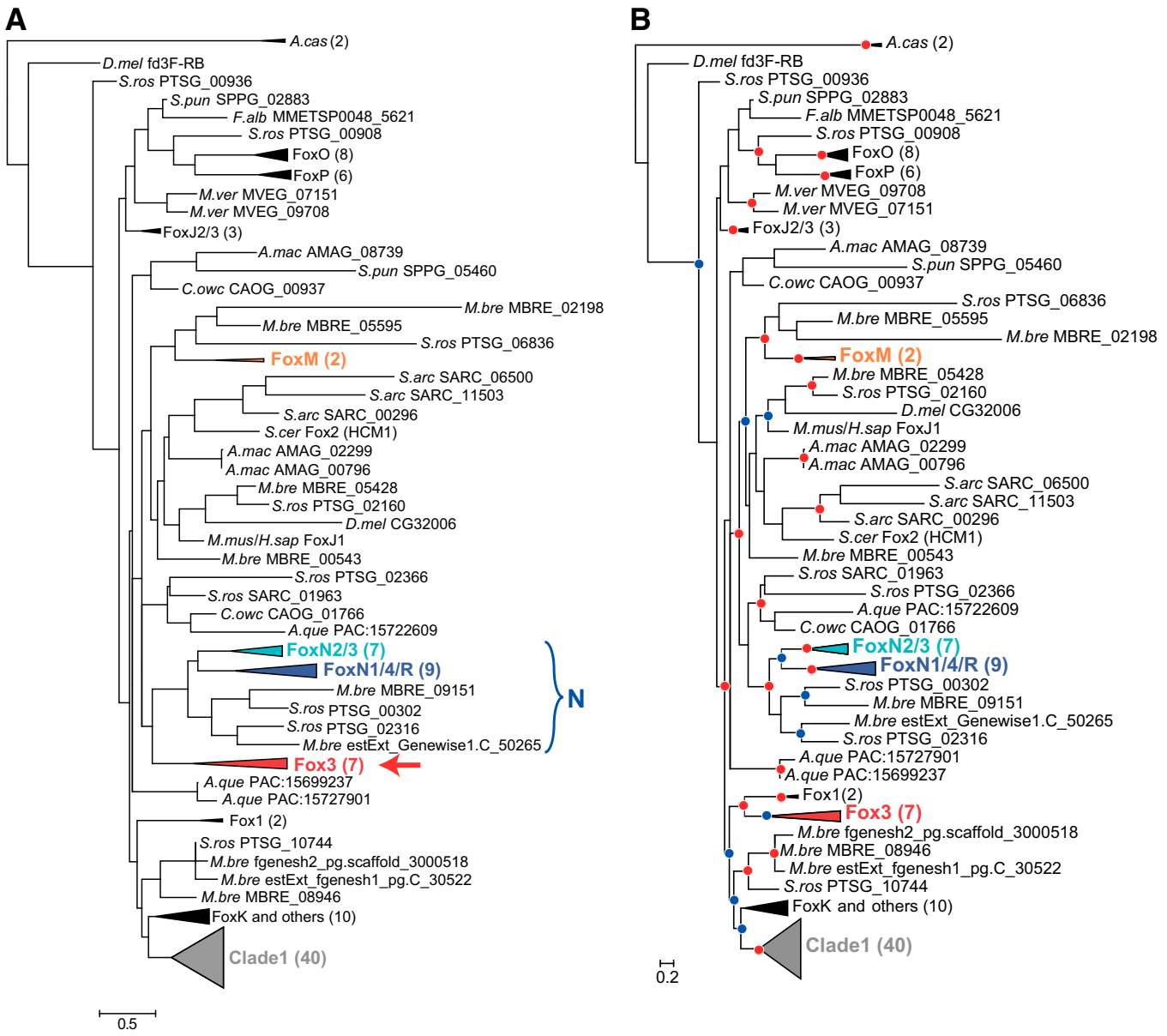


Fig. S3. (Continued)

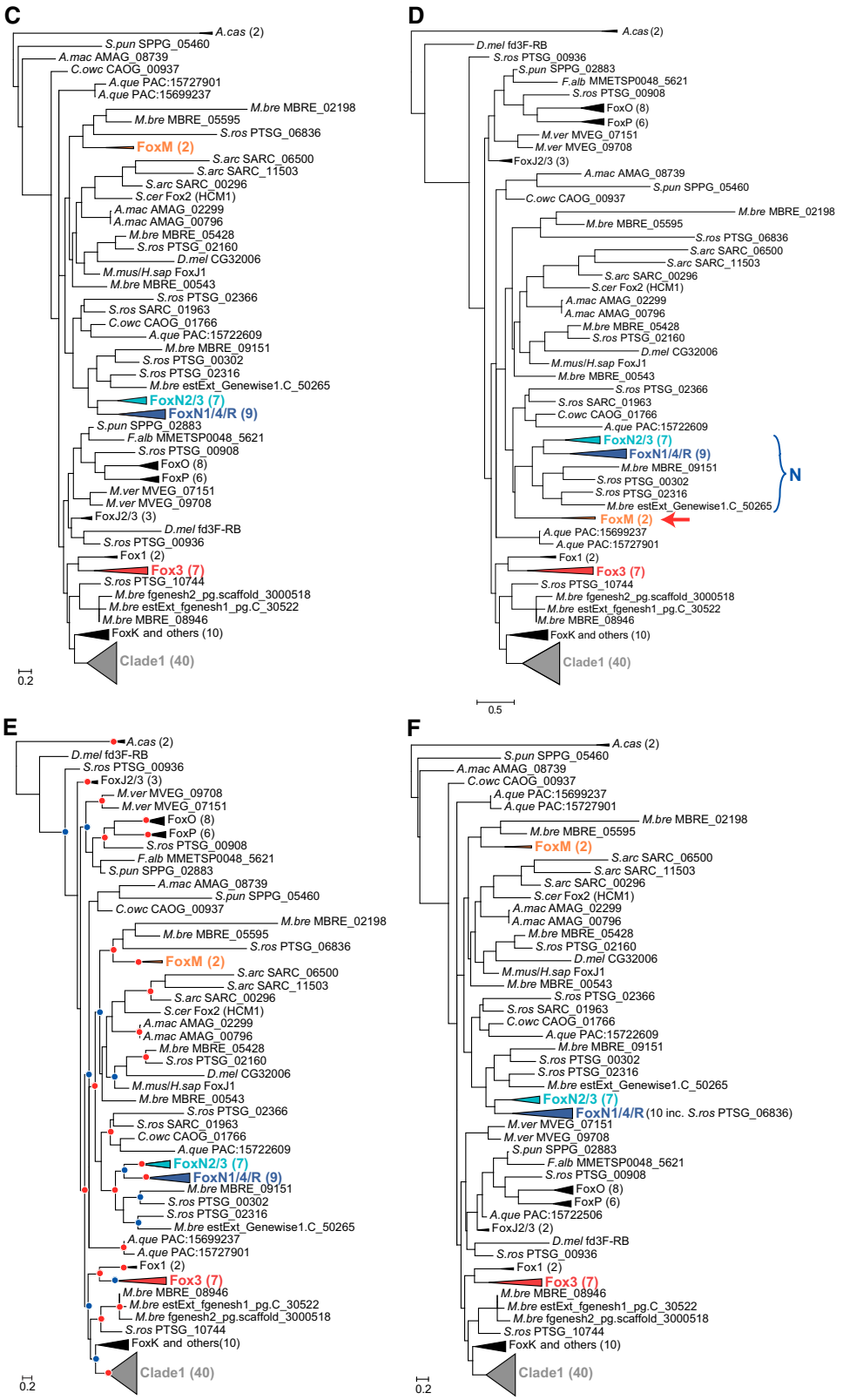


Fig. S3. ML phylogenetic trees testing monophyly of forkhead box N (FoxN) with Fox3 (*A*–*C*) and with FoxM (*D*–*F*). (*A*) A tree was constructed containing a monophyletic FoxN+3 clade, and then branch lengths were optimized and log likelihood was calculated and compared with the ML tree shown in Fig. 2 (see main text). (*B* and *C*) ML optimization of this starting tree with either PhyML (*B*) or RAxML (*C*) yielded trees in which each subfamily is more closely related to FkhP,S-only-binding clades than to each other. (*D*–*F*) Similarly, a starting tree was constructed containing a monophyletic FoxN+M clade (*D*) and ML optimized with PhyML (*E*) or RAxML (*F*).

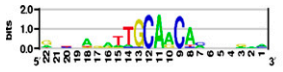
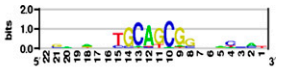
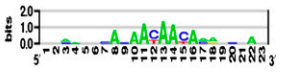
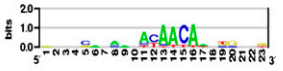
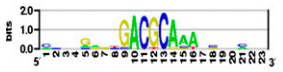
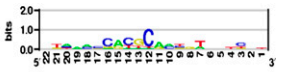
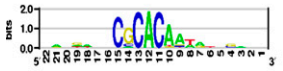
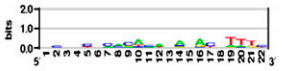
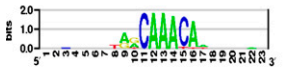
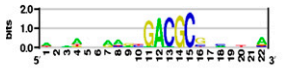
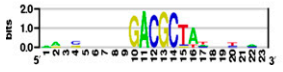
Protein Name	Primary Motif	Secondary Motif	Other published motifs, source
A. cas Fox			
A. mac Fox3			
A. mac Fox4			
A. gos Fox3			
A. nid Fox3			
A. nid Fox4			
D. mel jumu (N1/4)			
D. mel CHES-1-like (N2/3)			
H. sap FoxN2			

Fig. S4. (Continued)

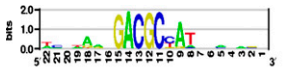
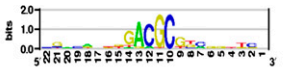
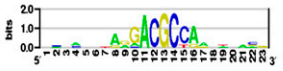
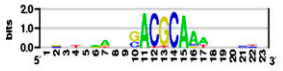
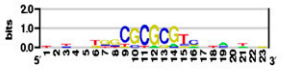
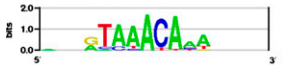
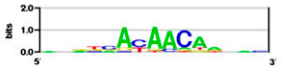
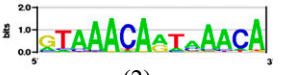
Protein Name	Primary Motif	Secondary Motif	Other published motifs, source
H. sap FoxN4			
H. sap FoxR1			
K. lac Fox3			
M. bre FoxN1/4			
M. mus FoxA2			 (1)
M. mus FoxJ1			
M. mus FoxJ3			 (2)
M. mus FoxJ3			 (2)
M. mus FoxJ3			 (2)

Fig. S4. (Continued)

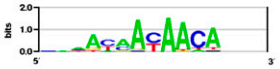
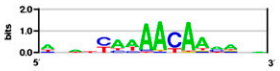
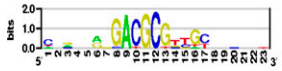
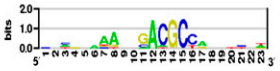
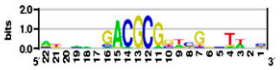
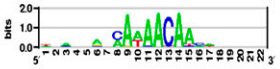
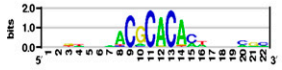
Protein Name	Primary Motif	Secondary Motif	Other published motifs, source
M. mus FoxJ3			 (2)
M. mus FoxK1			 (2)
			 (2)
M. mus FoxL1			
M. mus FoxN1			 (3)
M. mus FoxN4			
M. mus FoxM1			
M. gra Fox3			
N. vec FoxN2/3			

Fig. S4. (Continued)

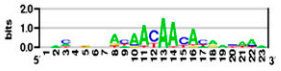
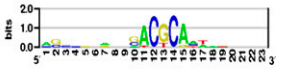
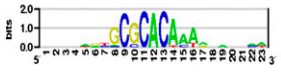
Protein Name	Primary Motif	Secondary Motif	Other published motifs, source
S. cer Fox1,Fkh1			
S. cer Fox1,Fkh2			
S. cer Fox3, Fhl1			
S. cer Fox2, Hcm1			
S. pur FoxM			
T. adh FoxN2/3			
T. mel Fox3			

Fig. S4. Primary and secondary protein-binding microarray (PBM)-derived DNA-binding site motifs for Forkhead proteins.

1. Tuteja G, White P, Schug J, Kaestner KH (2009) Extracting transcription factor targets from ChIP-Seq data. *Nucleic Acids Res* 37(17):e113.
2. Jolma A, et al. (2013) DNA-binding specificities of human transcription factors. *Cell* 152(1-2):327–339.
3. Schlake T, Schorpp M, Nehls M, Boehm T (1997) The nude gene encodes a sequence-specific DNA binding protein with homologs in organisms that lack an anticipatory immune system. *Proc Natl Acad Sci USA* 94(8):3842–3847.

Dataset S1. Fox domains and species used in full and representative ML analyses.

Dataset S1

Dataset S2. PBM experimental details. RefSeq and UniProt IDs are provided, if applicable.

Dataset S2

Dataset S3. PBM 8-mer enrichment (E) scores. Ungapped 8-mer E scores for all forkhead proteins considered in this study.

Datasets S3

Dataset S4. PBM-derived position weight matrices (PWMs) for primary and secondary motifs for each forkhead protein.

Datasets S4