Supplementary information

Species name	Three-character code	Species name (continued)	Code
Acyrthosiphon pisum	Api	*Homo sapiens*	Hsa
Aeropyrum pernix	Ape	*Methanobrevibacter smithii*	Msm
Anabaena variabilis	Ava	*Methanosarcina acetivorans*	Mac
Anopheles gambiae	Aga	*Monodelphis domestica*	Mdo
Apis mellifera	Ame	*Monosiga brevicollis*	Mbr
Arabidopsis thaliana	Ath	*Mus musculus*	Mmu
Atta cephalotes	Ace	*Mycoplasma genitalium*	Mge
Bacillus subtilis	Bsu	*Naegleria gruberi*	Ngr
Bigelowiella natans	Bna	*Nematostella vectensis*	Nve
Bombyx mori	Bmo	*Ornithorhynchus anatinus*	Oan
Branchiostoma floridae	Bfl	*Oryza sativa*	Osa
Buchnera aphidicola	Bap	*Pediculus humanus*	Phu
Caenorhabditis elegans	Cel	*Phaeodactylum tricornutum*	Ptr
Candidatus Korarchaeum cryptofilum	Cko	*Physcomitrella patens*	Ppa
Capitella teleta	Cte	*Pirellula staleyi*	Pst
Chlamydomonas reinhardtii	Cre	*Plasmodium falciparum*	Pfa
Ciona intestinalis	Cin	*Pyrococcus abyssi*	Pab
Cryptococcus neoformans	Cne	*Saccharomyces cerevisiae*	Sce
Culex quinquefasciatus	Cqu	*Schistosoma japonicum*	Sja
Danaus plexippus	Dpl	*Schizosaccharomyces pombe*	Spo
Deinococcus radiodurans	Dra	*Sulfolobus solfataricus*	Sso
Dictyostelium discoideum	Ddi	*Tribolium castaneum*	Tca
Drosophila melanogaster	Dme	*Trichoplax adhaerens*	Tad
Drosophila simulans	Dsi		
Emiliania huxleyi	Ehu		
Escherichia coli	Eco		
Geodermatophilus obscurus	Gob		
Glossina morsitans	Gmo		
Guillardia theta	Gth		
Haloferax volcanii	Hvo		

Supplementary Table 1. Species names in alphabetical order and their corresponding three-character code as used in Figure 1 and 4.

Age groups by Wolf et al. 2009
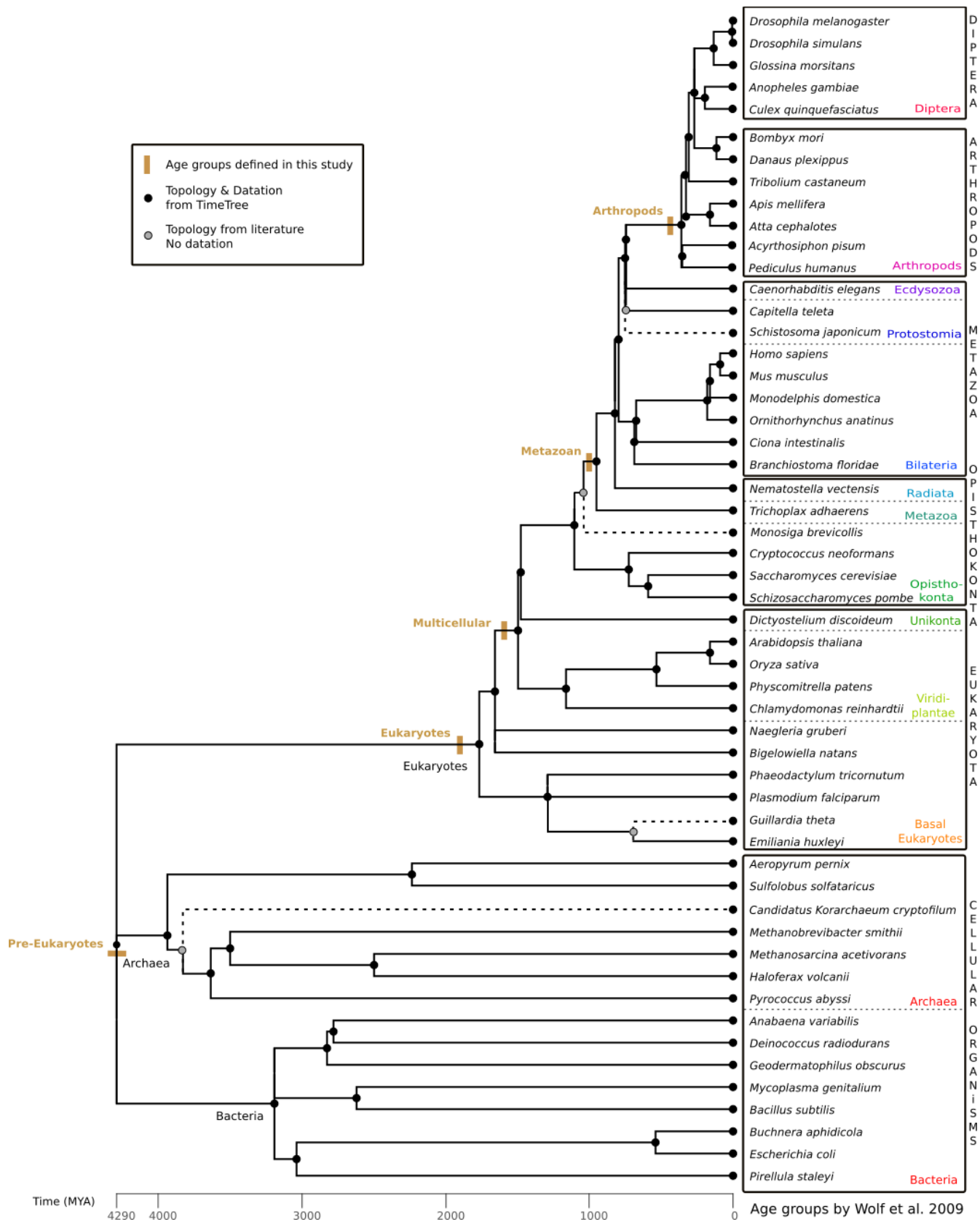
Supplementary Figure 1. Phylogenetic species tree of the 53 species as extracted from Timetree (http://www.treetime.org). Age groups defined in this study are represented as brown bars and labels on the tree. As a validation of our groups, on the right we indicate in black boxes the age groups used by (Wolf et al., 2009) to study homologs of *Drosophila melanogaster* proteins. In the boxes, colored names referred to the thirteen species groups that we defined in the main manuscript, see Fig. 1. At the bottom, time is indicated in Millions of Years Ago (MYA), with species divergence times extracted from Timetree. Species absent from TimeTree were manually added based on literature (dashed branches). For those, no dating data were available and nodes are arbitrarily placed on the timescale (grey nodes). Black nodes are dated by Timetree.

YELLOW	Dollo				
PAM	**Pre-Eukaryotes**	**Eukaryota**	**Opisthokonta**	**Metazoa**	**Arthopods**
Pre-Euk, I+II	3				
Euk, III	4	11			
Multicell, IV	1	5			
Metazoa, V		5	2	3	
Arthopods, VI		1		2	2

RED	Dollo				
PAM	**Pre-Eukaryotes**	**Eukaryota**	**Opisthokonta**	**Metazoa**	**Arthopods**
Pre-Euk, I+II	6				
Euk, III	7	20			
Multicell, IV	3	9			
Metazoa, V		11	4	9	
Arthopods, VI			1	5	8

| GREEN	Dollo				
PAM	**Pre-Eukaryotes**	**Eukaryota**	**Opisthokonta**	**Metazoa**	**Arthopods**
Pre-Euk, I+II	1				
Euk, III	1	3			
Multicell, IV		2			
Metazoa, V		6	1		
Arthopods, VI		1	1	2	7

Supplementary Table 2. Overlap between clustering by Partitioning Around Medoids (PAM) and the Dolly parsimony method for three chromatin types, respectively YELLOW, RED, and GREEN. Numbers indicate the number of shared chromatin-associated proteins, empty cells signal zero shared proteins. See also Figure 3 and 5B in the main text, and Supplementary Table 3.
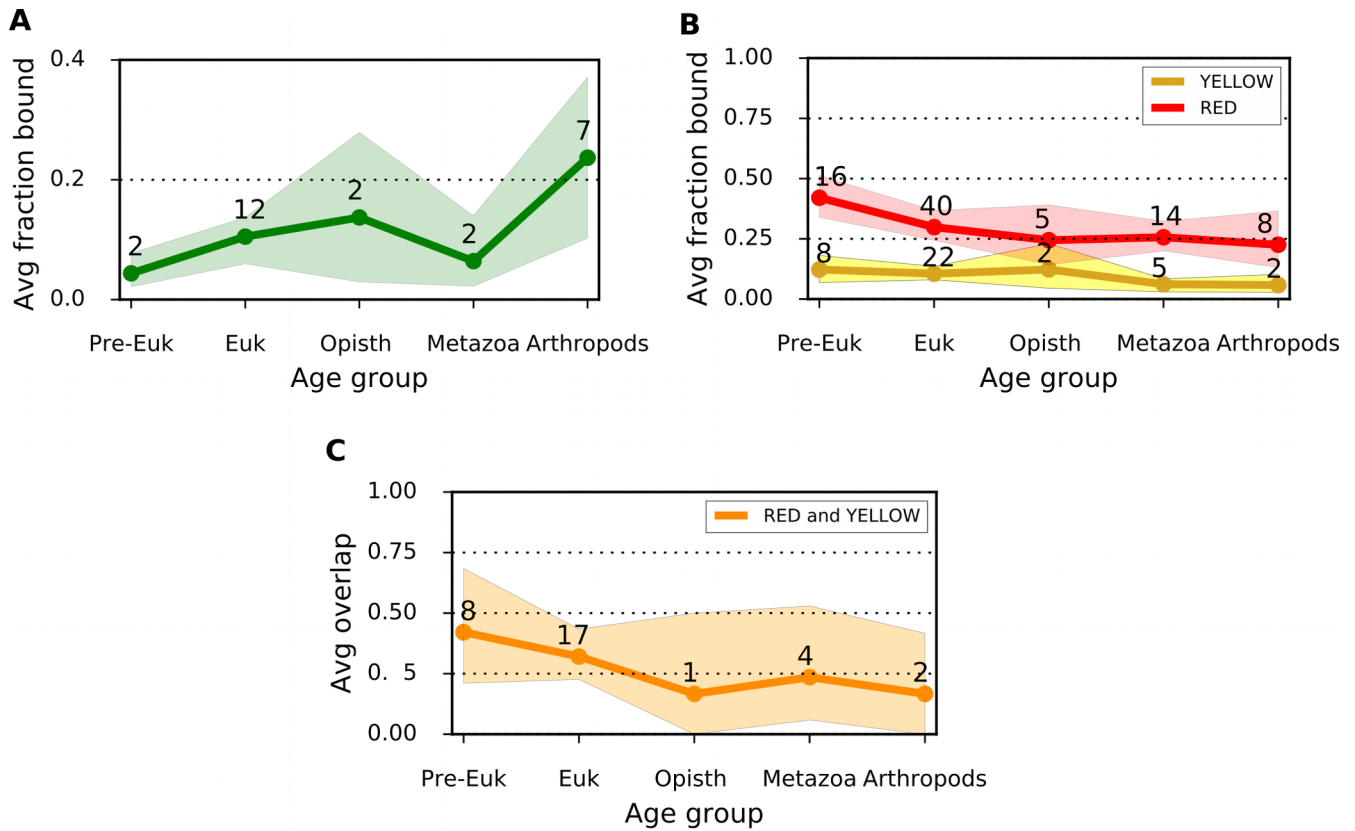
Supplementary Figure 2. Average fraction of genome bound by proteins over evolutionary age groups. The evolutionary age is determined by Dollo parsimony. In contrast in Fig. 3, we use Partitioning Around Medoids. **(A)** GREEN fraction bound. **(B)** RED and YELLOW fraction bound. **(C)** Average overlap between RED and YELLOW. See Methods and Fig. 3 in the main text for details.

DB[1]	A[2]	N[3]	O[4]	U[5]	AVGi[6]	AVGb[7]	Mann-Whitney Test
Panther7	W	38	E*	-	822.0	670.3	U = 6.6e+05 (p = 0.0741)
Multiparanoid	W	12	E***/O*	Dm*	854.8	781.7	U = 6.3e+05 (p = 0.0161)
Jaccard	W	12	E***	Dm*	1145.3	1015.9	U = 6.3e+05 (p = 0.0109)
Lens	W	12	E***	C*/Dm*	1145.3	1015.9	U = 6.3e+05 (p = 0.0109)
OthoMCL	W	12	E*/O*	-	630.9	622.7	U = 6.7e+05 (p = 0.129)
Panther7	D	38	E*	Di**	1329.4	1114.2	U = 6.2e+05 (p = 0.00757)
Multiparanoid	D	12	E***	C*/Dm**	1018.4	937.5	U = 6.3e+05 (p = 0.0111)
Jaccard	D	12	E***	Dm***	1316.7	1154.8	U = 6.1e+05 (p = 0.00341)
Lens	D	12	E***	C*/Dm***	1155.6	1076.8	U = 6.2e+05 (p = 0.00596)
OthoMCL	D	12	E***	C*/Dm*	880.7	817.9	U = 6.4e+05 (p = 0.0191)

Supplementary Table 3. Age enrichment tests for *D. melanogaster* CAPs using different algorithms. The column headings indicate the following, from left to right. DB: The database is DROME_PPODv4, clustered with the corresponding method. A: The algorithm used for enrichment tests, Wagner (W) and Dollo (D). N: Number of species in the species tree. O: overrepresented age groups, Eukaryota (E) and Opisthokonta (O). U: under-represented age groups, *D. melanogaster* (Dm), Cellular organism (C), and Diptera (Di). AVGi: average age input set. AVGb: average age background set. Fisher's exact test was used to calculate the significance of the differences for each age group: *P < 0.05; **P < 0.01; ***P < 0.001.
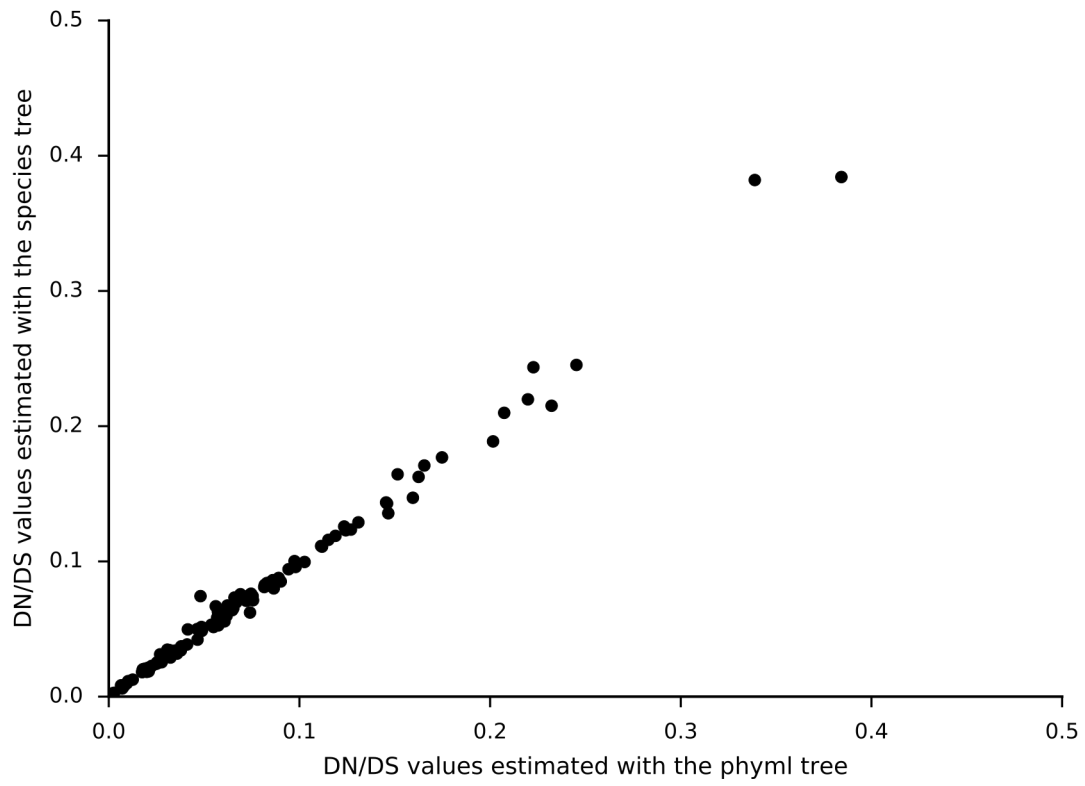
Supplementary Figure 3. Scatter plot of chromatin protein dN/dS ratios obtained using a gene tree topology (PhyML) and the Species tree topology (Timetree).

Foreground branch	Likelihood		2DL	Degrees of freedom	p-value
	Null Model	Model A			
1	-5180.81	-5174.19	13.24	1	$4.91 . 10^{-3}$
2	-5181.25	-5181.63	-0.76	1	1
3	-5178.25	-5175.12	6.26	1	0.22
4	-5180.81	-5174.19	13.24	1	$4.91 . 10^{-3}$
5	-5180.21	-5167.94	24.54	1	$1.31 . 10^{-5}$
6	-5181.38	-5177.30	8.16	1	0.07
7	-5180.17	-5172.50	15.34	1	$1.62 . 10^{-3}$
8	-5181.47	-5181.63	-0.32	1	1
9	-5181.63	-5181.63	0.00	1	1
10	-5180.24	-5162.70	35.08	1	$5.68 . 10^{-8}$
11	-5176.93	-5174.40	5.06	1	0.44
12	-5178.42	-5173.55	9.74	1	0.032
13	-5181.63	-5181.63	0.00	1	1
14	-5181.63	-5181.63	0.00	1	1
15	-5181.41	-5181.60	-0.38	1	1
16	-5181.63	-5181.63	0.00	1	1
17	-5180.50	-5181.47	-1.94	1	1
18	-5181.63	-5181.63	0.00	1	1

Supplementary Table 4. Summary of positive selection tests under the free-ratio model and branch-site model. Foreground branch numbers refer to a top-to-bottom and left-to-right numbering of the branches of the tree in Figure 6A. The red font indicates significant selection tests. P-values are adjusted for multiple testing with the Bonferroni correction.
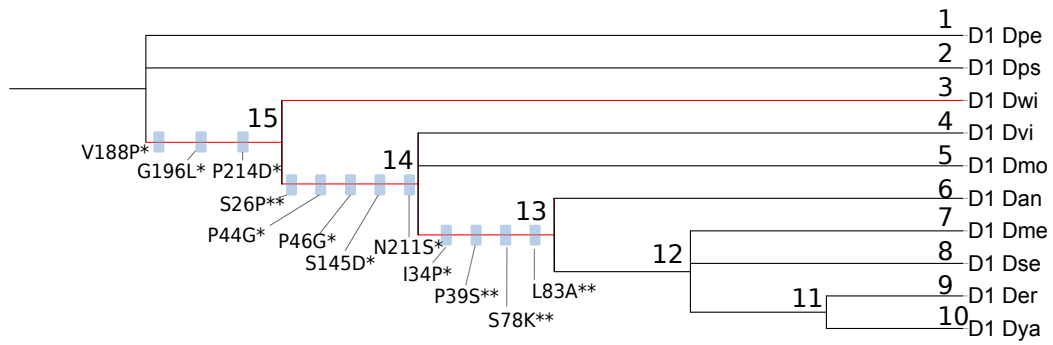
Supplementary Figure 4. Gene tree of D1 constructed with Phyml (with arbitrary branch lengths). The four branches with positive selection events are highlighted in red (p< 0.01, Bonferroni correction). On branches with more than one positively selected site, blue boxes indicate the amino acid substitution under positive selection, with the significance given as posterior probability of dN/dS > 1 (* for Pr > 0.95, ** for Pr > 0.99). Species are indicated by 3 letter abbreviations, see Figure 6A in the main text.

Foreground branch	Likelihood		2DL	Degrees of freedom	p-value
	Null Model	Model A			
1	-5198.60	-5198.60	0.00	1	1
2	-5198.60	-5198.60	0.00	1	1
3	-5195.26	-5188.23	14.06	1	$2.65 \cdot 10^{-3}$
4	-5194.61	-5192.83	3.56	1	0.88
5	-5198.53	-5198.04	0.98	1	1
6	-5192.72	-5190.66	4.12	1	0.63
7	-5198.49	-5198.49	0.00	1	1
8	-5198.60	-5198.60	0.00	1	1
9	-5198.12	-5198.12	0.00	1	1
10	-5198.60	-5181.63	0.00	1	1
11	-5198.60	-5198.60	0.00	1	1
12	-5196.71	-5194.95	3.52	1	0.91
13	-5194.00	-5184.47	19.06	1	$1.91 \cdot 10^{-4}$
14	-5193.77	-5187.46	12.62	1	$5.74 \cdot 10^{-3}$
15	-5195.54	-5188.97	13.14	1	$4.36 \cdot 10^{-3}$

Supplementary Table 5. Summary of positive selection tests using the gene tree topology of Supplementary Figure 4. Foreground branch numbers refer to branches of the tree. The red font indicates significant selection tests. P-values are adjusted for multiple testing with the Bonferroni correction.