
Supplementary Materials

Supplementary Tables

Supplementary Table S1: Summary of Hi-C sequencing data

Pair-end libraries	Insert size	Total data (G)	Read length (bp)	Q20 (%)	GC Content (%)	Sequence coverage (X)
Hi-C	350 bp	142.1	150	96.86	40.64	184.2

Supplementary Table S2: Final assembly statistics

	Scaffold		Contig	
	length(bp)	number	length(bp)	number
max_len	41,394,826	-	7,330,617	-
N10	35,960,341	2	4,076,292	14
N20	33,308,695	4	2,371,079	39
N30	31,281,249	7	1,710,000	74
N40	29,057,639	9	1,226,355	122
N50	27,816,280	11	878,406	191
N60	26,641,119	14	611,000	287
N70	24,171,097	17	350,058	443
N80	21,668,425	20	175,384	734
N90	14,785,410	24	60,477	1,422
Total_length	704,759,797		704,554,997	
number>=100bp	2,253		4,301	
number>=2,000bp	2,083		4,130	
=====				
GC_rate	0.396		0.396	
=====				
Total N bases: 204,800 ## Min N: 100 ## Max N: 100				

Supplementary Table S3: Catfish species used in this study

Species	Abbreviation
<i>Glyptosternom maculatum</i>	GM
<i>Exostoma labiatum</i>	EL
<i>Pareuchiloglanis feae</i>	PF
<i>Pareuchiloglanis kamengensis</i>	PK
<i>Glyptothorax quadriocellatus</i>	GQ
<i>Glyptothorax fukiensishonghensis</i>	GFH
<i>Glyptothorax interspinalum</i>	GI
<i>Glyptothorax cavia</i>	GC
<i>Glyptothorax laosensis</i>	GRT
<i>Glyptothorax zanaensis</i>	GZ
<i>Glyptothorax trilineatus</i>	GT
<i>Glyptothorax minimaculatus</i>	GMN
<i>Glyptothorax dorsalis</i>	GR
<i>Bagarius yarrelli</i>	BY
<i>Glyptothorax trilineatus</i>	GT
<i>Pseudecheneis sulcatus</i>	PS
<i>Leiocassis longirostris</i>	LL

Supplementary Table S4: Transcriptome sequencing and assembly for catfish species

Species	Tissue	Raw Reads	Clean Bases	Q20(%)	GC Content(%)	Unigene	Unigene (after filtering)*
<i>Glyptothorax trilineatus</i>	Liver	58,703,438	8.65G	97.32	44.99	53,166	7,493
<i>Pseudecheneis sulcatus</i>	Liver	52,732,074	7.63G	97.32	47.1	92,588	7,570
<i>Glyptothorax interspinalum</i>	Liver	49,713,012	7.2G	97.01	43.67	58,429	7,548
<i>Bagarius yarrelli</i>	Liver	46,516,790	6.51G	97.46	45.66	84,869	7,475
<i>Glyptothorax quadriocellatus</i>	Liver	55,338,112	8.04G	97.22	46.31	81,125	7,261
<i>Glyptothorax dorsalis</i>	Liver	45,723,950	6.65G	97	46.58	78,601	7,609
<i>Glyptothorax minimaculatus</i>	Muscle	51,562,436	7.45G	97.27	46.75	88,120	8,069
<i>Glyptothorax andersonii</i>	Muscle	53,253,548	7.78G	97.19	49.45	103,813	12,397
<i>Glyptothorax zanaensis</i>	Liver	44,037,460	6.41G	97.37	46.71	786,85	7,779
<i>Pareuchilog lanisfeae</i>	Liver	52,110,776	7.57G	96.77	47.38	146,849	8,773
<i>Exostoma labiatum</i>	Liver	48,712,936	7.04G	96.74	46.47	126,332	7,646
<i>Glyptothorax cavia</i>	Liver	43,842,934	6.34G	97.26	46.2	81,706	7,579
<i>Leiocassis longirostris</i>	Liver	45,610,394	6.64G	97.09	45.68	90,047	7,128
<i>Glyptothorax fukiensis honghensis</i>	Liver	52,346,474	7.62G	97.23	47.02	92,402	7,545
<i>Glyptosternom maculatum</i>	Liver	70,137,398	9.94G	95.11	46.2	118,505	17,618
<i>Pareuchiloglanis kamengensis</i>	Liver	50,645,710	7.33G	96.63	46.46	143,738	7,745
<i>Glyptothorax laosensis</i>	Muscle	59,004,652	8.63G	97.54	46.7	93,648	8,341

*Unigenes coding less than 50 amino acids were filtered out

Supplementary Table S5: Repeat content of in the genome for teleosts studied in this work

Species	content	DNA	LINE	SINE	LTR	Satellite	simple repeat	Total
<i>Astyanax mexicanus</i>	Length (bp)	1.5E+08	32195870	3771555	17988414	2684338	6789186	471880955
	% in genome	11.26488	2.411244	0.282463	1.347205	0.201038	0.508462	35.340556
<i>Cyprinus carpio</i>	Length (bp)	82902727	67595528	1542679	45256254	13811534	697930	259889515
	% in genome	4.837764	3.944517	0.090023	2.640915	0.805968	0.040727	15.165775
<i>Danio rerio</i>	Length (bp)	5.55E+08	105520422	12957877	277274398	22350368	11196029	907069760
	% in genome	33.07764	6.283957	0.771668	16.512257	1.33101	0.666746	54.017859
<i>Glyptosternum maculatum</i>	Length (bp)	28527736	175495547	5838349	66048667	2567022	4780706	248718345
	% in genome	3.756988	23.112055	0.768887	8.698343	0.338066	0.6296	32.755202
<i>Ictalurus punctatus</i>	Length (bp)	1.17E+08	29049039	14927122	25276956	1092575	946823	273503948
	% in genome	14.95077	3.708665	1.905733	3.227087	0.139488	0.12088	34.91801
<i>Oncorhynchus mykiss</i>	Length (bp)	3.53E+08	260406674	7619050	44460673	1161011	2721805	803162604
	% in genome	16.21284	11.950744	0.349658	2.040417	0.053282	0.124911	36.859236
<i>Oryzias latipes</i>	Length (bp)	79691181	64634399	8098210	18771797	3556777	148050	259149339
	% in genome	10.85627	8.805092	1.103213	2.557267	0.484537	0.020169	35.303704
<i>Pelteobagrus fulvidraco</i>	Length (bp)	1.15E+08	101671793	7367991	52409970	5432365	31757189	281933392
	% in genome	15.73402	13.874124	1.005435	7.15186	0.7413	4.333583	38.472607

Supplementary Table S6: Repeat contents of LINE categories in teleost studied in this work

Specie	Gmac	Ipun	Drer	Amex	Olat	Pful	Amol	Omyk	Ccar	Pnat
LINE/RTE-X	5969	856	274	569	453	10	54	42980	520	2279
LINE/Jockey	6	0	4	277	832	0	241	926	3	34
LINE/RTE-BovB	449773	3781	30374	640	18331	56850	3482	517	264	53950
LINE/L1-Tx1	3477	5893	5088	760	4594	6428	98	13949	681	3582
LINE/L1	204	4814	10005	5522	21211	5184	1504	11221	1289	11141
LINE/LINE	0	0	0	0	75	0	0	300	2	0
LINE/L2	251749	43750	212407	54372	36023	236956	28034	190971	38329	93025
LINE/CR1-Zenon	0	1	0	0	1331	21	0	176	0	0
LINE/R1	7	0	0	550	553	1	38	301	5	29
LINE/LOA	0	0	0	0	9	0	0	4	0	0
LINE/I-Nimb	1	353	4431	244	0	1531	456	1883	55	0
LINE/I	732	1473	21520	2859	895	311	1357	1608	507	194
LINE/Ambal	0	0	0	0	3	0	0	11	0	0
LINE/Proto2	0	0	0	0	866	0	265	4	23	0
LINE/Dong-R4	1229	0	2	0	9337	1	1799	3	0	0
LINE/R2-NeSL	0	0	36	0	1	0	0	0	0	0
LINE/Penelope	251	495	199	155	5700	693	445	4011	81	2599
LINE/Rex-Babar	32739	13517	56782	26763	19125	64839	9699	220302	12450	55885
LINE/Proto1	0	0	0	0	29	0	0	0	3	0
LINE/R2	3	18	52	0	371	22	0	675	24	0
LINE/L2-Hydra	0	0	0	0	1	0	0	0	0	0
LINE/CRE	0	0	0	0	3	0	0	0	0	0
LINE	63597	936	2436	142	183	38307	729	5414	169	78

LINE/R2-Hero	24	325	59	1200	351	33	0	497	97	416
LINE/RTE	0	669	6761	101	4416	16751	862	8	89	9294
LINE/DRE	0	0	0	0	25	0	0	15	0	0
LINE/RTE-RTE	0	27	0	0	43	0	0	0	3	0
LINE/Tad1	0	0	0	0	331	0	470	30	7	0
LINE/CR1	8	1	0	616	602	32	1	461	0	0

Supplementary Table S7: GO enrichment of expanded gene families

GO_accession	Description	Term_type	pValue	FDR
GO:0006334	nucleosome assembly	biological_process	1.69E-26	9.03E-24
GO:0031497	chromatin assembly	biological_process	1.69E-26	9.03E-24
GO:0034728	nucleosome organization	biological_process	1.69E-26	9.03E-24
GO:0006323	DNA packaging	biological_process	1.82E-26	9.03E-24
GO:0071103	DNA conformation change	biological_process	1.23E-19	2.64E-17
GO:0009409	response to cold	biological_process	3.16E-11	3.09E-09
GO:0048871	multicellular organismal homeostasis	biological_process	3.16E-11	3.09E-09
GO:0050826	response to freezing	biological_process	3.16E-11	3.09E-09
GO:0009266	response to temperature stimulus	biological_process	3.26E-11	3.12E-09
GO:0051704	multi-organism process	biological_process	9.97E-10	9.33E-08
GO:0008150	biological_process	biological_process	1.17E-09	1.07E-07
GO:0009628	response to abiotic stimulus	biological_process	3.14E-08	2.57E-06
GO:0004329	formate-tetrahydrofolate ligase activity	molecular_function	1.29E-07	9.68E-06
GO:0022402	cell cycle process	biological_process	2.32E-06	0.00016315
GO:0008509	anion transmembrane transporter activity	molecular_function	5.70E-06	0.0003766
GO:0002437	inflammatory response to antigenic stimulus	biological_process	8.38E-06	0.00049623

GO:0002438	acute inflammatory response to antigenic stimulus	biological_process	8.38E-06	0.00049623
GO:0002524	hypersensitivity	biological_process	8.38E-06	0.00049623
GO:0016064	immunoglobulin mediated immune response	biological_process	8.38E-06	0.00049623
GO:0016068	type I hypersensitivity	biological_process	8.38E-06	0.00049623
GO:0019724	B cell mediated immunity	biological_process	8.38E-06	0.00049623
GO:0004252	serine-type endopeptidase activity	molecular_function	8.39E-06	0.00049623
GO:0006950	response to stress	biological_process	8.77E-06	0.00051196
GO:0002250	adaptive immune response	biological_process	9.71E-06	0.00053848
GO:0002443	leukocyte mediated immunity	biological_process	9.71E-06	0.00053848
GO:0002449	lymphocyte mediated immunity	biological_process	9.71E-06	0.00053848
GO:0002460	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	biological_process	9.71E-06	0.00053848
GO:0002526	acute inflammatory response	biological_process	1.10E-05	0.00060523
GO:0002252	immune effector process	biological_process	1.38E-05	0.00074734
GO:0006760	folic acid-containing compound metabolic process	biological_process	5.83E-05	0.0030813
GO:0009396	folic acid-containing compound biosynthetic process	biological_process	5.83E-05	0.0030813
GO:0000226	microtubule cytoskeleton organization	biological_process	7.48E-05	0.0039097
GO:0051649	establishment of localization in cell	biological_process	7.91E-05	0.0040852
GO:0006461	protein complex assembly	biological_process	9.68E-05	0.0048558
GO:0070271	protein complex biogenesis	biological_process	9.68E-05	0.0048558
GO:0042559	pteridine-containing compound biosynthetic process	biological_process	9.73E-05	0.0048558
GO:0042558	pteridine-containing compound metabolic process	biological_process	9.97E-05	0.004923
GO:0007049	cell cycle	biological_process	0.000124	0.0060698
GO:0051641	cellular localization	biological_process	0.00013	0.0062648
GO:0016779	nucleotidyltransferase activity	molecular_function	0.000159	0.0076138

GO:0005272	sodium channel activity	molecular_function	0.000169	0.0079964
GO:0006820	anion transport	biological_process	0.000185	0.0086508
GO:0032991	macromolecular complex	cellular_component	0.000204	0.009432
GO:0004000	adenosine deaminase activity	molecular_function	0.00027	0.01226
GO:0008236	serine-type peptidase activity	molecular_function	0.000273	0.01226
GO:0017171	serine hydrolase activity	molecular_function	0.000273	0.01226
GO:0005622	intracellular	cellular_component	0.00032	0.014254
GO:0004175	endopeptidase activity	molecular_function	0.000326	0.014366
GO:0016459	myosin complex	cellular_component	0.000525	0.022905
GO:0016772	transferase activity, transferring phosphorus-containing groups	molecular_function	0.000576	0.024898
GO:0015031	protein transport	biological_process	0.000595	0.025232
GO:0045184	establishment of protein localization	biological_process	0.000595	0.025232
GO:0034061	DNA polymerase activity	molecular_function	0.000667	0.02802
GO:0042398	cellular modified amino acid biosynthetic process	biological_process	0.000794	0.033026
GO:0065008	regulation of biological quality	biological_process	0.000808	0.033299
GO:0008104	protein localization	biological_process	0.000958	0.039135
GO:0006955	immune response	biological_process	0.000977	0.03953
GO:0044430	cytoskeletal part	cellular_component	0.001008	0.040432
GO:0035556	intracellular signal transduction	biological_process	0.001095	0.043532
GO:0043227	membrane-bounded organelle	cellular_component	0.001129	0.044105
GO:0043231	intracellular membrane-bounded organelle	cellular_component	0.001129	0.044105
GO:0007017	microtubule-based process	biological_process	0.001274	0.049345

Supplementary Table S8: KEGG pathway enrichment of expanded gene families

Pathway	Pathway ID	Input number	Background number	pValue	FDR
One carbon pool by folate	dre00670	5	19	2.63E-06	8.43E-05
Tight junction	dre04530	9	174	6.12E-05	0.000978425
ErbB signaling pathway	dre04012	7	107	0.000104858	0.00111849
Endocytosis	dre04144	10	286	0.000517205	0.004137643
Calcium signaling pathway	dre04020	8	263	0.004325479	0.027683065
mTOR signaling pathway	dre04150	4	74	0.006467962	0.0344958
VEGF signaling pathway	dre04370	4	86	0.010623867	0.04856625
Phosphatidylinositol signaling system	dre04070	4	90	0.01231871	0.04927484

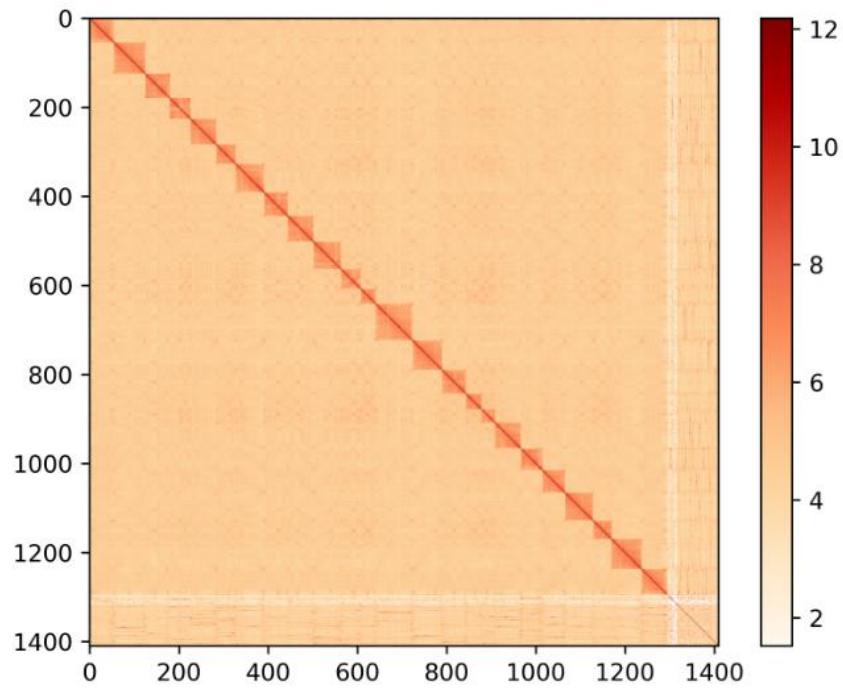
Supplementary Table S9: The whole-genome resequencing data for population genetics analysis for *G. maculatum*.

Hom and Het represented the homozygous and heterozygous locus, respectively.

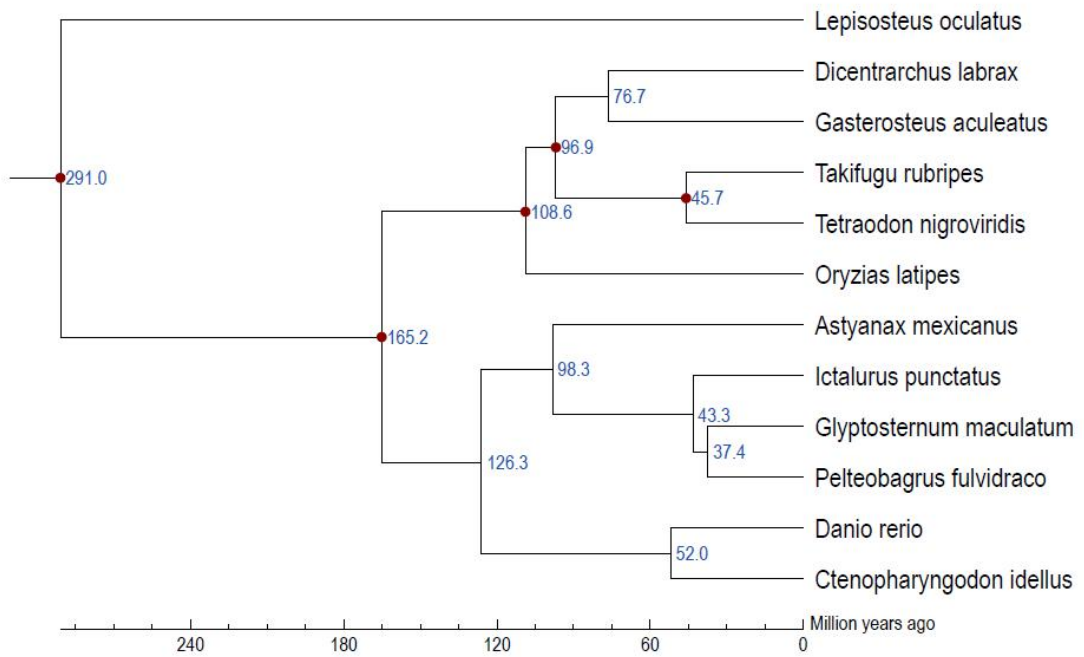
group	Sample	BioSample	SRR	Hom	Het	Total
GM4300	DES02225	SAMN12748442	SRR10119705	71,255	178,695	249,950
GM4300	DES02226	SAMN12748443	SRR10119704	73,551	174,949	248,500
GM4300	DES02227	SAMN12748444	SRR10119703	71,190	178,000	249,190
GM4300	DES02228	SAMN12748445	SRR10119702	74,070	179,038	253,108
GM4300	DES02229	SAMN12748446	SRR10119701	71,170	177,872	249,042
GM4300	DES02230	SAMN12748447	SRR10119700	73,880	175,473	249,353
GM4300	DES02231	SAMN12748448	SRR10119699	73,352	176,126	249,478
GM4300	DES02232	SAMN12748449	SRR10119698	71,129	177,430	248,559
GM4300	DES02233	SAMN12748450	SRR10119697	74,157	176,860	251,017
GM4300	DES02234	SAMN12748451	SRR10119695	74,060	176,376	250,436
GM4500	DSW40269	SAMN12748434	SRR10119714	71,316	176,811	248,127
GM4500	DSW40270	SAMN12748435	SRR10119713	63,473	193,992	257,465
GM4500	DSW40271	SAMN12748436	SRR10119712	74,289	176,718	251,007
GM4500	DSW40272	SAMN12748437	SRR10119711	71,727	177,896	249,623
GM4500	DSW40273	SAMN12748438	SRR10119710	71,912	179,563	251,475
GM4500	DSW40274	SAMN12748439	SRR10119709	69,260	181,003	250,263
GM4500	DSW40275	SAMN12748440	SRR10119708	67,628	183,486	251,114
GM4500	DSW40276	SAMN12748432	SRR10119716	73,141	178,011	251,152
GM4500	DSW40277	SAMN12748433	SRR10119715	64,348	190,190	254,538
GM4500	DSW40278	SAMN12748441	SRR10119706	70,031	179,381	249,412
GM2900	DSW49759	SAMN12748423	SRR10119707	74,380	177,348	251,728
GM2900	DSW49760	SAMN12748424	SRR10119696	73,490	178,704	252,194
GM2900	DSW49761	SAMN12748422	SRR10119718	75,038	176,227	251,265
GM2900	DSW49762	SAMN12748425	SRR10119693	74,336	178,330	252,666
GM2900	DSW49763	SAMN12748426	SRR10119692	75,921	175,159	251,080
GM2900	DSW49764	SAMN12748427	SRR10119691	75,497	177,443	252,940

GM2900	DSW49765	SAMN12748428	SRR10119690	74,979	176,768	251,747
GM2900	DSW49766	SAMN12748429	SRR10119689	75,279	176,260	251,539
GM2900	DSW49767	SAMN12748430	SRR10119688	74,334	178,836	253,170
GM2900	DSW49768	SAMN12748431	SRR10119717	76,746	173,860	250,606
outgroup	JI08262401	SAMN12748452	SRR10119694	449,450	739	450,189

Supplementary Figures



Supplementary Figure S1: Contact matrix generated from Hi-C data.

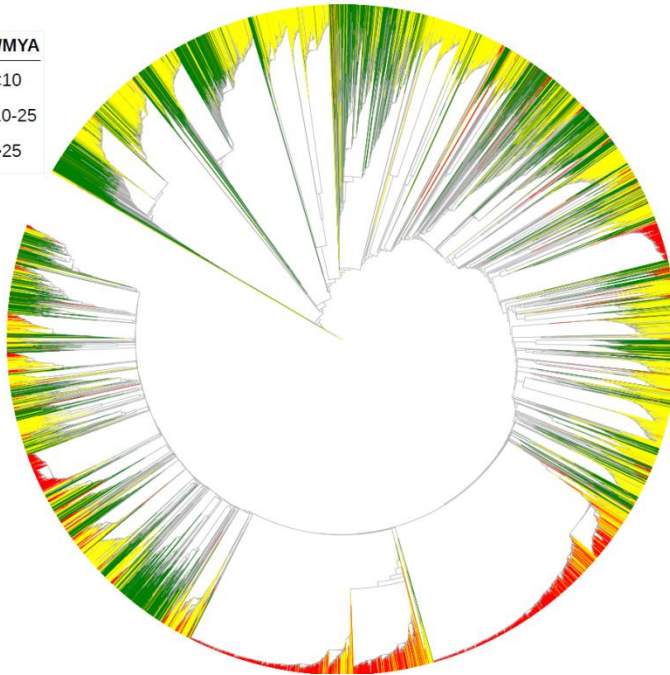


Supplementary Figure S2: Phylogenetic analysis of the *G. maculatum* with close-related teleosts

Tree scale: 0.1 ⇐

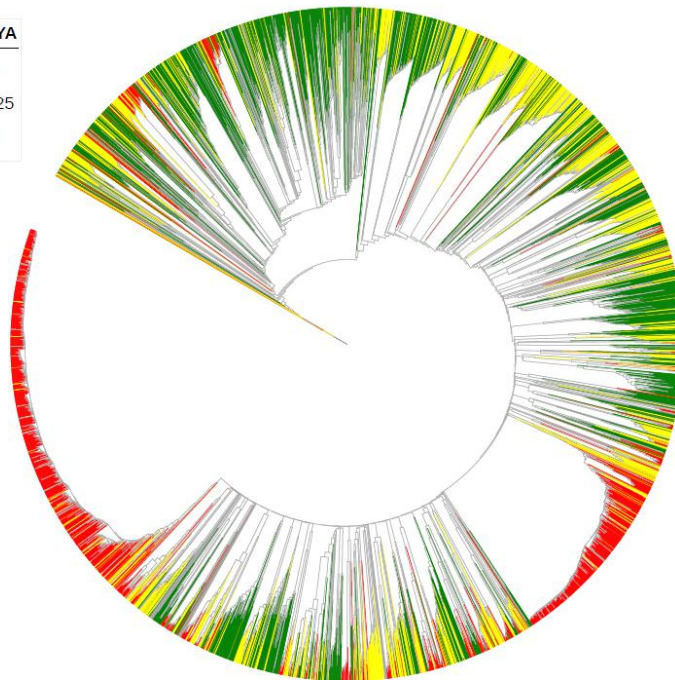
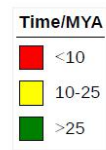
Time/MYA

■	<10
■	10-25
■	>25



Supplementary Figure S3: The phylogenetic analysis of LINE/RTE-Bovb in *G. maculatum* genome. The green, yellow and red color represented the repeat elements inserted 25, 10-25 and less than 10 Ma, respectively.

Tree scale: 0.1



Supplementary Figure S4: The phylogenetic analysis of LINE/L2 in *G. maculatum* genome. The green, yellow and red color represented the repeat elements inserted 25, 10-25 and less than 10 Ma, respectively.