Supplement to “Mutation rate and the cost of complexity” by R. Haygood

G, X, and B

Table S1 lists the species used to characterize relationships among haploid genome size G, cell-type number X, and body size B. The first 22 species are animals, the next 7 are brown algae, and the last 8 are plants. Each species is listed with its phylum or, for plants, division, its estimated G derived from the Animal Genome Size Database (Gregory, 2001) or the Plant DNA C-values Database (Bennett and Leitch, 2003), and its estimated X and B from Bell and Mooers (1997) and/or McCarthy and Enquist (2005). For a species with ploidy level P (2 for diploid, 4 for tetraploid, etc.), $G = 2 \times \text{C-value} / P$. All species are known or assumed to be diploid, except *Lemna minor*, whose estimated G derives from the measured C-value of a hexaploid strain. For several species, the databases yield multiple estimates of G, of which I have taken the geometric mean. For several animals, Bell and Mooers and McCarthy and Enquist give different estimates of X or B, of which I have taken the geometric mean.

Figure S1 presents a cladogram of these species. For animals listed by Bell and Mooers, the cladogram differs from that of Bell and Mooers in agreement with the consensus that bilaterian animals consist of deuterostomes (including chordates) and protostomes, which consist of ecdysozoans (including arthropods and nematodes) and lophotrochozoans (including annelids and platyhelminths) (Aguinaldo et al., 1997; de Rosa et al., 1999). Trichotomies reflect uncertainty about the ancestry of placozoans and gastrotrichs. For brown algae and plants, the cladogram differs from that of Bell and Mooers in the additions of *Cutleria* sp. according to Rousseau and de Reviers (1999) and *Hummia onusta* according to Siemer et al. (1998) and in a trichotomy reflecting uncertainty about the ancestry of brown algae.

$N_e u$, X, and B

Table S2 lists the species used to characterize relationships among effective population mutation rate per site $N_e u$, cell-type number X, and body size B, grouped by phylum or, for plants, division. For each higher taxon, each species listed by Lynch (2006) is listed with its estimated $N_e u$, and each species listed by Bell and Mooers (1997) and/or McCarthy and Enquist (2005) is listed with its estimated X and B. The averages and coefficients of variation of $\log[N_e u]$, $\log[X]$, and $\log[B]$ within each higher taxon are also listed. Lynch actually gives estimates of nucleotide diversity, whose expected value for a species with ploidy level P is $2PN_e u$, which I have reduced to estimates of $N_e u$. For several species, Lynch gives multiple estimates, of which I have taken the geometric mean.

Figure S2 presents a cladogram of these phyla and divisions.

References

Aguinaldo, A. M. A., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A., and Lake, J. A., 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**:489–493.

Bell, G., and Mooers, A. O., 1997. Size and complexity among multicellular organisms. *Biological Journal of the Linnean Society* **60**:345–363.

Bennett, M. D., and Leitch, I. J., 2003. Plant DNA C-values database. http://www.rbgkew.org/cval/homepage.html.

de Rosa, R., Grenier, J. K., Andreeva, T., Cook, C. E., Adoutte, A., Akam, M., Carroll, S. B., and Balavoine, G., 1999. Hox genes in brachiopods and priapulids and protostome evolution. *Nature* **399**:772–776.

Gregory, T. R., 2001. Animal genome size database. http://www.genomesize.com.

Lynch, M., 2006. The origins of eukaryotic gene structure. *Molecular Biology and Evolution*. In press.

McCarthy, M. C., and Enquist, B. J., 2005. Organismal size, metabolism and the evolution of complexity in metazoans. *Evolutionary Ecology Research* **7**:681–696.

Rousseau, F., and de Reviers, B., 1999. Phylogenetic relationships within the Fucales (Phaeophyceae) based on combined partial SSU + LSU rDNA sequence data. *European Journal of Phycology* **34**:53–64.

Siemer, B. L., Stam, W. T., and Olsen, J. L., 1998. Phylogenetic relationships of the brown algal orders Ectocarpales, Chordariales, Dictyosiphonales, and Tilopteridales (Phaeophyceae) based on RUBISCO large subunit and spacer sequences. *Journal of Phycology* **34**:1038–1048.

Table S1: Species used to characterize relationships among G, X, and B

species	phylum	G[a]	X	B[b]
Hirudo medicinalis	Annelida	0.225	26	20000
Lumbricus terrestris	Annelida	0.587	41	10000
Araneus diadematus	Arthropoda	2.51	45	1400
Callinectes sapidus	Arthropoda	1.97	69	320000
Carcinus maenas	Arthropoda	1.05	69	140000
Periplaneta americana	Arthropoda	2.66	50	3200
Bufo regularis	Chordata	3.74	154	280000
Canis familiaris	Chordata	3.04	99	50000000
Gallus domesticus	Chordata	1.22	178	850000
Homo sapiens	Chordata	3.42	210	66000000
Lacerta vivipara	Chordata	1.33	162	83000
Morone saxatilis	Chordata	0.899	133	860000
Mus musculus	Chordata	3.20	145	75000
Rana temporaria	Chordata	4.13	153	23000
Scyliorhinus canicula	Chordata	6.37	122	1300000
Sphenodon punctatus	Chordata	4.89	162	500000
Hydra attenuata	Cnidaria	1.81	15	0.063
Turbanella cornuta	Gastrotricha	0.186	18	0.0071
Caenorhabditis elegans	Nematoda	0.0925	24	0.00089
Trichoplax adhaerens	Placozoa	0.0391	4	0.00032
Dugesia lugubris	Platyhelminthes	1.76	14	190
Schmidtea (Dugesia) mediterranea	Platyhelminthes	0.704	14	3.2
Ectocarpus siliculosus	Phaeophyceae	0.245	4	0.32
Colpomenia sinuosa	Phaeophyceae	0.294	5	2000
Scytosiphon lomentaria	Phaeophyceae	0.245	4	790
Hummia onusta	Phaeophyceae	0.270	5	100
Cutleria sp.	Phaeophyceae	0.441	7	3200
Fucus vesiculosus	Phaeophyceae	0.539	7	3200000
Ascophyllum nodosum	Phaeophyceae	0.809	6	630000

species	division	G[a]	X	B[b]
Funaria hygrometrica	Bryophyta	0.395	20	250
Polytrichum commune	Bryophyta	0.456	26	1000
Pinus monophylla	Coniferophyta	29.6	30	–[c]
Equisetum palustre	Equisetophyta	12.3	16	56000
Lemna minor	Magnoliophyta	0.475	18	0.79
Sagittaria lancifolia	Magnoliophyta	13.0	42	100000
Wolffia arrhiza	Magnoliophyta	1.60	5	0.010
Psilotum nudum	Psilophyta	71.2	17	100000

[a]Billions of base pairs, to three significant figures.

[b]Cubic millimeters, to two significant figures.

[c]The value given by Bell and Mooers (1997) is conspicuously incorrect, and the correct value is not readily available, so *P. monophylla* is omitted from analyses involving B.

Figure S1: Cladogram of species in Table S1

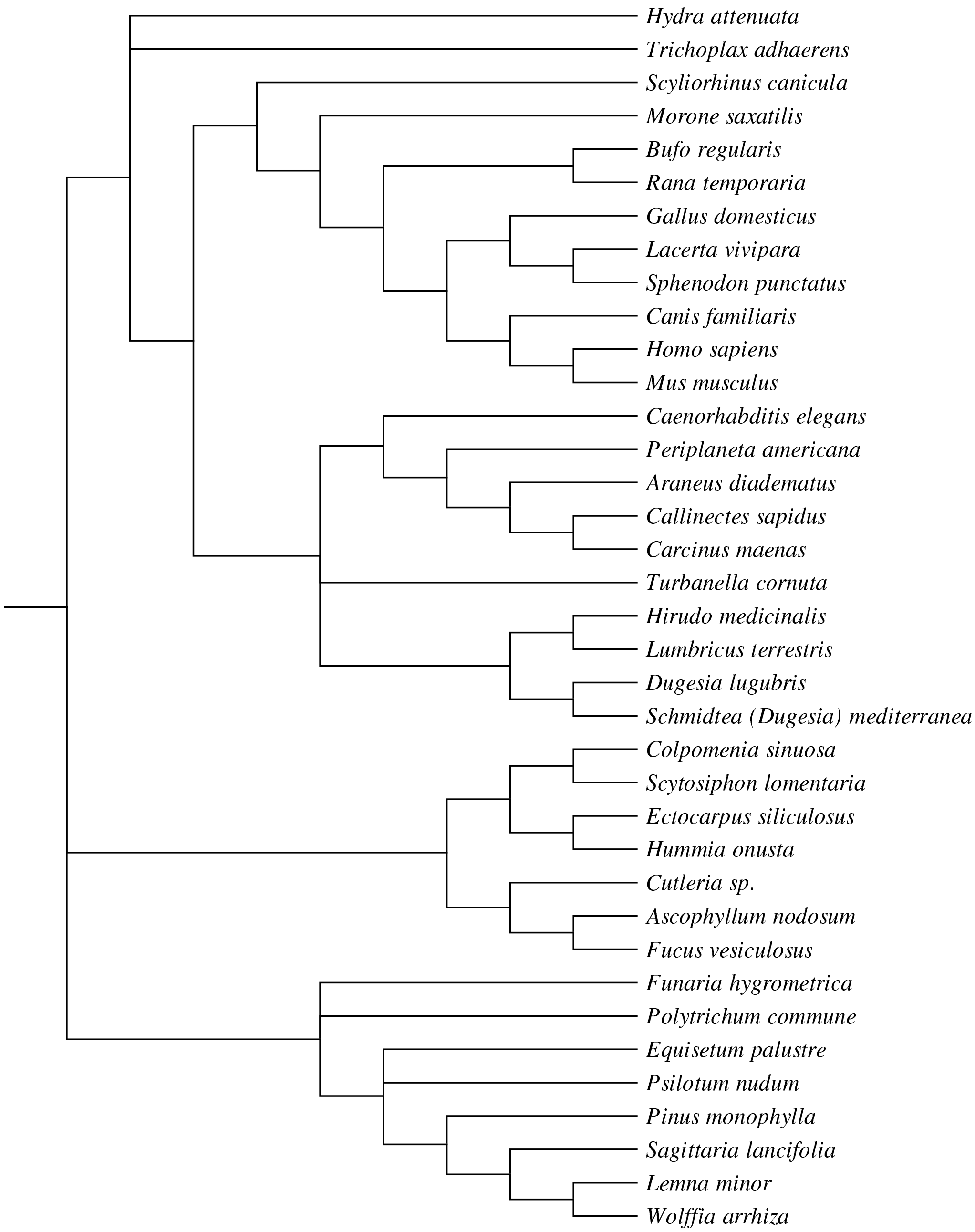

Table S2: Species used to characterize relationships among $N_e u$, X, and B

Acrasiomycota

species	$N_e u$	species	X	B[a]
Dictyostelium discoideum	0.00817	*Acrasis rosea*	2	0.00071
		Dictyostelium discoideum	3	0.013
		Dictyostelium minutum	2	0.0014
$\overline{\log[N_e u]} = -2.09$	$\mathrm{CV}\{\log[N_e u]\} = 0$	$\overline{\log[X]} = 0.36$	$\mathrm{CV}\{\log[X]\} = 0.28$	
		$\overline{\log[B]} = -2.6$	$\mathrm{CV}\{\log[B]\} = 0.19$	

Arthropoda

species	$N_e u$	species	X	B[a]
Aedes aegypti	0.01435	*Araneus diadematus*	45	1400
Anopheles arabiensis	0.00520	*Callinectes sapidus*	69	320000
Anopheles gambiae	0.00587	*Carcinus maenas*	69	140000
Artemia franciscana	0.00365	*Periplaneta americana*	50	3200
Balanus glandula	0.00751	*Pseudocharopinus dentatus*	31	28
Ceratitis capitata	0.00791			
Drosophila kikkawai	0.00848			
Drosophila melanogaster	0.00395			
Drosophila miranda	0.00128			
Drosophila persimilis	0.00209			
Drosphila pseudoobscura	0.00274			
Drosophila simulans	0.00690			
Gryllus firmis	0.00170			
Gryllus pennsylvanicus	0.00174			
Heliconius erato	0.01422			
Heliconius melpomene	0.00611			
Lutzomyia longipalpis	0.01049			
$\overline{\log[N_e u]} = -2.32$	$\mathrm{CV}\{\log[N_e u]\} = 0.140$	$\overline{\log[X]} = 1.7$	$\mathrm{CV}\{\log[X]\} = 0.085$	
		$\overline{\log[B]} = 3.8$	$\mathrm{CV}\{\log[B]\} = 0.44$	

Ascomycota

species	$N_e u$	species	X	B[a]
Auxarthron zuffianum	0.06614	*Gymnoascus reessii*	5	0.016
Candida albicans	0.00174	*Leptosphaeria* sp.	9	0.011
Candida glabrata	0.00354			
Coccidioides immitis	0.00445			
Fusarium cerealis	0.00484			
Fusarium graminearum	0.00390			
Histoplasma capsulatum	0.01276			
Mycosphaerella graminicola	0.00873			
Neurospora crassa	0.00793			
Neurospora discreta	0.00653			
Neurospora intermedia	0.00403			
Neurospora sitophila	0.00051			
Saccharomyces cerevisiae	0.00932			
Saccharomyces paradoxus	0.00097			
Uncinocarpus reesii	0.01412			
$\overline{\log[N_e u]} = -2.29$	$\mathrm{CV}\{\log[N_e u]\} = 0.221$	$\overline{\log[X]} = 0.83$	$\mathrm{CV}\{\log[X]\} = 0.22$	
		$\overline{\log[B]} = -1.9$	$\mathrm{CV}\{\log[B]\} = 0.026$	

Basidiomycota

species	$N_e u$	species	X	B[a]
Cryptococcus neoformans	0.00776	*Sphaerolobus stellatus*	9	1.3
Heterobasidion annosum	0.01546			
$\overline{\log[N_e u]} = -1.96$	$\mathrm{CV}\{\log[N_e u]\} = 0.108$	$\overline{\log[X]} = 0.95$	$\mathrm{CV}\{\log[X]\} = 0$	
		$\overline{\log[B]} = 0.10$	$\mathrm{CV}\{\log[B]\} = 0$	

Bryophyta

species	$N_e u$	species	X	B[a]
Ceratodon purpureus	0.00589	*Fegatella conica*	15	3.2
		Funaria hygrometrica	20	250
		Monoclen forsteri	19	3.2
		Pogonatum stevensii	21	710
		Polytrichum commune	26	1000
		Sphagnum recurvum	11	890
		Symphogyna brogniarti	13	0.45
$\overline{\log[N_e u]} = -2.23$	$\mathrm{CV}\{\log[N_e u]\} = 0$	$\overline{\log[X]} = 1.2$	$\mathrm{CV}\{\log[X]\} = 0.11$	
		$\overline{\log[B]} = 1.7$	$\mathrm{CV}\{\log[B]\} = 0.18$	

Chlorophyta

species	$N_e u$	species	X	B[a]
Chlamydomonas debaryana	0.00733	*Astrophomene gubernaculum*	2	0.000045
Chlamydomonas reinhardtii	0.01122	*Eudorina illinoisiensis*	2	0.00010
		Fritschiella tuberosa	5	0.00020
		Microthamnion kutzingianus	3	0.000063
		Pleodorina sphaerica	2	0.00035
		Ulothrix zonata	3	0.000032
		Volvox aureus	2	0.00056
$\overline{\log[N_e u]} = -2.04$	$\mathrm{CV}\{\log[N_e u]\} = 0.0639$	$\overline{\log[X]} = 0.41$	$\mathrm{CV}\{\log[X]\} = 0.37$	
		$\overline{\log[B]} = -3.9$	$\mathrm{CV}\{\log[B]\} = 0.23$	

Chordata

species	$N_e u$	species	X	B[a]
Aethia cristatella	0.00329	*Bufo regularis*	154	280000
Aethia pusilla	0.00096	*Canis familiaris*	99	50000000
African cichlids	0.00123	*Gallus domesticus*	178	850000
Agelaius phoeniceus	0.00155	*Homo sapiens*	210	66000000
Canis familiaris	0.00029	*Lacerta vivipara*	162	83000
Eleutherodactylus sp.	0.00053	*Lebistes reticulatus*	138	2000
Emys marmorata	0.00012	*Morone saxatilis*	133	860000
Erinaceus concolor	0.00221	*Mus musculus*	145	75000
Erinaceus europaeus	0.00165	*Rana temporaria*	153	23000
Ficedula albicollis	0.00067	*Salmo gairdneri*	116	250000
Ficedula hypoleuca	0.00063	*Scoliodon terraenovae*	122	6300000
Fugu rubripes	0.00200	*Scyliorhinus canicula*	122	1300000
Gallus gallus	0.00166	*Sphenodon punctatus*	162	500000
Gorilla gorilla	0.00041			
Homo sapiens	0.00028			
Loxodonta africana	0.00047			
Loxodonta cyclotis	0.00074			
Mirounga leonine	0.00021			
Mus domesticus	0.00240			
Oncorhynchus tshawytscha	0.00086			
Ovis aries	0.00102			
Pan paniscus	0.00021			
Pan troglodytes	0.00035			
Prochilodus rubrotaeniatus	0.00148			
Rattus norvegicus	0.00032			
Salmo salar	0.00013			
$\overline{\log[N_e u]} = -3.17$	$\mathrm{CV}\{\log[N_e u]\} = 0.128$	$\overline{\log[X]} = 2.2$	$\mathrm{CV}\{\log[X]\} = 0.040$	
		$\overline{\log[B]} = 5.7$	$\mathrm{CV}\{\log[B]\} = 0.22$	

Coniferophyta

species	$N_e u$	species	X	B[a]
Cryptomeria japonica	0.00098	*Pinus monophylla*	30	–[b]
Pinus pinaster	0.00165			
Pinus sylvestris	0.00097			
$\overline{\log[N_e u]} = -2.94$	$\mathrm{CV}\{\log[N_e u]\} = 0.0454$	$\overline{\log[X]} = 1.5$	$\mathrm{CV}\{\log[X]\} = 0$	

Echinodermata

species	$N_e u$	species	X	B[a]
Strongylocentrotus franciscanis	0.00315	*Asterias rubens*	46	3500000
		Echinus esculentus	43	2900000
		Holothuria forskali	48	290000
$\overline{\log[N_e u]} = -2.50$	$\mathrm{CV}\{\log[N_e u]\} = 0$	$\overline{\log[X]} = 1.7$	$\mathrm{CV}\{\log[X]\} = 0.015$	
		$\overline{\log[B]} = 6.2$	$\mathrm{CV}\{\log[B]\} = 0.098$	

Magnoliophyta

species	$N_e u$	species	X	B[a]
Aegilops tauschii	0.00599	*Croomia pauciflora*	42	16000
Arabidopsis thaliana	0.00311	*Fuirena ciliaris*	44	25000
Arabidopsis halleri	0.00305	*Lemna minor*	18	0.79
Arabidopsis lyrata petraea	0.00739	*Lomandra hermaphroditicum*	36	35000
Arabidopsis lyrata lyrata	0.00105	*Mammillaria elongata*	27	63000
Arabis gemmifera	0.00080	*Petermannia cirrhosa*	39	25000
Betula pendula	0.00110	*Sagittaria lancifolia*	42	100000
Dioscorea tokoro	0.00108	*Selenipedium palmifolium*	35	13000
Dunnia sinensis	0.00192	*Wolffia arrhiza*	5	0.010
Glycine max	0.00206	*Wolffia microscopica*	7	0.071
Helianthus annuus	0.00860	*Wolfiella welwitschii*	8	0.40
Hordeum vulgare	0.00259			
Leavenworthia crassa	0.00744			
Leavenworthia stylosa	0.00797			
Leavenworthia uniflora	0.00927			
Lycopersicon chmielewskii	0.00039			
Lycopersicon pimpinellifolium	0.00043			
Lycopersicon chilense	0.00165			
Lycopersicon hirsutum	0.00149			
Lycopersicon peruvianum	0.00680			
Manihot esculenta	0.00410			
Mimulus guttatus	0.01940			
Mimulus nasutus	0.00168			
Oryza rufipogon	0.00099			
Oryza sativa	0.00158			
Phaseolus vulgaris	0.01275			
Populus tremula	0.00409			
Schiedea globosa	0.00077			

species	$N_e u$			
Silene latifolia	0.00426			
Silene dioica	0.00627			
Solanum pimpinellifolium	0.00161			
Sorghum bicolor	0.00047			
Zea mays	0.00451			
Zea diploperennis	0.00313			
$\overline{\log[N_e u]} = -2.58$	$\text{CV}\{\log[N_e u]\} = 0.170$	$\overline{\log[X]} = 1.3$	$\text{CV}\{\log[X]\} = 0.27$	
		$\overline{\log[B]} = 2.5$	$\text{CV}\{\log[B]\} = 0.33$	

Mollusca

species	$N_e u$	species	X	B[a]
Biomphalaria pfeifferi	0.00973	*Anodonta cygnea*	33	5300
Crassostrea virginica	0.02711	*Helix aspersa*	41	28000
		Sepia officinalis	78	23000
$\overline{\log[N_e u]} = -2.39$	$\text{CV}\{\log[N_e u]\} = 0.132$	$\overline{\log[X]} = 1.7$	$\text{CV}\{\log[X]\} = 0.12$	
		$\overline{\log[B]} = 4.2$	$\text{CV}\{\log[B]\} = 0.095$	

Nematoda

species	$N_e u$	species	X	B[a]
Caenorhabditis briggsae	0.00092	*Caenorhabditis elegans*	24	0.00089
Caenorhabditis elegans	0.00023			
Caenorhabditis remanei	0.00714			
Haemonchus contortus	0.01009			
Haemonchus placei	0.01188			
$\overline{\log[N_e u]} = -2.55$	$\text{CV}\{\log[N_e u]\} = 0.297$	$\overline{\log[X]} = 1.4$	$\text{CV}\{\log[X]\} = 0$	
		$\overline{\log[B]} = -3.0$	$\text{CV}\{\log[B]\} = 0$	

[a]Cubic millimeters, to two significant figures.

[b]The value given by Bell and Mooers (1997) is conspicuously incorrect, and the correct value is not readily available, so *P. monophylla* is omitted from analyses involving B.

Figure S2: Cladogram of phyla and divisions in Table S2