

SUPPLEMENTARY INFORMATION

Accessory Tables

Table A1. C-values vs genome size.

Species	C value	Genome size (Gb)
Mouse	3.3	2.5
Opossum	4.3	3.4
Chimpanzee	3.76	3.0
Chicken	1.25	1.0

Table A2. Sequencing data by insert size.

Insert Size(kb)	Count number	Assembled Num	Trimmed bases	Paired read num	Paired bases
1.5-2.0	13688	12769	7864169	11266	7147302
2.0-2.5	30860	27868	16566633	23460	14456677
2.0-3.0	62952	57781	37405066	51330	33991476
2.0-4.0	123069	114217	78470191	104279	72503023
2.5-3.0	31829	28639	17244301	24360	15097909
3.0-3.5	15760	14514	8987717	12686	8086119
3.0-4.0	1298245	1166093	681919515	961656	578014809
3.5-4.0	573	321	132130	270	108856
4.0-4.3	11964296	11048205	7422999427	9907029	6759757330
4.3-6.5	7564965	7058321	5164607068	6203560	4585400739
6.5-9.5	4787407	4352061	2959216736	3667469	2531831207
40	652982	493809	304165771	311447	196298750
147	407649	379614	269859963	297082	213731642
total	26954275	24754212	16969438687	21575894	15016425839

Table A3. Sequence contiguity by species.

Platypus Contiguity stat	Ctg N50 len	Sctg N50 Len	Ctg N50 Num	Sctg N50 Num
platypus_051215(1.7G)	13,026	365,576	33,796	736
Mouse 7X(2.5G)	24,775	11,479,562	29,556	66
Chimp050130(2.9G) 6.7X	28,231	13,733,919	29,244	60
Chicken040612(1G) 6.6X	38,867	10,918,580	6,835	26
chimp050412(no fosmid 3.0G) 6.7X	28,238	3,323,878	29,362	211
Platypus Contiguity stat	Largest Ctg	Ave. Ctg	Largest Sctg	Ave. Sctg
platypus_051215(1.7G)	245,967	6,278	14,341,343	28,661
Mouse 7X(2.5G)	225,618	11,576	44,729,073	78,832
Chimp050130(2.9G) 6.7X	380,600	9,425	59,100,486	29,107
Chicken040612(1G) 6.6X	441,790	12,007	50,876,398	42,110
chimp050412(no fosmid 3.0G) 6.7X	288,304	8,765	25,457,179	21,251

Table A4. Total sequence assigned to platypus chromosomes.

Sequence per
Chromosome

Chr	Sequence*
1	47594283
2	54797317
3	59581953
4	58987262
5	24609220
6	16302927
7	40039088
10	11243762
11	6809224
12	15872666
14	2696122
15	3786880
17	1399469
18	6611290
20	1816412
X1	45541551
X2	5652501
X3	5951358
X5	27786739
Un	1559714169

*Sequence lengths include gap size estimates

Table A5. G+C content distribution of finished BACs: “Covered” vs “Not covered” regions.

%G+C	Not covered		Covered	
	num	%age	num	%age
10-19	2	0.1	0	0
20-29	8	0.6	0	0
30-39	66	4.9	54	3.7
40-49	592	44	1045	70.7
50-59	438	32.6	283	19.1
60-69	181	13.5	85	5.8
70-79	51	3.8	11	0.7
80-89	6	0.4	0	0

Table A6. Repeat content based on the platypus RepeatMasker library.

Covered			
Covered regions of finished platypus BACs aligning to the WGS sequence			
total_length:	11687519		
GC_level:	44.38		
bases_masked:		5268945	45.08
SINEs:	13041	2516541	21.53
LINEs:	12251	2479459	21.21
LTR_elements:	25	13698	0.12
DNA_elements:	354	72473	0.62
Unclassified:	443	51518	0.44

Total_interspersed_repeats:	NA	5133689	43.92
Small_RNA:	0	0	0

Satellites:	71	7117	0.06
Simple_repeats:	1907	77311	0.66
Low_complexity:	795	51330	0.44

Not Covered (Gaps)

(Regions of finished platypus BACs not aligning to the WGS sequence)

Total_length:	1213497		
GC_level:	49.85		
Bases_masked:		765383	63.07

SINES:	1635	271373	22.36
LINEs:	1463	454461	37.45
LTR_elements:	5	2028	0.17
DNA_elements:	23	4300	0.35
Unclassified:	36	2948	0.24

Total_interspersed_repeats:	NA	735110	60.58
-----------------------------	----	--------	-------

Small_RNA:	0	0	0
------------	---	---	---

Satellites:	4	613	0.05
Simple_repeats:	279	12463	1.03
Low_complexity:	215	17256	1.42

Table A7. Platypus segmental duplication analysis by chromosome assignment.

chrom	size	size (gap-free)	All WGAC (>90% > 1 kb)	WGAC (<94% <10 kb)	WGAC (>94% > 10 kb)	WSSD	WGAC+WSSD	Total SD	% SD
	47,594,283	44,500,396	2,409,545	44,445	55,609	0	0	44,445	0.10%
chr2	54,797,317	50,943,648	2,947,515	61,079	107,719	74,917	53,300	114,379	0.22%
chr3	59,581,953	55,753,503	3,970,872	157,160	254,692	170,811	130,311	287,471	0.52%
chr4	58,987,262	55,339,946	3,373,590	101,053	126,600	0	0	101,053	0.18%
chr5	24,609,220	22,751,819	1,051,886	19,770	23,747	0	0	19,770	0.09%
chr6	16,302,927	15,302,526	1,075,579	31,053	63,676	73,252	58,449	89,502	0.58%
chr7	40,039,088	37,716,602	2,428,143	70,516	120,910	78,862	55,796	126,312	0.33%
chr10	11,243,762	10,686,721	728,385	16,340	11,036	0	0	16,340	0.15%
chr11	6,809,224	6,329,107	319,445	6,126	0	0	0	6,126	0.10%
chr12	15,872,666	15,079,184	1,186,676	48,638	128,452	72,624	48,514	97,152	0.64%
chr14	2,696,122	2,492,320	38,725	0	10,912	15,000	0	0	0.00%
chr15	3,786,880	3,543,753	154,816	5,858	0	0	0	5,858	0.17%
chr17	1,399,469	1,274,415	32,044	2,187	0	0	0	2,187	0.17%
chr18	6,611,290	6,314,477	553,310	2,073	21,649	0	0	2,073	0.03%
chr20	1,816,412	1,654,293	92,022	0	0	0	0	0	0.00%
chrX1	45,541,551	42,774,007	2,477,972	120,243	122,529	0	0	120,243	0.28%
chrX2	5,652,501	5,467,671	414,593	0	0	23,373	0	0	0.00%
chrX3	5,951,358	5,824,245	507,442	1,854	0	0	0	1,854	0.03%
chrX5	27,786,739	25,788,079	199,997	14,873	0	10,133	0	14,873	0.06%
subtotal	437,080,024	409,536,712	23,962,557	703,268	1,047,531	518,972	346,370	1,049,638	0.26%
all contigs	894,359,526	786,450,919	121,274,573	24,988,227	14,207,330	12,907,044	7,460,009	32,448,236	4.13%
all ultras	689,150,142	646,450,220	45,273,432	2,250,071	3,692,305	3,145,198	2,219,933	4,470,004	0.69%
total	2,020,589,692	1,842,437,851	190,510,562	27,941,566	18,947,166	16,571,214	10,026,312	37,967,878	2.06%

Total SD=total duplication content based on all WGAC duplications <94% (column 6) and all duplications >94% and > 10

kb that are confirmed by both WGAC and WSSD. % SD represents a lower bound estimate of duplication content.

Table A8. Chimeric clones in chimp and platypus assemblies.

	Chimp	Platypus	Platypus/Chimp
Chimera	83,570	196,953	2.36
Contig Ends	1,083,550	1,092,049	1.01
Both	2,395	5,213	2.18

*both=chimeric reads at both ends of a given contig

Table A9. Orientation of reads at contig ends.

count	platy_start	percentage
119,456	0 0 0	23.46
56,575	0 0 1	11.11
76,137	0 1 0	14.95
21,594	0 1 1	4.24
41,472	1 0 0	8.14
13,198	1 0 1	2.59
139,724	1 1 0	27.44
41,127	1 1 1	8.08
509,283	total	100.00

count	chimp_start	percentage
111,557	0 0 0	22.12
44,473	0 0 1	8.82
97,379	0 1 0	19.31
32,511	0 1 1	6.45
58,706	1 0 0	11.64
23,660	1 0 1	4.69
98,851	1 1 0	19.60
37,155	1 1 1	7.37
504,292	total	100.00

count	platy_end	percentage
41,391	0 0 0	8.13
138,810	0 0 1	27.26
13,440	0 1 0	2.64
42,258	0 1 1	8.30
123,856	1 0 0	24.32
71,021	1 0 1	13.95
57,381	1 1 0	11.27
21,126	1 1 1	4.15
509,283	total	100.00

count	chimp_end	percentage
56,761	0 0 0	11.26
94,667	0 0 1	18.77
23,784	0 1 0	4.72
36,854	0 1 1	7.31
117,301	1 0 0	23.26
97,954	1 0 1	19.42
44,467	1 1 0	8.82
32,504	1 1 1	6.45
504,292	total	100.00

Table A10. Ensembl gene sets used for orthology analysis

Species	Genes	Transcripts	Assembly
<i>H. sapiens</i>	22810	43605 ncbi36 (hg18)	
<i>M. musculus</i>	24442	32078 ncbi36 (m36)	
<i>C. familiaris</i>	19314	25568 canFam2, Broad Institute (May 2005)	
<i>M. domestica</i>	19597	32690 monDom4, Broad Institute	
<i>O. anatinus</i>	18597	27557 WashU assembly	
<i>G. gallus</i>	16715	22186 galGal3, Chicken genome sequencing consortium (May 2006)	

Table A11. Simple 1:1 orthologs across the amniota. Each set contains exactly one gene from each species.

Species	(1:1)n ortholog sets	
<i>H.sapiens</i>	22,611	100%
+ <i>M. musculus</i>	14,905	66%
+ <i>C. familiaris</i>	13,788	61%
+ <i>M. domestica</i>	12,042	53%
+ <i>O. anatinus</i>	9,264	41%
+ <i>G. gallus</i>	7,587	34%

Table A12. Different ways to estimate genome-wide substitution rates.

Method	Aggregation	Ortholog sets	Fit on tree
Method 1 / Pairwise	Median over gene pairs	Pairwise	No
Method 2 / PairwiseStrict	Median over gene pairs	Multiple	No
Method 3 / TreeMedian	Median over gene trees	Multiple	Yes
Method 4 / Sample Average	20 samples of 200 concatenated gene alignments	Multiple	Yes

Table A13. Estimated genome-wide d_S and d_N for different methods (Supplementary Notes S18). Pairwise rates between genomes have been read off the tree for tree based methods.

Species	Species	Pairwise		PairwiseStrict		TreeMedian		SampleAverage	
		dS	dN	dS	dN	dS	dN	dS	dN
<i>C. familiaris</i>	<i>G. gallus</i>	1.741	0.053	1.547	0.043	1.411	0.032	1.314	0.056
<i>C. familiaris</i>	<i>H. sapiens</i>	0.369	0.041	0.338	0.032	0.315	0.022	0.344	0.044
<i>C. familiaris</i>	<i>M. domestica</i>	1.113	0.063	0.986	0.051	0.867	0.036	0.879	0.067
<i>C. familiaris</i>	<i>M. musculus</i>	0.667	0.088	0.637	0.074	0.596	0.062	0.605	0.106
<i>C. familiaris</i>	<i>O. anatinus</i>	1.461	0.103	1.277	0.087	1.126	0.076	1.121	0.129
<i>G. gallus</i>	<i>H. sapiens</i>	1.642	0.094	1.469	0.078	1.381	0.062	1.292	0.107
<i>G. gallus</i>	<i>M. domestica</i>	1.586	0.099	1.426	0.083	1.345	0.088	1.254	0.145
<i>G. gallus</i>	<i>M. musculus</i>	1.891	0.111	1.687	0.092	1.662	0.102	1.553	0.167
<i>G. gallus</i>	<i>O. anatinus</i>	1.707	0.102	1.497	0.086	1.442	0.088	1.311	0.145
<i>H. sapiens</i>	<i>M. domestica</i>	1.037	0.095	0.918	0.080	0.837	0.082	0.858	0.136
<i>H. sapiens</i>	<i>M. musculus</i>	0.575	0.112	0.543	0.097	0.534	0.105	0.530	0.163
<i>H. sapiens</i>	<i>O. anatinus</i>	1.406	0.124	1.220	0.107	1.096	0.119	1.100	0.185
<i>M. domestica</i>	<i>M. musculus</i>	1.305	0.119	1.199	0.103	1.118	0.105	1.119	0.163
<i>M. domestica</i>	<i>O. anatinus</i>	1.308	0.111	1.152	0.097	1.061	0.099	1.062	0.154
<i>M. musculus</i>	<i>O. anatinus</i>	1.696	0.107	1.454	0.093	1.377	0.105	1.361	0.158

Table A14. Branch-specific genome-wide d_N/d_S values across the amniota for two different methods (Methods 3 and 4; see **Table A13**).

<i>Branch</i>	<i>TreeMedian</i>		<i>SampleAverage</i>	
	Median	Stddev	Mean	Stddev
branch1	0.095	76.03	0.199	0.020
branch2	0.068	22.58	0.130	0.010
branch3	0.098	69.73	0.195	0.024
<i>H. sapiens</i>	0.068	15.23	0.132	0.010
<i>M. Musculus</i>	0.056	12.5	0.105	0.008
<i>C. familiaris</i>	0.061	21.68	0.128	0.009
<i>M. domestica</i>	0.065	19.96	0.125	0.008
<i>O. anatinus</i>	0.065	13.92	0.132	0.010
<i>G. gallus</i>	0.062	10.86	0.123	0.008

Table A15. Alignment of a region of the *PTPN4* gene showing a three-amino-acid deletion believed to have occurred after marsupials and eutherians diverged from monotremes.

SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>sapiens</i> (human)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>troglodytes</i> (chimp)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>gorilla</i> (gorilla)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>pygmaeus</i> (orang_sumatran)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>leucogenys</i> (gibbon)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>mulatta</i> (rhesus)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>jacchus</i> (marmoset_nwm)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>syrichta</i> (tarsier)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>murinus</i> (mouse_lemur)
SPSKPLVRKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>belangeri</i> (tree_shrew)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>musculus</i> (mouse)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>norvegicus</i> (rat)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>porcellus</i> (guinea_pig)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>tridecemlineatus</i> (ground_squirrel)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>ordii</i> (kangaroo_rat)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>cuniculus</i> (rabbit)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>princeps</i> (pika)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>familiaris</i> (dog)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>catus</i> (cat)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>caballus</i> (horse)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>lucifugus</i> (microbat)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>vampyrus</i> (macrobat)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>taurus</i> (cow)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>aries</i> (sheep)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>truncatus</i> (dolphin)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>scrofa</i> (pig)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>novemcinctus</i> (armadillo)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>hoffmanni</i> (sloth)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>capensis</i> (hyrax)
SPSKPLARKLM---DWEIVSRISLSDDRLETQSLPSRSPPGTPN	<i>telfairi</i> (tenrec)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>domestica</i> (opossum)
SPSKPLARKLM---DWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>eugenii</i> (wallaby)
SPSKPLARKLMMSGMDWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>anatinus</i> (platypus)
SPSKPLARKLMMSGMDWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>gallus</i> (chicken)
SPSKPLARKLMMSGMDWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>platyrhynchos</i> (duck)
SPSKPLARKVMMSGMDWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>guttata</i> (finch)
SPSKPLARKLISGMDWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>carolinensis</i> (lizard)
SPSKPLSRKLVSEMDWEVVSRLNSISDDRLETQSLPSRSPPGTPN	<i>tropicalis</i> (frog)
SPSKPLARKLLGGTDWESVSRNSISDDRLETQSLPTRSPPGTPN	<i>rubripes</i> (fugu)
SPSKPLARKLLGGTDWESVSRNSISDDRLETQSLPTRSPPGTPN	<i>nigroviridis</i> (pufferfish)
SPSKPLARKLMGGTDWESVSRNSISDDRLETQSLPTRSPPGTPS	<i>latipes</i> (medaka)
SPSKPLARKLMGVTDWEAVSRNSISDDRLETQSLPTRSPPGTPN	<i>rerio</i> (zebrafish)
SPSKPLARKLVGGTDWESVSRNSISDDRLETQSLPTRSPPGTPN	<i>aculeatus</i> (stickleback)