

Supplementary Materials

Supplementary Table 1 Hi-C–anchored chromosome-scale assembly of *T. rotundifolia*.

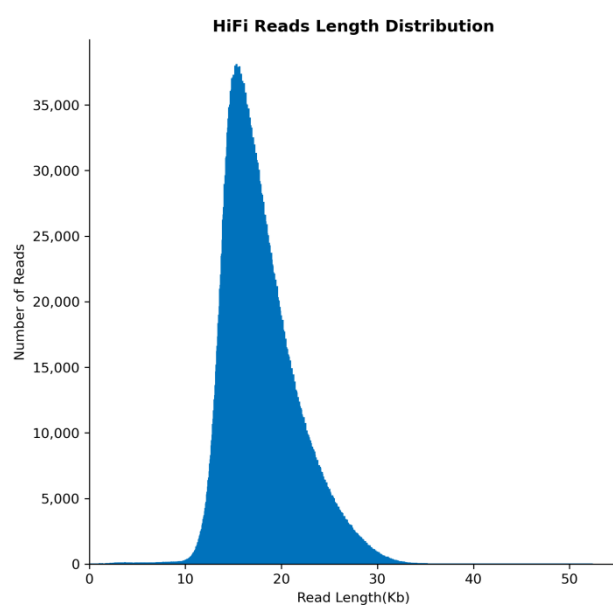
Supplementary Table 2 Structural statistics of gene models predicted using different annotation strategies in *T. rotundifolia*.

Supplementary Table 3 Functional annotation coverage of predicted genes across public databases.

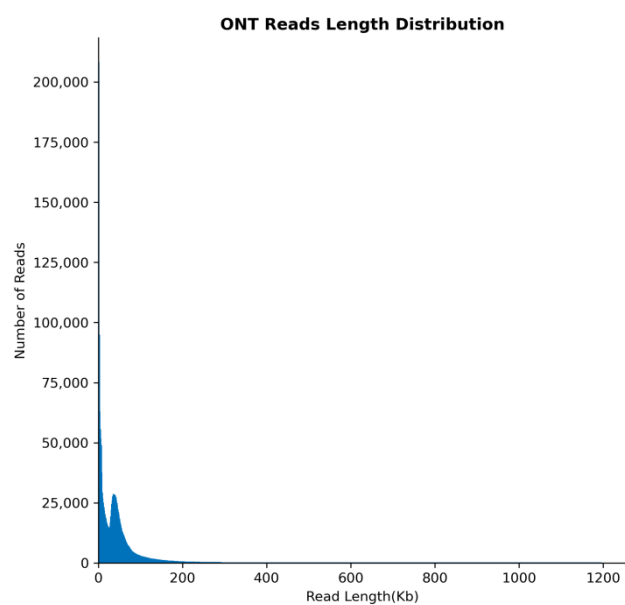
Supplementary Table 4 Per-chromosome base-level accuracy (QV) and estimated error rates of the *T. rotundifolia* genome assembly.

Supplementary Table 5 BUSCO completeness assessment of the *T. rotundifolia* genome assembly and annotated protein-coding genes.

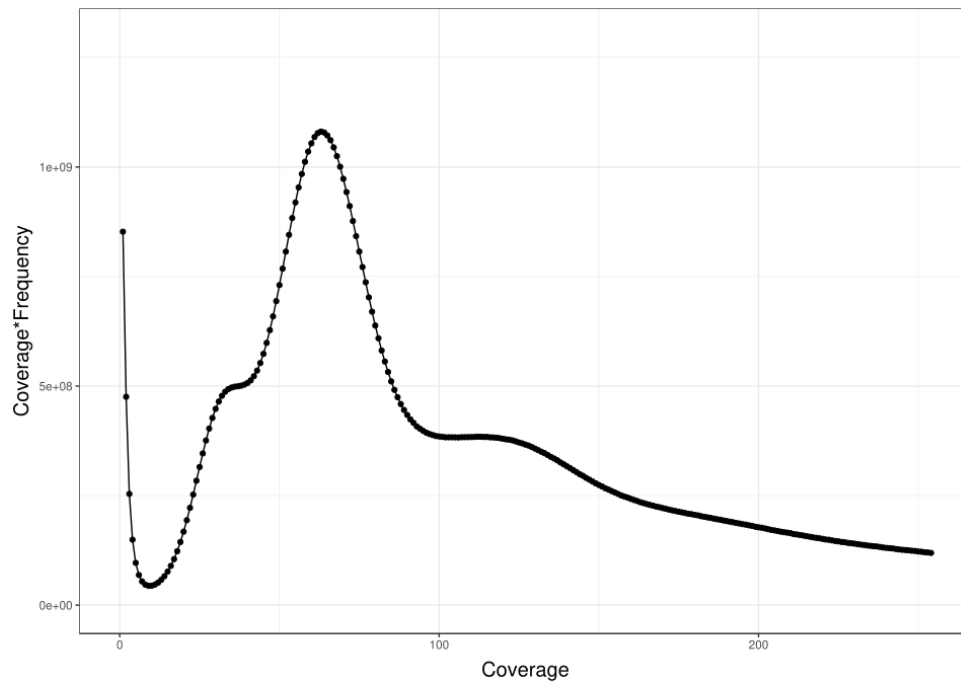
Supplementary Table 6 Detection of canonical telomeric repeats at chromosome termini in the *T. rotundifolia* genome assembly.



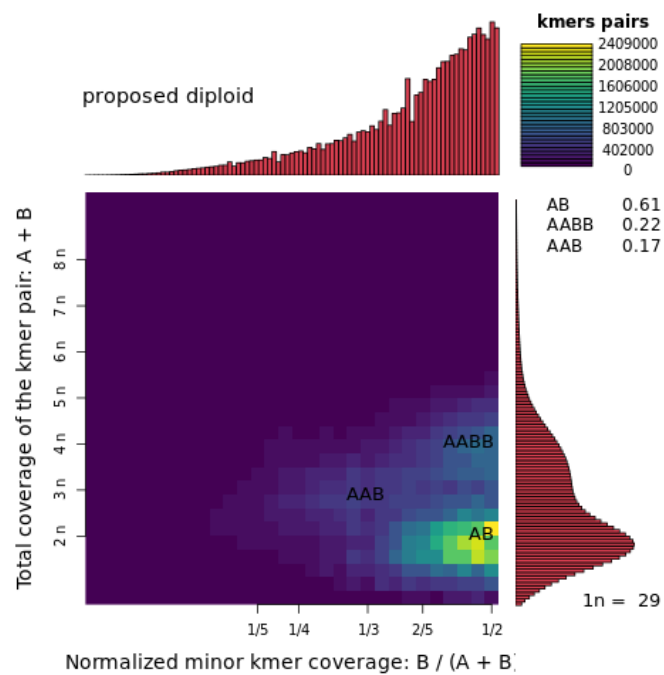
Supplementary Figure S1 The HiFi reads sequence length distribution of *G. rochechouartii* genome.



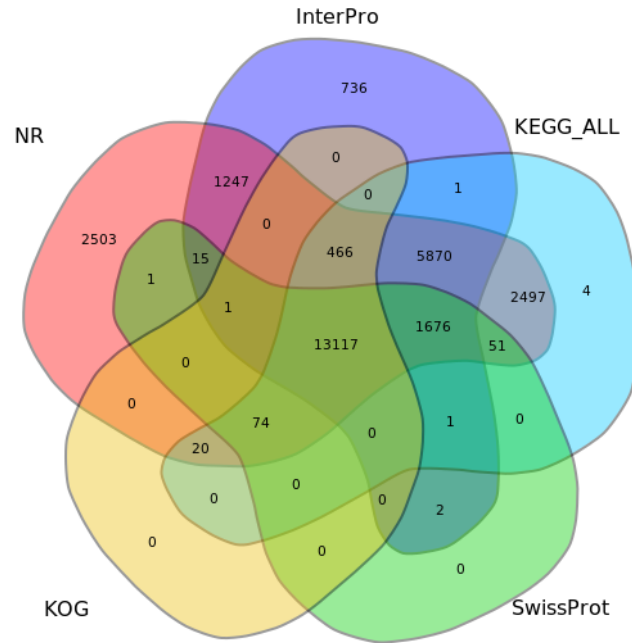
Supplementary Figure S2 The ONT length distribution of *G. rochechouartii* genome.



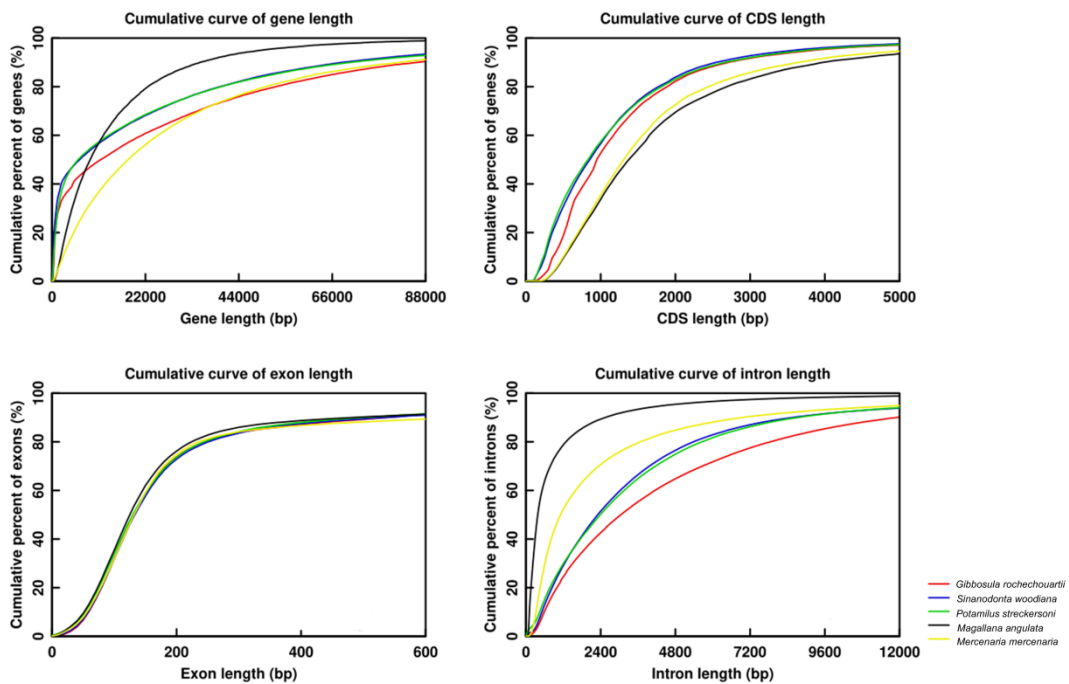
Supplementary Figure S3 Distribution of k-mer depth and frequency at 17 k-mer for *G. rochechouartii* genome.



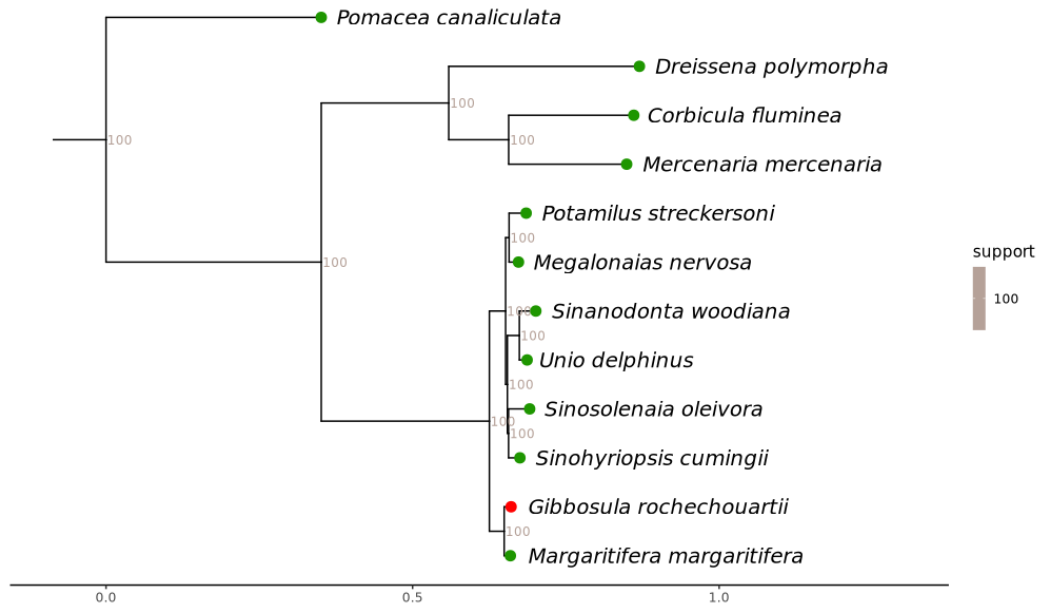
Supplementary Figure S4 Smudgeplots K-mer graph for genome characterization.



Supplementary Figure S5 The venn diagram of annotation results of five databases for *G. rochechouartii* genome.



Supplementary Figure S6 Comparison of gene length (A), CDS length (B), exon length (C) and intron length (D) among four related gene sets of *G. rochechouartii*.



Supplementary Figure S7 The phylogenetic tree constructed based on Universal Single-Copy Orthologs (USCOs).

Supplementary Table S1 The ONT data statistics of *G. rochechouartii* genome.

Length Cutoff	Total number	Total number (>2kb)	Total bases (bp)	Max length (bp)	Mean length (bp)	N50 (bp)	N90 (bp)	GC content (%)
all	1,715,465	1,438,915	61,034,820,494	1,196,242	35,579	60,567	25,763	36
>=50kb	413,172	413,172	37,253,444,225	1,196,242	90,164	93,642	55,433	36

Supplementary Table S2 The Preliminary genome assembly results of *G. rochechouartii* genome.

Mode	Total number	Total number (>= 2 kb)	Total length (bp)	max length (bp)	Mean length (bp)	N50 (bp)	N90 (bp)	GC content (%)
Hifiasm	77	77	3,036,315,524	333,463,444	39,432,669	136,783,532	54,436,468	35.9

Supplementary Table S3 The Hi-C assisted assembly of *G. rochechouartii* genome.

Genome version	Sequence Category	Scaffold				Contig			
		Sequence length (bp)	Sequence number	N50 (bp)	N90 (bp)	Sequence length (bp)	Sequence number	N50 (bp)	N90 (bp)
before anchored	whole genome	3,036,315,524	77	136,783,532	54,436,468	3,036,315,524	77	136,783,532	54,436,468
	whole genome	2,989,716,572	55	155,482,520	129,327,949	2,989,714,572	75	121,126,762	41,240,849
after anchored	chromosome	2,982,894,982	19	155,482,520	129,327,949	2,982,892,982	39	127,875,216	41,240,849
	free sequence	6,821,590	36	359,049	77,387	6,821,590	36	359,049	77,387

Supplementary Table S4 The details of chromosome and number of telomere characteristic sequences for *G. rochechouartii* genome.

Chromosome	Length(bp)	Number of contigs	Number of gaps	Number of start repeat unit	Number of end repeat unit
chr1	155,546,315	2	1	22	62
chr2	154,795,352	1	0	590	407
chr3	129,289,573	3	2	0	1,175
chr4	175,308,678	2	1	750	1,513
chr5	147,200,883	1	0	0	0
chr6	186,210,341	2	1	0	0
chr7	185,483,780	1	0	0	0
chr8	194,196,100	1	0	0	0
chr9	139,128,551	1	0	35	0
chr10	121,285,161	1	0	825	78
chr11	166,223,848	1	0	634	764
chr12	147,976,893	1	0	289	631
chr13	149,331,216	1	0	505	0
chr14	136,659,519	1	0	29	702
chr15	136,857,721	1	0	0	266
chr16	166,597,950	1	0	881	0
chr17	166,270,117	1	0	638	0
chr18	123,072,924	3	2	0	688
chr19	201,955,789	1	0	462	163

Supplementary Table S5 Statistical results of repetitive sequences.

Type	Repeat Size (bp)	% of genome
Trf	231,164,683	7.74
Repeatmasker	316,743,746	10.61
Proteinmask	299,697,429	10.04
<i>De novo</i>	1,434,883,546	48.06
Total	1,829,241,468	61.27

Supplementary Table S6 Statistics of non-coding RNA annotation.

	Type	Copy	Average length (bp)	Total length (bp)	Percentage of genome (%)
miRNA	-	725	74	53,779	0.001801
tRNA	-	1,483	74	110,346	0.003696
rRNA	rRNA	526	293	153,957	0.005157
	18S	57	1,653	94,235	0.003156
	28S	48	160	7,684	0.000257
	5.8S	56	154	8,604	0.000288
	5S	365	119	43,434	0.001455
	snRNA	396	152	60,309	0.00202
	CD-box	73	96	7,015	0.000235
snRNA	HACA-box	37	217	8,029	0.000269
	splicing	285	158	45,140	0.001512
	scaRNA	1	125	125	0.000004

Supplementary Table S7 Species taxonomic information and accession code of all samples used in this study.

Species	Version
<i>Sinanodonta woodiana</i>	GCA_046254885.1
<i>Sinosolenia oleivora</i>	GCA_039955035.1
<i>Corbicula fluminea</i>	GCA_001632725.1
<i>Unio delphinus</i>	GCA_029339505.1
<i>Dreissena polymorpha</i>	GCA_020536995.1
<i>Margaritifera margaritifera</i>	GCA_029931535.1
<i>Megalonias nervosa</i>	GCA_016617855.1
<i>Sinohyriopsis cumingii</i>	GCA_028554795.2
<i>Potamilus streckersoni</i>	GCA_016746295.1
<i>Mercenaria mercenaria</i>	GCF_021730395.1
<i>Pomacea canaliculata</i>	GCF_003073045.1

Supplementary Table S8 Calibration points used for divergence time estimation.

Taxon1	Taxon2	Divergence time (Mya)	CI(max-min) (Mya)
<i>Pomacea canaliculata</i>	<i>Corbicula fluminea</i>	535.6	541.7-515.0
<i>Sinohyriopsis cumingii</i>	<i>Mercenaria mercenaria</i>	440	510.9-230.0
<i>Mercenaria mercenaria</i>	<i>Dreissena polymorpha</i>	303	407.1-170.0
<i>Gibbosula rochechouartii</i>	<i>Margaritifera margaritifera</i>	171.6	178-168

Supplementary Table S9 Software, versions, and parameters used in this study.

Software	Version	Parameters
FastQC	0.11.3	default
Blastn	2.11.0+	-evalue 0.00001 -max_hsps 1
GCE	1.0.0	-k 17
smudgeplot	0.2.3dev	-k21 -m100 -ci1 -cs10000
hifiasm	0.19.9	--n-hap X
purge_haplotigs	1.1.3	-a 70
HiCUP	0.7.2	--NM 3
Juicer	1.6	default
juicebox	1.11.08	default
minimap2	2-2.24	ont:-ax map-ont ccs:-ax map-hifi
Blastn	2.11.0+	-evalue 0.00001 -max_hsps 1
TGS-GapCloser	v1.2.0	--min_nread 10
Pilon	1.23	--fix all --changes
TRF	4.09	2 7 7 80 10 50 2000 -d -h
RepeatModeler	open-1.0.11	default
LTR_FINDER_parallel	1.0.7	default
RepeatMasker	open-4.0.9	-nolow -no_is -norna species model: <i>Pomacea canaliculata</i> --uniqueGeneId=true
AUGUSTUS	3.3.2	--noInFrameStop=true --gff3=on --strand=both
Genscan	-	HumanIso.smat
HISAT2	2.1.0	--dta
Stringtie	1.3.5	default
TransDecoder	5.5.0	default
BUSCO	5.7.1	Genome BUSCO : --miniprot -m genom e -f --offline Protein BUSCO : -m protein -f --offline max_dna_len=3000000 min_contig=10000 pred_flank=500 min_protein=30
Maker2	2.31.10	-q
tRNAscan-SE	1.3.1	cmScan --rfam --nohmmonly
Rfam	14.8	-evalue 1e-5
blastp	2.11.0+	-S blast -I 1.5
orthofinder	2.5.5	-t=c
Gblocks	0.91b	-f a -N 100 -m GTRGAMMA
RAxML	8.2.12	clock = 3;model = 0
memctree	4.9	