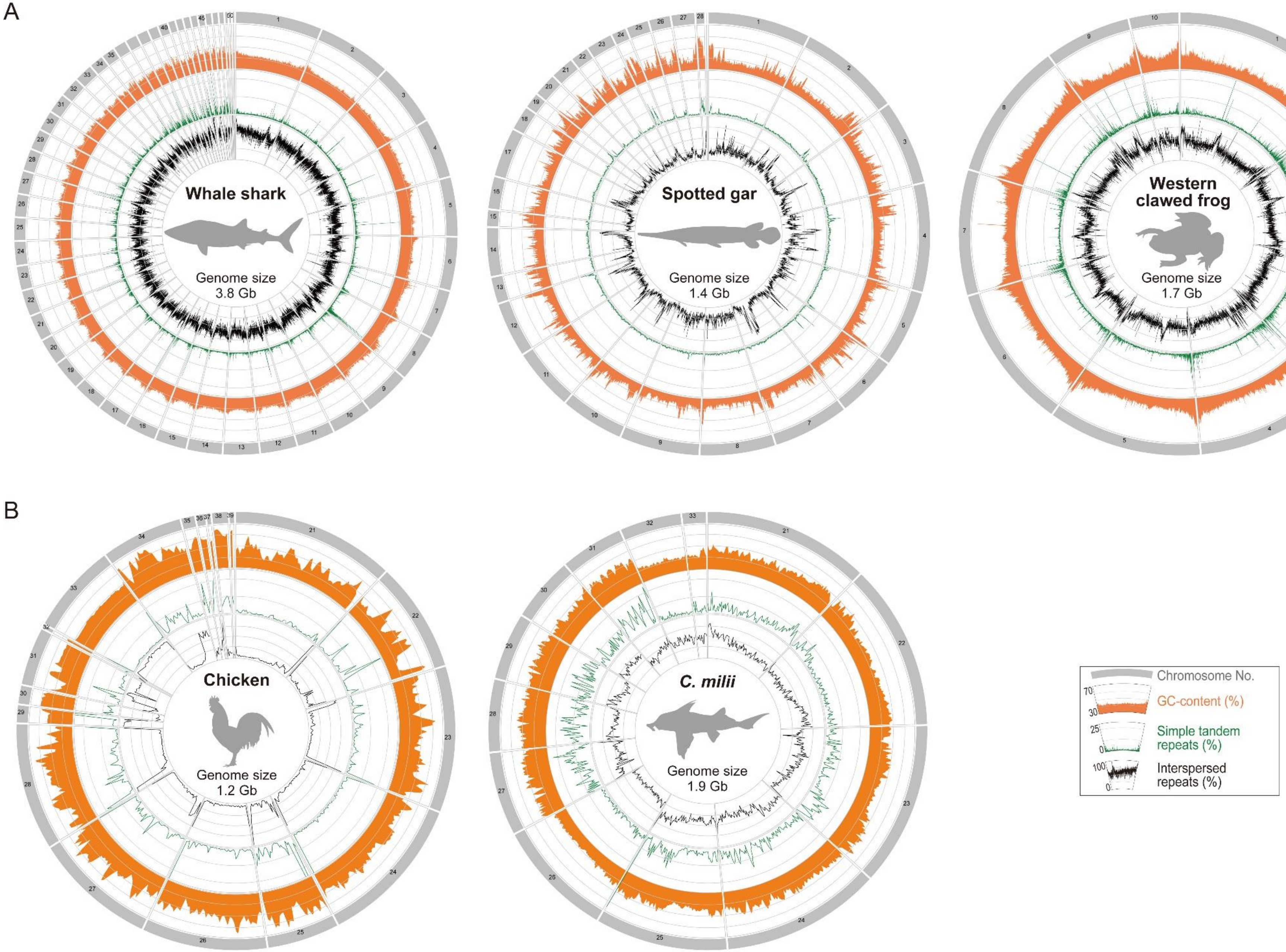
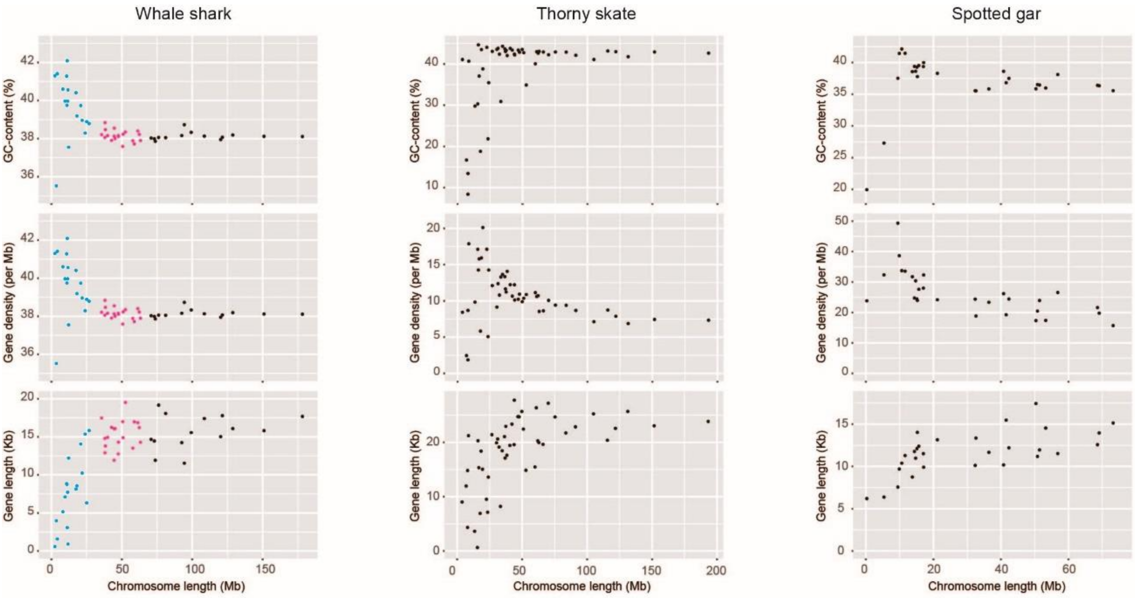
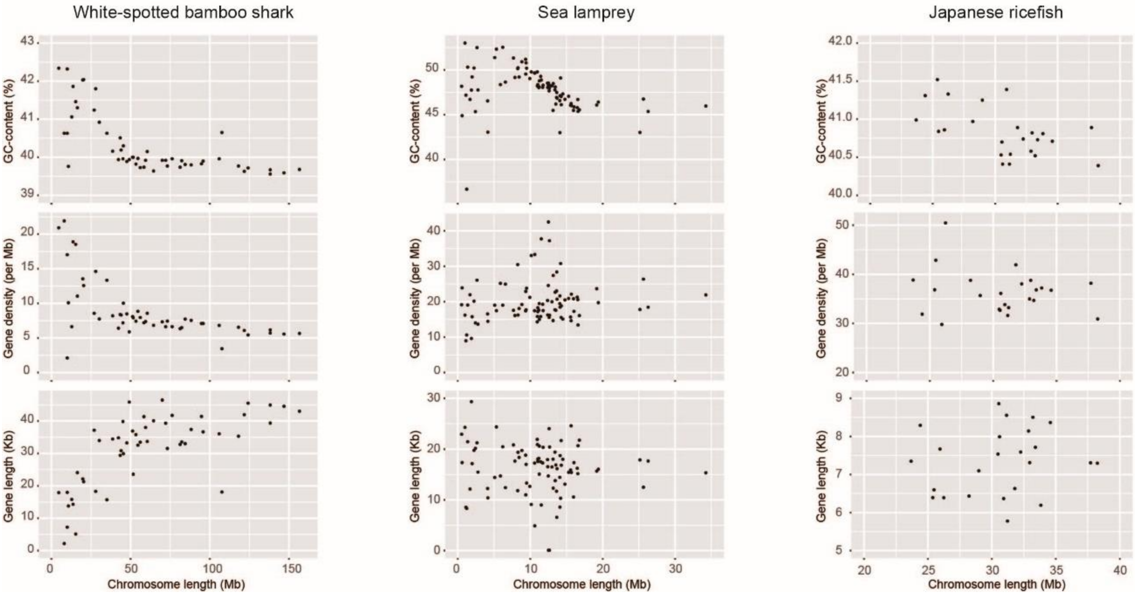
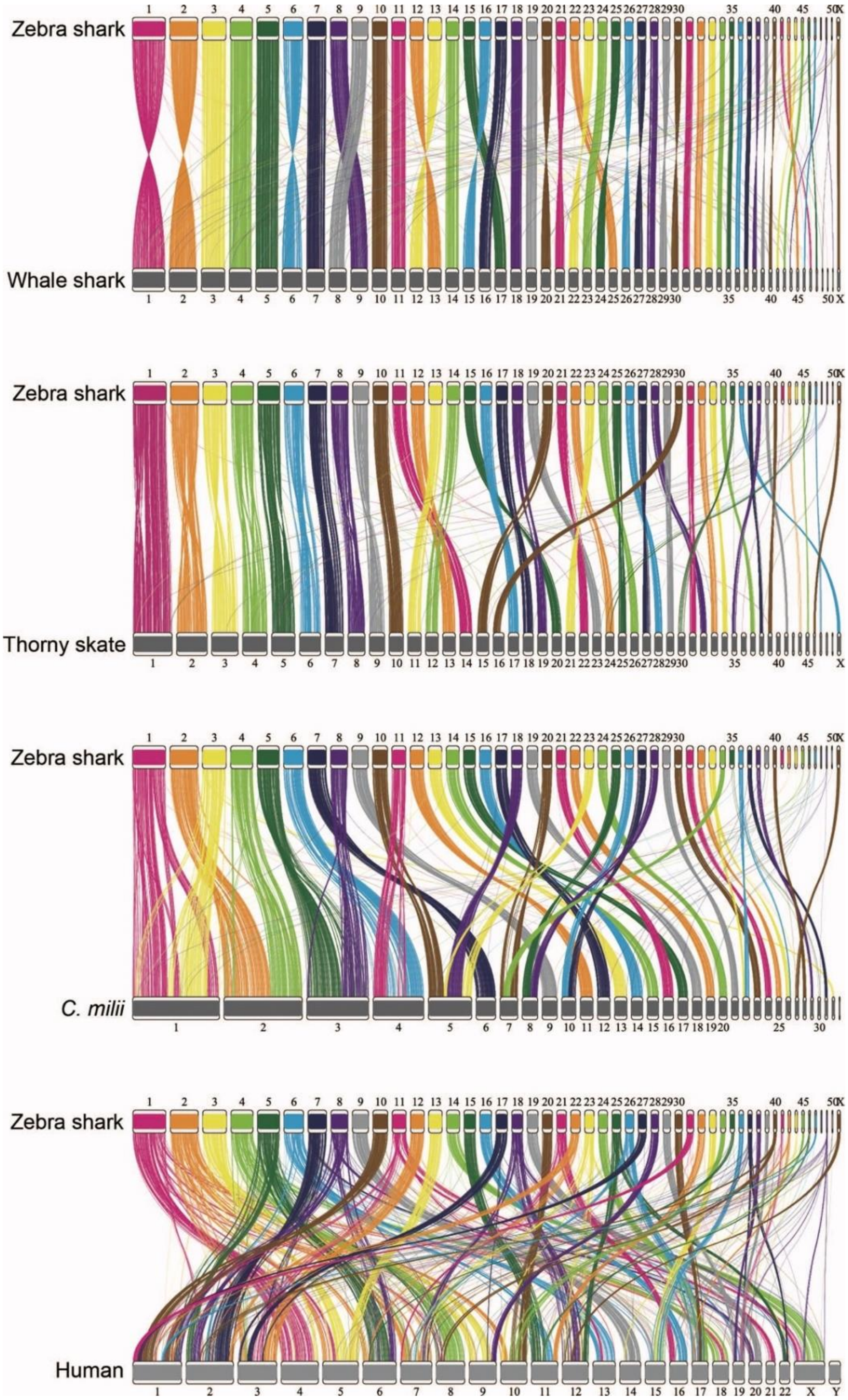

A
Whale shark
Genome size
3.8 Gb
Spotted gar
Genome size
1.4 Gb
Western clawed frog
Genome size
1.7 Gb
B
Chicken
Genome size
1.2 Gb
C. milii
Genome size
1.9 Gb
Chromosome No.
70
30
GC-content (%)
25
0
Simple tandem repeats (%)
100
0
Interspersed repeats (%)

Supplementary Figure 1. Chromosomal sequence characterization for selected vertebrate species. (*A*) Whole genome views. (*B*) Views focusing on relatively small chromosomes of the zebra shark and chicken. See Figure 2A for the whole genome views. The orange areas show GC-content (30–70%), while the green and black lines show content of simple tandem repeats (0–25%) and interspersed repeats (0–100%), respectively, in 100-kb-long non-overlapping windows (also see Figure 2A).

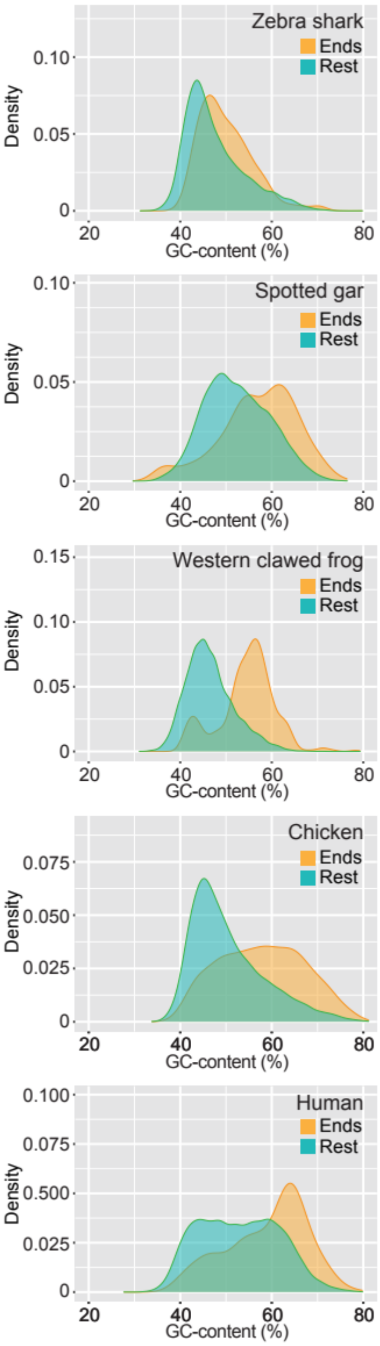

Whale shark
Thorny skate
Spotted gar
GC-content (%)
Gene density (per Mb)
Gene length (Kb)
Chromosome length (Mb)
Chromosome length (Mb)
Chromosome length (Mb)


White-spotted bamboo shark
Sea lamprey
Japanese ricefish
GC-content (%)
Gene density (per Mb)
Gene length (Kb)
Chromosome length (Mb)
Chromosome length (Mb)
Chromosome length (Mb)

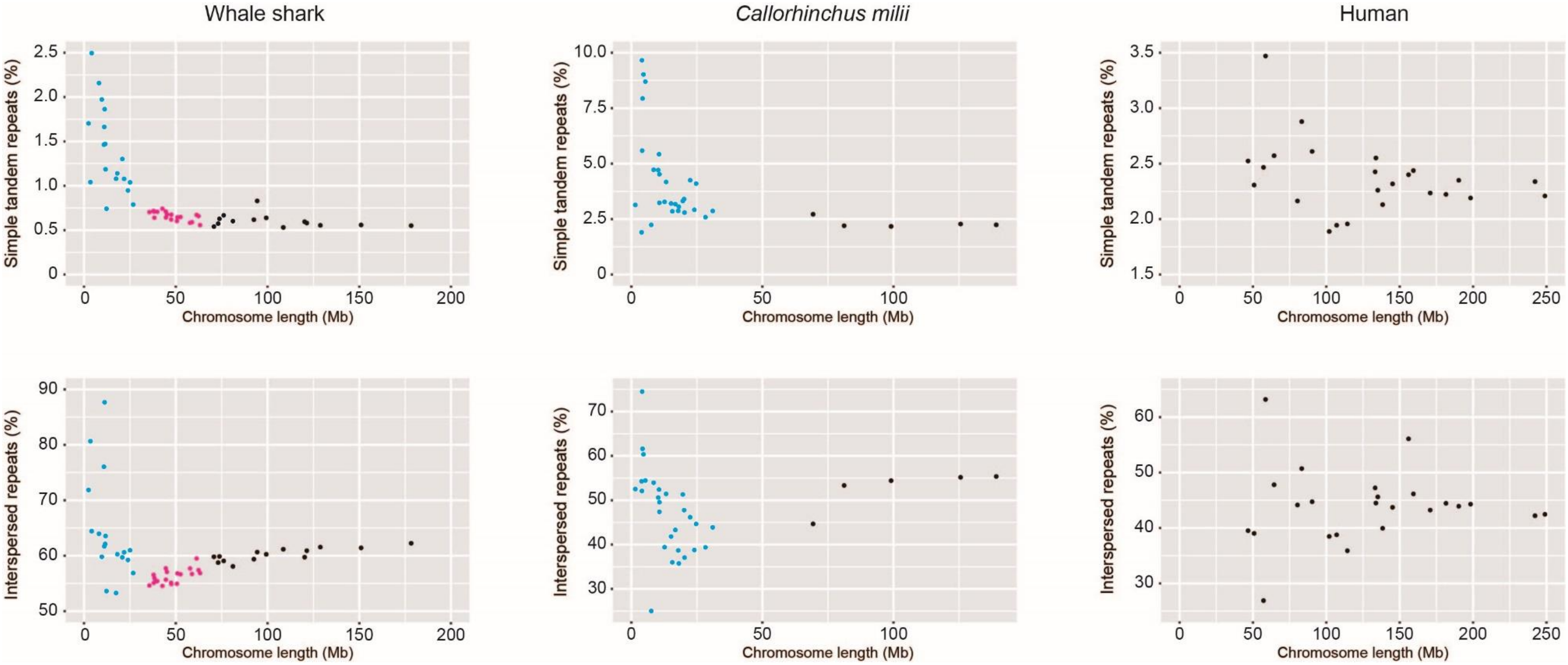
Supplementary Figure 2. Two-dimensional plot of GC-content and median values of gene length, gene density, and the number of synonymous substitutions per synonymous site (K_S) for protein-coding genes on individual chromosomes. See Methods for statistical tests for correlation of these features with chromosome length.



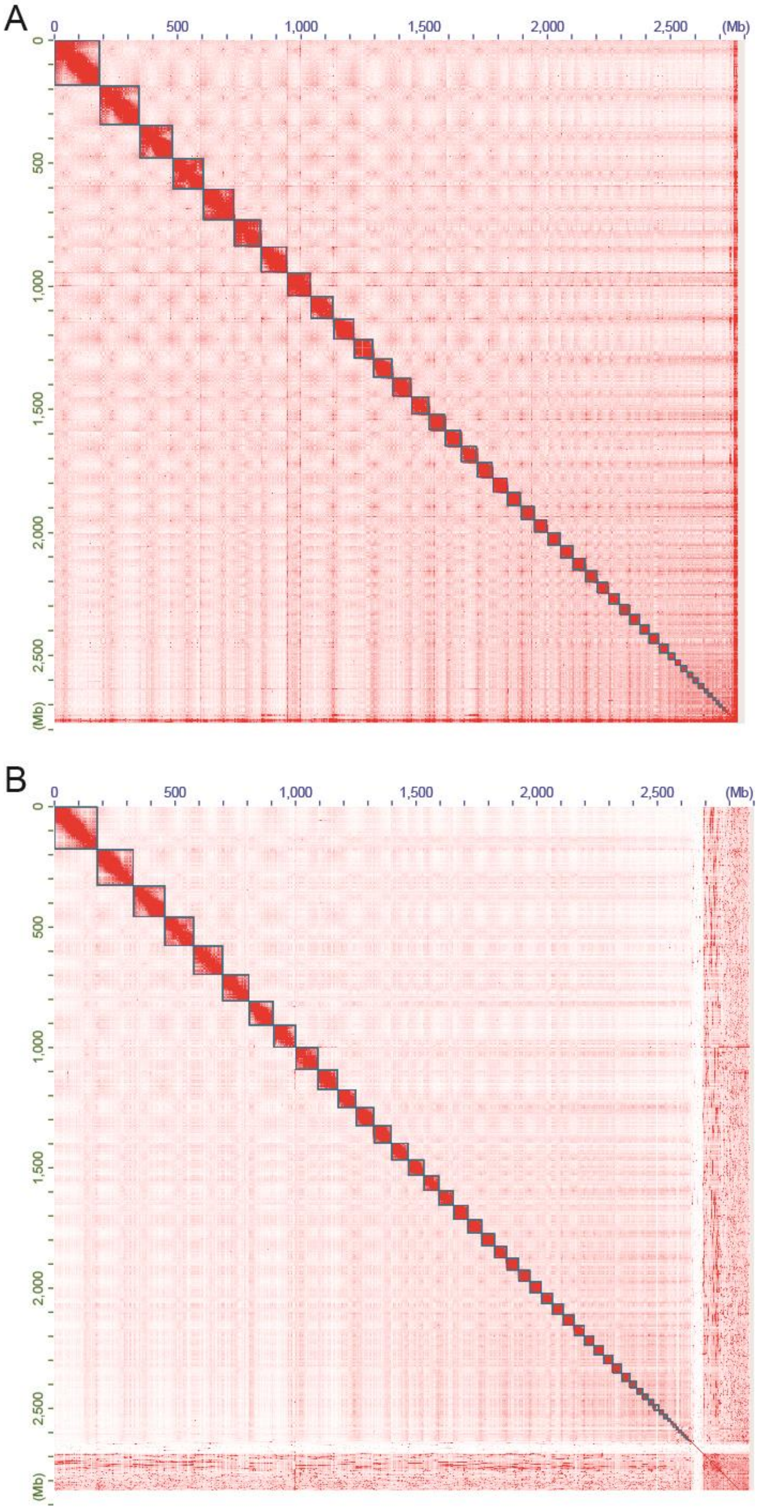
Supplementary Figure 3. Chromosomal homology between the zebra shark and other vertebrates. Conserved synteny is visualized with the inter-specific correspondence of one-to-one orthologs (see Methods).



Supplementary Figure 4. Intrachromosomal heterogeneity of GC-content in protein-coding regions of individual predicted genes. GC-content of protein-coding genes is compared between 1 Mb ends of relatively large chromosomes and their remainders for different size categories of chromosomes.



Supplementary Figure 5. Two-dimensional plots of chromosome lengths and the coverage of interspersed repeats and simple tandem repeats. See Methods and Supplementary Table 3 for more details about statistical tests.



Supplementary Figure 6. Chromatin contact heatmap based on Hi-C data. (*A*) Zebra shark. (*B*) Whale shark. Color intensities correspond to the frequency of chromatin contacts detected by Hi-C data (see Methods for details).

Supplementary Table 1. Comparative statistics of the newly obtained and previously released elasmobranch genome assemblies.

Species		**Zebra shark**		White-spotted bamboo shark		Brownbanded bamboo shark		White shark		Thorny skate	
Assembly Name		**sSteFas1.1**		ASM401019v2 (GCF_004010195.1)		Cpunctatum_v2.1		GCA_003604245.1		sAmbRad1.1.pri (GCF_010909765.2)	
Type		Genome	Gene model	Genome	Gene model	Genome	Gene model	Genome	Gene model	Genome	Gene model
Selected program		BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5
Selected reference gene set		Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata
# of core genes detected	**Complete**	3039 (90.61%)	3012 (89.80%)	3012 (89.80%)	3169 (94.48%)	2933 (87.45%)	2947 (87.87%)	3075 (91.68%)	1560 (46.51%)	3052 (91.00%)	3195 (95.26%)
	Complete + Fragmented	3204 (95.53%)	3206 (95.59%)	3187 (95.02%)	3263 (97.29%)	3155 (94.07%)	3155 (94.07%)	3214 (95.83%)	2243 (66.88%)	3162 (94.28%)	3236 (96.48%)
# of missing core genes		150 (4.47%)	148 (4.41%)	167 (4.98%)	91 (2.71%)	199 (5.93%)	199 (5.93%)	140 (4.17%)	1111 (33.12%)	192 (5.72%)	118 (3.52%)
Average # of orthologs per core genes		1.03	1.09	1.03	2.37	1.02	1.04	1.04	1.04	1.03	2.15
% of detected core genes that have more than 1 ortholog		2.34	7.9	1.93	52.57	1.91	2.75	2.57	3.27	2.23	49.39
# of sequences		1266	33222	101977	47685	125458	33501	9222	24520	957	38895
Total length (nt/aa)		2771866316	13169207	3556477427	31931996	3441189827	12025038	3915280839	8641475	2558784953	28873284
Longest sequence (nt/aa)		187018787	30503	156605281	38279	56086953	8400	23593172	11289	193168304	35076
Shortest sequence (nt/aa)		5000	5	1000	31	500	30	10005	51	1363	31
Mean sequence length (nt/aa)		2189468	396	34875	670	27429	359	424559	352	2673756	742
Median sequence length (nt/aa)		6877	251	1621	482	936	214	51597	243	51440	507
N50 sequence length (nt)		76588082		73167256		9192170		2991981		62052929	
L50 sequence count		13		17		90		330		13	
# of sequences > 1Kb		1266 (100.0%)		101512 (99.5%)		58399 (46.5%)		9222 (100.0%)		957 (100.0%)	
# of sequences > 10Kb		337 (26.6%)		10379 (10.2%)		3085 (2.5%)		9222 (100.0%)		945 (98.7%)	
# of sequences > 100Kb		59 (4.7%)		100 (0.1%)		1253 (1.0%)		3312 (35.9%)		245 (25.6%)	
# of sequences > 1Mb		50 (3.9%)		52 (0.1%)		495 (0.4%)		888 (9.6%)		64 (6.7%)	
# of sequences > 10Mb		45 (3.6%)		47 (0.0%)		82 (0.1%)		34 (0.4%)		45 (4.7%)	
Sum length of sequences > 1Mb (bp)		2756734042 (99.5%)		3104319730 (87.3%)		2965606038 (86.2%)		2920922205 (74.6%)		2482724239 (97.0%)	
Sum length of sequences > 10Mb (bp)		2723511891 (98.3%)		3069888471 (86.3%)		1650831574 (48.0%)		460085648 (11.8%)		2407838606 (94.1%)	
GC-content (%)		41.78		42.06		41.79		43.87		44.3	
# of gaps (>=5 Ns)		86090		155015		443025		173480		3749	

Completeness of gene models were assessed using peptide sequences of predicted genes.

Supplementary Table 2. Compatative statistics of the newly obtained and previously released whale shark genome assemblies.

Assembly		**sRhiTyp1.1 (this study)**		Rtypus_kobe_v1.0		RhiTyp_1.0_HiC (DNA ZOO)	GCA_001642345.3 ASM164234v3 (Tan et al., 2021)	GCF_001642345.1 ASM164234v2		GCA_013626285.1 RhiTyp_1.0 (Weber et al., 2020)
Type		Genome	Gene model	Genome	Gene model	Genome	Genome	Genome	Gene model	Genome
Selected program		BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5	BUSCO_v5
Selected reference gene set		Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata	Vertebrata
# of core genes detected	**Complete**	2565 (76.48%)	2609 (77.79%)	2605 (77.67%)	2663 (79.40%)	3088 (92.07%)	2779 (82.86%)	2630 (78.41%)	2758 (82.23%)	3069 (91.50%)
	Complete + Partial	3080 (91.83%)	3084 (91.95%)	3068 (91.47%)	3071 (91.56%)	3219 (95.97%)	3125 (93.17%)	3103 (92.52%)	3176 (94.69%)	3217 (95.92%)
# of missing core genes		274 (8.17%)	270 (8.05%)	286 (8.53%)	283 (8.44%)	135 (4.03%)	229 (6.83%)	251 (7.48%)	178 (5.31%)	137 (4.08%)
Average # of orthologs per		1.03	1.1	1.03	1.03	1.03	1.03	1.03	1.45	1.03
% of detected core genes that		2.57	8.85	1.92	2.67	2.62	2.12	2.24	26.14	2.61
# of sequences		16776	36824	155537	27005	133476	39171	57334	27896	136451
Total length (nt/aa)		2881325413	12601623	2659892559	9576545	2822097320	2961781270	2931599576	13150867	2820592179
Longest sequence (nt/aa)		178270547	34351	3410558	22026	182888402	2759877	1482643	23396	16092075
Shortest sequence (nt/aa)		5000	5	500	3	200	666	1000	31	200
Mean sequence length (nt/aa)		171750	342	17101	355	21143	75612	51132	471	20671
Median sequence length		8723	211	1809	229	385	12614	15389	354	401
N50 sequence length (nt)		70795558		325616		70985326	344655	144422		3127311
L50 sequence count		14		2117		14	2425	5485		238
# of sequences > 1K (nt)		16776 (100.0%)		118303 (76.1%)		28321 (21.2%)	39167 (100.0%)	57333 (100.0%)		31297 (22.9%)
# of sequences > 10K (nt)		6359 (37.9%)		13782 (8.9%)		1105 (0.8%)	24329 (62.1%)	39558 (69.0%)		4101 (3.0%)
# of sequences > 100K (nt)		241 (1.4%)		6132 (3.9%)		84 (0.1%)	7248 (18.5%)	8565 (14.9%)		1823 (1.3%)
# of sequences > 1M (nt)		51 (0.3%)		221 (0.1%)		51 (0.0%)	214 (0.5%)	10 (0.0%)		690 (0.5%)
# of sequences > 10M (nt)		46 (0.3%)		0 (0.0%)		47 (0.0%)	0 (0.0%)	0 (0.0%)		24 (0.0%)
Sum length of sequences > 1M (nt)		2639639612 (91.6%)		306535345 (11.5%)		2691409604 (95.4%)	274947285 (9.3%)	11332151 (0.4%)		2226036735 (78.9%)
Sum length of sequences > 10M (nt)		2611710806 (90.6%)		0 (0.0%)		2674204998 (94.8%)	0 (0.0%)	0 (0.0%)		295806019 (10.5%)
GC-content (%)		41.51		40.83		41.74	41.77	41.78		41.74
# of gaps (>=5 N's)		95171		180448		59795	15895	10		56773

Completeness of gene models were assessed using peptide sequences of predicted genes.

Supplementary Table 3. Statistical tests for correlation with chromosome length.

A. Zebra shark

	GC-content	Gene length	Gene density	Ks	Simple tandem repeats	Interspersed repeats
Correlation coefficient	-0.5716	0.7572	-0.5777	-0.3979	-0.6138	0.4840
Interpretation	Moderate Negative	Strong Positive	Moderate Negative	Weak Negative	Moderate Negative	Moderate Positive
Logarithmic approximation	y = -0.45ln(x) + 41.122	y = 10.626ln(x) - 10.991	y = -8.42ln(x) + 45.982	y = -0.027ln(x) + 0.1891	y = -0.78ln(x) + 3.9934	y = 1.5193ln(x) + 52.433
Coefficient of determination, R^2	0.3527	0.8378	0.5674	0.3281	0.7442	0.1799
p-value	1.455E-05	1.970E-10	1.120E-05	4.216E-03	2.134E-06	3.697E-04
Significant difference	**	**	**	**	**	**

B. *Callorhinchus milii*

	GC-content	Gene length	Gene density	Ks	Simple tandem repeats	Interspersed repeats
Correlation coefficient	-0.4354	0.7554	-0.5548	-0.4487	-0.4308	0.1205
Interpretation	Moderate Negative	Strong Positive	Moderate Negative	Moderate Negative	Moderate Negative	No
Logarithmic approximation	y = -1.431ln(x) + 47.824	y = 2.4301ln(x) + 4.2049	y = -6.936ln(x) + 47.866	y = -0.086ln(x) + 0.5379	y = -1.101ln(x) + 6.9893	y = -1.712ln(x) + 52.778
Coefficient of determination, R^2	0.3161	0.733	0.4115	0.4667	0.3202	0.0378
p-value	1.132E-02	3.760E-07	8.069E-04	9.997E-03	1.233E-02	5.041E-01
Significance of difference	*	**	**	**	*	n.s.

C. Chicken

	GC contents	Gene length	Gene density	Ks	Simple tandem repeats	Interspersed repeats
Correlation coefficient	-0.5373	0.6696	-0.5099	-0.2691	-0.2832	-0.1998
Interpretation	Moderate Negative	Moderate Positive	Moderate Negative	Weak Negative	Weak Negative	No
Logarithmic approximation	y = -2.947ln(x) + 53.947	y = 1.3267ln(x) + 5.0556	y = -16.06ln(x) + 85.579	y = -0.085ln(x) + 0.4454	y = -1.126ln(x) + 5.6236	y = -6.328ln(x) + 32.726
Coefficient of determination, R^2	0.6117	0.5632	0.6847	0.2626	0.4856	0.2329
p-value	2.919E-04	1.687E-06	6.609E-04	1.125E-01	7.282E-02	2.103E-01
Significance of difference	**	**	**	n.s	n.s	n.s.

D. Human

	GC contents	Gene length	Gene density	Ks	Simple tandem repeats	Interspersed repeats
Correlation coefficient	0.1676	0.4945	-0.2824	-0.5869	-0.3439	0.0465
Interpretation	No	Moderate Positive	Weak Negative	Moderate Negative	Weak Negative	No
Logarithmic approximation	y = 2.1872ln(x) + 28.093	y = 4.0045ln(x) - 4.3211	y = -3.654ln(x) + 30.186	y = -0.063ln(x) + 0.4389	y = -0.265ln(x) + 3.6278	y = 0.9987ln(x) + 39.123
Coefficient of determination, R^2	0.0389	0.2628	0.0957	0.4461	0.1577	0.0051
p-value	4.336E-01	1.404E-02	1.812E-01	2.570E-03	9.988E-02	8.290E-01
Significance of difference	n.s	*	n.s	**	n.s	n.s.

E. Whale shark

	GC-content	Gene length	Gene density	Ks	Simple tandem repeats	Interspersed repeats
Correlation coefficient	-0.4612	0.6327	-0.5988	-0.4118	-0.6325	-0.1986
Interpretation	Moderate Negative	Moderate Positive	Moderate Negative	Moderate Negative	Moderate Negative	No
Logarithmic approximation	y = -0.694ln(x) + 41.165	y = 6.7389ln(x) - 2.4441	y = -5.853ln(x) + 35.247	y = -0.029ln(x) + 0.1931	y = -0.389ln(x) + 2.2918	y = -3.121ln(x) + 71.304
Coefficient of determination, R^2	0.3289	0.7284	0.5623	0.3722	0.6757	0.2329
p-value	6.586E-04	6.335E-07	3.454E-06	2.677E-03	6.389E-07	1.624E-01
Significance of difference	**	**	**	**	**	n.s.

The tests A-D were performed for Figures 2B, 4D and Supplementary Figure 5, while those in E-P were performed for Supplementary Figure 2. Correlation coefficients in the range between ±0.0 and ±0.2 indicate no correlation, ±0.2 to ±0.4 indicate weak positive/negative correlation, ±0.4 to ±0.7 indicate moderate positive/negative correlation, and ±0.7 to ±1.0 indicate strong positive/negative correlation. The results are showed in the “Interpretation” columns. R^2 values approaching 1 indicate high linearity. Significance of difference was represented as an asterisk in the correlation coefficients: *, p-value < 0.05; **, p-value < 0.01; n.s., not significant.

F. White-spotted bamboo shark

	GC-content	Gene length	Gene density
Correlation coefficient	-0.6753	0.6873	-0.6057
Interpretation	Moderate Negative	Moderate Positive	Moderate Negative
Logarithmic approximation	y = -0.719ln(x) + 43.077	y = 10.576ln(x) - 9.4156	y = -3.56ln(x) + 22.489
Coefficient of determination, R^2	0.5941	0.6354	0.5125
p-value	5.463E-08	2.544E-08	2.487E-06
Significance of difference	**	**	**

G. Thorny skate

	GC-content	Gene length	Gene density
Correlation coefficient	0.3841	0.5412	-0.3419
Interpretation	Weak Positive	Moderate Positive	Weak Negative
Logarithmic approximation	y = 5.7015ln(x) + 18.312	y = 5.1051ln(x) + 0.0208	y = -0.763ln(x) + 13.479
Coefficient of determination, R^2	0.3098	0.4309	0.0315
p-value	6.443E-03	5.939E-05	1.619E-02
Significance of difference	**	**	*

H. Sea lamprey

	GC-content	Gene length	Gene density
Correlation coefficient	-0.3119	-0.1110	0.1680
Interpretation	Weak Negative	No	No
Logarithmic approximation	y = -0.335ln(x) + 48.568	y = -0.907ln(x) + 18.077	y = 1.4659ln(x) + 16.893
Coefficient of determination, R^2	0.0144	0.0236	0.0455
p-value	3.658E-03	3.120E-01	1.243E-01
Significance of difference	**	n.s.	n.s.

I. Spotted gar

	GC-content	Gene length	Gene density
Correlation coefficient	0.0287	0.6232	-0.6635
Interpretation	No	Moderate Positive	Moderate Negative
Logarithmic approximation	y = 1.9528ln(x) + 30.955	y = 1.6303ln(x) + 6.6291	y = -2.977ln(x) + 35.273
Coefficient of determination, R^2	0.2534	0.4869	0.2076
p-value	3.658E-03	3.120E-01	1.243E-01
Significance of difference	n.s.	**	**

J. Medaka

	GC-content	Gene length	Gene density
Correlation coefficient	-0.5622	0.1606	-0.2008
Interpretation	Moderate Negative	No	Weak Negative
Logarithmic approximation	y = -1.422ln(x) + 45.69	y = 1.0579ln(x) + 3.7387	y = -6.895ln(x) + 59.937
Coefficient of determination, R^2	0.327	0.0259	0.0406
p-value	4.241E-03	4.534E-01	3.469E-01
Significance of difference	**	n.s.	n.s.

K. Stickleback

	GC-content	Gene length	Gene density
Correlation coefficient	-0.3402	0.0144	-0.1206
Interpretation	Weak Negative	No	No
Logarithmic approximation	y = -0.746ln(x) + 46.546	y = 0.2458ln(x) + 4.9873	y = -6.06ln(x) + 75.383
Coefficient of determination, R^2	0.1542	0.0061	0.0169
p-value	1.214E-01	9.493E-01	5.929E-01
Significance of difference	n.s.	n.s.	n.s.

L. Western clawed frog

	GC-content	Gene length	Gene density
Correlation coefficient	-0.9340	0.4344	-0.6244
Interpretation	Strong Negative	Moderate Positive	Moderate Negative
Logarithmic approximation	y = -2.788ln(x) + 54.646	y = 1.9828ln(x) + 1.2458	y = -7.254ln(x) + 55.945
Coefficient of determination, R^2	0.9482	0.128	0.3865
p-value	7.650E-05	2.096E-01	5.366E-02
Significance of difference	**	n.s.	n.s.

M. Two-lined caecilian

	GC-content	Gene length	Gene density
Correlation coefficient	-0.2417	0.5683	-0.5862
Interpretation	Weak Negative	Moderate Positive	Moderate Negative
Logarithmic approximation	y = -0.108ln(x) + 44.776	y = 12.809ln(x) - 26.843	y = -4.064ln(x) + 29.007
Coefficient of determination, R^2	0.0334	0.5694	0.4971
p-value	3.187E-01	1.113E-02	8.351E-03
Significance of difference	n.s.	*	**

N. Indian cobra

	GC-content	Gene length	Gene density
Correlation coefficient	-0.1715	0.2039	-0.5474
Interpretation	No	Weak Positive	Moderate Negative
Logarithmic approximation	y = -0.437ln(x) + 40.582	y = 0.313ln(x) + 10.862	y = -4.768ln(x) + 36.031
Coefficient of determination, R^2	0.0252	0.0282	0.5702
p-value	4.826E-01	4.024E-01	1.528E-02
Significance of difference	n.s.	n.s.	*

O. Goode's thornscrub tortoise

	GC-content	Gene length	Gene density
Correlation coefficient	-0.3186	0.5463	-0.6520
Interpretation	Weak Negative	Moderate Positive	Moderate Negative
Logarithmic approximation	y = -1.473ln(x) + 50.446	y = 3.7851ln(x) + 4.1224	y = -4.696ln(x) + 32.402
Coefficient of determination, R^2	0.1653	0.4812	0.6983
p-value	1.292E-01	5.752E-03	5.557E-04
Significance of difference	n.s.	**	**

P. Common wall lizard

	GC-content	Gene length	Gene density
Correlation coefficient	-0.7173	0.3581	-0.4530
Interpretation	Strong Negative	Weak Positive	Moderate Negative
Logarithmic approximation	y = -2.256ln(x) + 54.074	y = 2.427ln(x) + 5.3606	y = -7.508ln(x) + 50.022
Coefficient of determination, R^2	0.6844	0.1728	0.2899
p-value	5.467E-04	1.322E-01	5.143E-02
Significance of difference	**	n.s.	n.s.

Supplementary Table 4. Statistical tests for the inter- and intra-chromosomal comparisons of sequence trends.

A. Zebra shark

	GC-content	Gene length	Gene density	K_s	Simple tandem repeats	Interspersed repeats
p -value (Kruskal-Wallis test)	0.0000	2.277E-174	1.043E-05	1.210E-39	6.474E-07	1.665E-05
Significance of difference	**	**	**	**	**	**
eMAC-End vs eMAC-Rest *p* -value (Mann-Whitney U test)	8.490E-17	1.549E-01	1.750E-01	6.814E-02	2.594E-04	1.019E-03
Significance of difference	**	n.s.	n.s.	n.s.	**	**
eMAC-End vs eMAC-Rest rank-biserial correlation (Mann-Whitney U test)	0.272	-0.046	0.306	0.151	0.816	0.735
Effect	Small	N/A	Medium	Small	Large	Large
eMAC-End vs eMIC *p* -value (Mann-Whitney U test)	1.111E-16	1.295E-09	1.750E-04	7.503E-02	6.547E-03	4.729E-05
Significance of difference	**	**	**	n.s.	*	**
eMAC-End vs eMIC rank-biserial correlation (Mann-Whitney U test)	-0.275	0.201	-0.798	-0.150	-0.580	0.866
Effect	Small	Small	Large	Small	Large	Large
eMAC-Rest vs eMIC *p* -value (Mann-Whitney U test)	0	8.648E-176	3.353E-05	1.811E-40	2.554E-06	1.464E-02
Significance of difference	**	**	**	**	**	*
eMAC-Rest vs eMIC rank-biserial correlation (Mann-Whitney U test)	-0.458	0.249	-0.882	-0.264	-1.000	0.521
Effect	Medium	Small	Large	Small	Large	Large

B. Chicken

	GC-content	Gene length	Gene density	K_s	Simple tandem repeats	Interspersed repeats
p -value (Kruskal-Wallis test)	0.0000	2.520E-269	1.292E-05	7.531E-133	8.663E-05	1.651E-03
Significance of difference	**	**	**	**	**	**
MAC-End vs MAC-Rest *p* -value (Mann-Whitney *U* test)	2.547E-109	2.828E-45	1.827E-04	3.617E-37	1.827E-04	5.828E-04
Significance of difference	**	**	**	**	**	**
MAC-End vs MAC-Rest rank-biserial correlation (Mann-Whitney *U* test)	0.455	-0.289	1.000	0.390	1.000	0.920
Effect	Medium	Small	Large	Medium	Large	Large
MAC-End vs MIC *p* -value (Mann-Whitney *U* test)	1.588E-02	3.303E-01	4.344E-01	3.005E-01	1.343E-02	5.642E-03
Significance of difference	*	n.s.	n.s.	n.s.	*	*
MAC-End vs MIC rank-biserial correlation (Mann-Whitney *U* test)	0.050	0.020	0.181	0.035	0.562	0.629
Effect	N/A	N/A	Small	N/A	Large	Large
MAC-Rest vs MIC *p* -value (Mann-Whitney *U* test)	0	1.770E-258	1.006E-05	4.761E-109	2.516E-03	3.005E-01
Significance of difference	**	**	**	**	**	n.s.
MAC-Rest vs MIC rank-biserial correlation (Mann-Whitney *U* test)	-0.436	0.324	-1.000	-0.362	-0.686	-0.238
Effect	Medium	Medium	Large	Medium	Large	Small

The tests were performed for Figure 4B and 4D. Significance of difference was shown with the number of asterisks: *, *p* -value < 0.05/number of tests; **, *p* -value < 0.01/number of tests; n.s., not significant. Effect size in the range between ±0.0 and ±0.1 indicate no effect, ±0.1 to ±0.3 indicate small effect, ±0.3 to ±0.5 indicate medium effect, and ±0.5 to ±1.0 indicate large effect. N/A, not applicable.

Supplementary Table 5. Statistical tests for intragenomic difference of sequence properties in zebra shark.

	K_s	K_a	K_a/K_s	GC-content
p -value (Kruskal-Wallis test)	3.590E-51	1.174E-03	2.273E-06	0.000E+00
Significance of difference	**	**	**	**
eMAC vs eMID *p* -value (Mann-Whitney U test)	7.946E-01	1.330E-02	6.624E-03	1.179E-50
Significance of difference	n.s.	n.s.	*	**
eMAC vs eMID rank-biserial correlation (Mann-Whitney U test)	0.003	0.031	0.034	-0.036
Effect	N/A	N/A	N/A	N/A
eMAC vs eMIC *p* -value (Mann-Whitney U test)	2.195E-47	1.600E-02	4.097E-07	0.000E+00
Significance of difference	**	n.s.	**	**
eMAC vs eMIC rank-biserial correlation (Mann-Whitney U test)	-0.294	-0.049	0.103	-0.824
Effect	Small	N/A	Small	Large
eMAC vs X *p* -value (Mann-Whitney U test)	2.698E-03	8.250E-01	3.319E-01	6.464E-79
Significance of difference	*	n.s.	n.s.	**
eMAC vs X rank-biserial correlation (Mann-Whitney U test)	0.221	0.016	-0.072	-0.988
Effect	Small	N/A	N/A	Large
eMID vs eMIC *p* -value (Mann-Whitney U test)	2.805E-46	1.930E-04	8.651E-04	0.000E+00
Significance of difference	**	**	**	**
eMID vs eMIC rank-biserial correlation (Mann-Whitney U test)	-0.296	-0.077	0.069	-0.883
Effect	Small	N/A	N/A	Large
eMID vs X *p* -value (Mann-Whitney U test)	3.319E-03	8.842E-01	1.902E-01	3.792E-61
Significance of difference	*	n.s.	n.s.	**
eMID vs X rank-biserial correlation (Mann-Whitney U test)	0.217	-0.011	-0.097	-0.992
Effect	Small	N/A	N/A	Large
eMIC vs X *p* -value (Mann-Whitney U test)	3.392E-09	3.692E-01	5.455E-02	1.689E-02
Significance of difference	**	n.s.	n.s.	n.s.
eMIC vs X rank-biserial correlation (Mann-Whitney U test)	0.447	0.068	-0.145	-0.997
Effect	Medium	N/A	Small	Large
eMAC vs PAR *p* -value (Mann-Whitney U test)				3.10533E-24
Significance of difference				**
eMAC vs PAR rank-biserial correlation (Mann-Whitney U test)				-0.99951
Effect				Large
eMID vs PAR *p* -value (Mann-Whitney U test)				1.814E-23
Significance of difference				**
eMID vs PAR rank-biserial correlation (Mann-Whitney U test)				-1.000
Effect				Large
eMIC vs PAR *p* -value (Mann-Whitney U test)				8.158E-16
Significance of difference				**
eMIC vs PAR rank-biserial correlation (Mann-Whitney U test)				-1.000
Effect				Large
X vs PAR *p* -value (Mann-Whitney U test)				5.022E-17
Significance of difference				**
X vs PAR rank-biserial correlation (Mann-Whitney U test)				-1.000
Effect				Large

These tests were performed for Figure 5D. Significance of difference was shown with the number of asterisks: *, *p* -value < 0.05/number of tests; **, *p* -value < 0.01/number of tests; n.s., not significant. Effect size in the range between ±0.0 and ±0.1 indicate no effect, ±0.1 to ±0.3 indicate small effect, ±0.3 to ±0.5 indicate medium effect, and ±0.5 to ±1.0 indicate large effect. N/A, not applicable.

Supplementary Table 6A. Paired-end libraries prepared for genome sequencing.

Species	Library ID	Starting total DNA amount	DNA shearing condition						# of PCR cycles	AMPure XP concentration for size selection	Peak insert length
		(ng)	Machine*	Peak power	Duty factor	Cycle/Burst	Duration (second)	Times repeated	(cycle)		(bp)
Rhincodon typus	P465_01_1**										523
Rhincodon typus	P438_01_1	100	E220	175	10	200	60	2	4	2.0x	280
Rhincodon typus	P438_02_1	100	E220	175	10	200	60	2	4	2.0x	273
Stegostoma tigrinum	P167_01_3	50.00	S220	140	10	200	55	2	3	1.0x, 0.6x	228
Stegostoma tigrinum	P167_12_3	238.00	S220	105	5	200	50	3	0	1.0x, 0.65x	446
Stegostoma tigrinum	P377_02_2	50.00	E220	140	10	200	180	2	5	1.0x, 0.5x	400
Stegostoma tigrinum	P167_11_4	343.80	S220	105	5	200	50	3	0	0.9x, 0.65x	689
Stegostoma tigrinum	P167_01_5	50.00	S220	105	5	200	50	3	3	0.65x, 1.0x	617
Stegostoma tigrinum	P167_11_5	89.64	S220	105	5	200	50	3	0	0.65x	823
Stegostoma tigrinum	P167_01_4	29.16	S220	140	10	200	55	2	5	0.6x, 0.2x	366
Stegostoma tigrinum	P377_02_3***	50.00	E220	140	10	200	180	2	6	1.0x, 0.5x	436
Stegostoma tigrinum	P167_11_1	381.60	S220	105	5	200	50	3	0	0.65x, 0.2x	341
Stegostoma tigrinum	P167_01_6	50.00	S220	105	5	200	50	3	4	0.6x, 0.3x	316
Stegostoma tigrinum	P167_11_2	90.72	S220	105	5	200	50	3	2	0.55x, 0.2x	348
Stegostoma tigrinum	P377_02_1****	500.00	E220	140	10	200	180	2	6	1.0x, 0.5x	496
Stegostoma tigrinum	P489_01_1	50.00	E220	175	5	200	55	2	5	1.0x, 0.5x	510

*S220: S220 Focused-ultrasonicator, Covaris; E220: E220 Focused-ultrasonicator, Covaris.

**This library was prepared using Chromium Gnome Reagent Kits and the preparation procedure followed the official manual provided by the manufacturer.

***This library was prepared uisng QIAseq FX DNA Library Kits (QIAGEN, 180473).

****This library was prepared with heating 500 ng of DNA for five minutes at 90℃ followed by size fractionation with 0.5 - 1.0 concentration of AMPure XP beads.

Supplementary Table 6B. Mate-pair libraries prepared for zebra shark genome sequencing.

Species	Library ID	Targeted mate distance range (Kbp)	Starting total DNA amount	DNA shearing condition						# of PCR cycles	Peak insert length
			(μg)	machine*	Peak power	Duty factor	Cycle/Burst	Duration (second)	Times repeated	(cycles)	(bp)
Stegostoma tigrinum	P380_01_1	7-10	4	S220	240	20	200	40	1	8	536
Stegostoma tigrinum	P380_01_2	12-18	4	S220	240	20	200	40	1	13	681

Supplementary Table 6C. RNA-seq libraries prepared for transcriptome data acquisition.

Species	Library ID	Sex	Stage	Tissue	Starting total RNA amount (ng)	Peak insert length (bp)	RNA fragment condition	# of PCR cycles	AMPure XP concentration for size selection	
									After end repair	After PCR
Rhincodon typus	P381_01_1	female	juvenile	Gill	1000	323	94°C, 4 min	9	1.8x	1.0x
Rhincodon typus	P381_02_1	female	juvenile	Eye	1000	312	94°C, 4 min	9	1.8x	1.0x
Rhincodon typus	P381_03_1	female	juvenile	Stomach	1000	325	94°C, 4 min	9	1.8x	1.0x
Rhincodon typus	P381_04_1	female	juvenile	Intestine	1000	306	94°C, 4 min	10	1.8x	1.0x
Rhincodon typus	P381_05_1	female	juvenile	Pituitary	1000	271	94°C, 4 min	10	1.8x	1.0x
Rhincodon typus	P381_06_1	female	juvenile	Liver	1000	351	94°C, 4 min	10	1.8x	1.0x
Rhincodon typus	P381_07_1	female	juvenile	Kidney	1000	278	94°C, 4 min	10	1.8x	1.0x
Rhincodon typus	P381_09_1	female	juvenile	Spleen	1000	248	94°C, 4 min	13	1.8x	1.0x
Rhincodon typus	P381_10_1	female	juvenile	Intestine_upper	1000	284	94°C, 4 min	10	1.8x	1.0x
Rhincodon typus	P381_11_1	female	juvenile	Pancreas	1000	256	94°C, 4 min	13	1.8x	1.0x
Rhincodon typus	P433_01_1	female	juvenile	Follicle	1000	256	94°C, 4 min	4	1.8x	1.0x
Rhincodon typus	P433_02_1	female	juvenile	Follicle-large	1000	272	94°C, 4 min	6	1.8x	1.0x
Rhincodon typus	P433_03_1	female	juvenile	Uterus_anterior	1000	256	94°C, 4 min	7	1.8x	1.0x
Rhincodon typus	P433_04_1	female	juvenile	Uterus_middle	1000	236	94°C, 4 min	9	1.8x	1.0x
Rhincodon typus	P433_05_1	female	juvenile	Uterus_posterior	1000	238	94°C, 4 min	9	1.8x	1.0x
Stegostoma tigrinum	P430_01_1	female	juvenile	Blood	1000	256	94°C, 4 min	8	1.8x	1.0x
Stegostoma tigrinum	P430_02_1	female	juvenile	Eye	1000	279	94°C, 4 min	6	1.8x	1.0x
Stegostoma tigrinum	P430_03_1	female	juvenile	Olfactory bulb	1000	277	94°C, 4 min	6	1.8x	1.0x
Stegostoma tigrinum	P430_04_1	female	juvenile	Muscle	1000	253	94°C, 4 min	8	1.8x	1.0x
Stegostoma tigrinum	P430_05_1	female	juvenile	Cerebrum	1000	271	94°C, 4 min	7	1.8x	1.0x
Stegostoma tigrinum	P430_06_1	female	juvenile	Cerebellum	1000	238	94°C, 4 min	9	1.8x	1.0x
Stegostoma tigrinum	P430_07_1	female	juvenile	Diencephalon	1000	260	94°C, 4 min	7	1.8x	1.0x
Stegostoma tigrinum	P430_08_1	female	juvenile	Mesencephalon	1000	267	94°C, 4 min	9	1.8x	1.0x
Stegostoma tigrinum	P430_09_1	female	juvenile	Medulla oblong	1000	281	94°C, 4 min	7	1.8x	1.0x
Stegostoma tigrinum	P430_10_1	female	juvenile	Gill	1000	244	94°C, 4 min	9	1.8x	1.0x
Stegostoma tigrinum	P430_11_1	female	juvenile	Heart	1000	246	94°C, 4 min	8	1.8x	1.0x
Stegostoma tigrinum	P430_12_1	female	juvenile	Spleen	1000	233	94°C, 4 min	11	1.8x	1.0x
Stegostoma tigrinum	P430_13_1	female	juvenile	Liver	1000	249	94°C, 4 min	8	1.8x	1.0x
Stegostoma tigrinum	P430_14_1	female	juvenile	Gallbladder	1000	251	94°C, 4 min	8	1.8x	1.0x
Stegostoma tigrinum	P430_15_1	female	juvenile	Stomach	1000	242	94°C, 4 min	8	1.8x	1.0x
Stegostoma tigrinum	P430_16_1	female	juvenile	Intestine	1000	248	94°C, 4 min	8	1.8x	1.0x
Stegostoma tigrinum	P430_17_1	female	juvenile	Colon	1000	267	94°C, 4 min	8	1.8x	1.0x
Stegostoma tigrinum	P430_18_1	female	juvenile	Kidney	1000	285	94°C, 4 min	6	1.8x	1.0x
Stegostoma tigrinum	P430_19_1	female	juvenile	Rectal gland	1000	260	94°C, 4 min	7	1.8x	1.0x
Stegostoma tigrinum	P430_20_1	female	juvenile	Ovary	1000	245	94°C, 4 min	7	1.8x	1.0x
Stegostoma tigrinum	P430_21_1	female	juvenile	Fallopian tube	1000	272	94°C, 4 min	6	1.8x	1.0x
Stegostoma tigrinum	P430_22_1	female	juvenile	Olfactory epithelium	1000	259	94°C, 4 min	7	1.8x	1.0x
Stegostoma tigrinum	P183_06_1	Unknown	Embryo	Embryo	500	195	94°C, 8 min	8	1.8x	1.0x
Stegostoma tigrinum	P183_07_1	Unknown	Embryo	Embryo	500	174	94°C, 8 min	9	1.8x	1.0x

Supplementary Table 7. Primer sequences used in genomic qPCR.

Gene	Chromosomal location	Oligonucleotide primer sequence (5′ to 3′)	
		Forward	Reverse
HoxA1	Autosome (Scaffold 2)	TTTGCAGAAAGCAGAGCTGA	GCCCAGTGCAATGGTAATTT
FoxG1	Autosome (Scaffold 9)	GCTTTGCAACGGGATATGTT	CGAGGCAACTGGATATGGAT
Gdf11	X (Scaffold 41)	GCCAAACAAAAGGCTTTCAG	CGGCTAAAGGACGAATGGTA
Gli1	X (Scaffold 41)	GACTGGAATGCCCAAACAGT	AATCTCAGGGGATCTCAGCA
Stat3	X (Scaffold 41)	TCGGAAATCCTTCGCTTCTA	TTCCAAATGCCTCATCCTTC
Dip2c	PAR (Scaffold 41)	GCTTAGGACGTGCTGTCTCC	AAGGTGTTGACAGGGTGGAC
Aladin	PAR (Scaffold 41)	CTCACATCCCCTTCTGCTTC	TCAGAGTTCGGAGAAGAGGTC
Creb1	PAR (Scaffold 41)	CAGTGCCGGAAGAAGAAGAA	CTCCTCGATCAGGGTCTTGT