

Table S1. The outlook of the phenotypic characteristics of published pineapple genomes.

Variety	Phenotype	Leaf spine	Leaf strip	Leaf color	Plant architecture	Peduncle length	Peduncle diameter	Fruit size	Eyes profile	Unripe fruit skin color	Nutrition growth period
BL		Present	Absent	Red	Semi-erect	Short	Large	Middle	Rrominent	Black	Middle
LY		Absent	Absent	Red	Erect	Long	Samll	Very samll	Flat	Red	Short
YLL		Present	Absent	Green	Semi-erect	Long	Samll	Very samll	Flat	Red	Very short
F153		Absent	Absent	Green	Semi-erect	Short	Large	Large	Flat	Black	Middle
MD2 (2016)		Absent	Absent	Green	Semi-erect	Short	Large	Large	Flat	Green	Middle
CB5 (2019)		Present	Absent	Green	Flattened	Long	Middle	Middle	Rrominent	Red	Long
CB5 (2022)		Present	Present	Chimera	Flattened	Long	Middle	Middle	Rrominent	Red	Long

Table S2. Sequencing data statistics

	BL		LY		YLL	
	Bases(G)	Depth(X)	Bases(G)	Depth(X)	Bases(G)	Depth(X)
HiFi	29.2	69	27.2	64	27.9	66
Hi-C	48.5	114	58.2	137	44.6	105
ONT	43.3	102	37	87	35.5	83
CUT Tag-Seq	7.6	17.9	7.5	17.7	9.2	21.7

note: Genome size was evaluated as about 423.8Mb

Table S3. Data statistics of genome assembly.

		BL		LY		YLL		Ref	MD2 (2016)	F153	CB5 (2019)	CB5 (2022)	Yugufa	
		hap1	hap2	hap1	hap2	hap1	hap2						3H	3C
contigs	Num. of contigs	404	184	465	216	695	287	\	\	\	\	\	\	\
	total length of cotnigs	417,549,365	427,304,259	430,198,546	414,602,401	450,481,546	419,792,559	\	\	\	\	\	\	\
	maximum contigs length	18,130,374	20,988,244	19,487,342	19,997,820	20,842,862	18,210,815	\	\	\	\	\	\	\
	mean contigs length	1,033,538	2,322,306	925,158	1,919,456	648,175	1,462,692	\	\	\	\	\	\	\
	contigs N50	13,660,326	15,220,253	13,918,218	13,829,636	12,546,884	11,638,651	\	\	\	\	\	\	\
chromosomes/sc affolds	Num. of chromosomes	25	25	25	25	25	25	25	8,448	25	25	25	25	25
	total length of anchored(bp)	399,020,914	417,431,254	415,015,407	406,882,445	425,105,455	405,024,226	423,847,674	524,069,662	315,839,004	486,611,595	483,169,765	401,640,934	431,138,038
	Anchored rate(%)	95.6	97.7	96.5	98.1	94.4	96.5	\	\	82.7	94.8	\	\	\
	Num. of Gaps	31	10	20	19	39	31	0	8,645	4,553	1,867	479	1,763	1,804
	GC content(%)	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.38	0.39	0.4	0.39	0.39
	Maximum chr. length(bp)	20,885,627	22,116,788	20,182,354	20,446,592	20,842,862	20,395,506	20,988,244	1,287,057	17,334,668	27,132,755	31,887,959	23,875,565	33,039,201
	Mean chr. length(bp)	15,960,837	16,697,250	16,600,616	16,275,298	17,004,218	16,200,969	16,953,907	62,035	12,633,560	19,464,464	19,326,790	16,065,637	17,245,522
	Median chr. length(bp)	16,215,998	16,847,914	17,185,802	16,622,272	17,119,228	16,077,226	17,604,455	24,890	12,612,916	20,178,443	19,385,557	15,426,415	16,674,444
	N50	16,448,991	17,047,075	17,185,802	16,692,395	17,119,228	16,701,519	18,210,815	153,084	13,096,602	20,311,332	20,383,276	15,497,171	17,173,289

Table S4. Manully checked and gathered a set of sequences to form the reference genome.

F153	bl.hap1		bl.hap2		ly.hap1		ly.hap2		yll.hap1		yll.hap2	
	id	length	id	length	id	length	id	length	id	length	id	length
contig1	h1tg000010l	15,805,296	h2tg000029l	14,158,845			h2tg000021l	13,829,636				
contig10			h2tg000010l	16,304,956			h2tg000009l	16,622,272	h1tg000037l	18,825,551		
contig11	h1tg000012l	17,953,106	h2tg000007l	19,023,145			h2tg000002l	16,692,395	h1tg000008l	16,622,846		
contig12	h1tg000017l	16,215,998	h2tg000014l	16,215,827	h1tg000001l	15,761,836	h2tg000028l	16,400,829	h1tg000033l	17,604,455		
contig13									h1tg000001l	14,656,638		
contig14			h2tg000002l	17,189,056			h2tg000019l	19,188,885				
contig15	h1tg000035l	15,594,272	h2tg000001l	15,220,253	h1tg000009l	13,885,599						
contig16	h1tg000019l	17,579,196	h2tg000012l	16,847,914								
contig17			h2tg000011l	12,894,463							h2tg000020l	12,740,338
contig18			h2tg000023l	14,493,109	h1tg000018l	15,668,925			h1tg000011l	15,789,271		
contig19	h1tg000029l	13,660,326	h2tg000024l	14,717,538								
contig2			h2tg000027l	18,373,142	h1tg000005l	17,406,915	h2tg000003l	17,217,520	h1tg000010l	17,119,228	h2tg000027l	17,727,638
contig20	h1tg000004l	13,785,295	h2tg000019l	13,895,535							h2tg000014l	13,316,313
contig21	h1tg000001l	14,862,238			h1tg000011l	18,850,705	h2tg000029l	14,922,600	h1tg000005l	18,890,200		
contig22					h1tg000002l	13,918,218			h1tg000017l	12,546,884		
contig23					h1tg000015l	13,557,329					h2tg000030l	13,004,031
contig24	h1tg000011l	12,353,176	h2tg000022l	12,409,377	h1tg000016l	12,428,122	h2tg000022l	13,801,193	h1tg000063l	9,846,527	h2tg000054l	8,752,364
contig25			h2tg000009l	16,284,938	h1tg000008l	15,915,826	h2tg000025l	14,874,981	h1tg000027l	14,680,481		
contig3					h1tg000003l	19,297,848						
contig4									h1tg000006l	20,842,862		
contig5			h2tg000020l	20,988,244	h1tg000006l	18,852,848	h2tg000001l	19,997,820				
contig6	h1tg000005l	18,130,374	h2tg000025l	17,952,691			h2tg000016l	18,892,189			h2tg000021l	15,012,427
contig7	h1tg000016l	17,337,001	h2tg000003l	17,375,442							h2tg000005l	18,210,815
contig8					h1tg000022l	19,487,342	h2tg000023l	19,494,550	h1tg000025l	18,088,540		
contig9	h1tg000015l	16,448,991	h2tg000015l	17,047,075	h1tg000024l	17,113,717	h2tg000011l	17,027,274				

Table S5. Reference genome information statistics.

Chr	Length (bp)	raw gene num	manual check gene num
contig01	15,805,296	876	878
contig02	18,373,142	1,032	1,037
contig03	19,297,848	1,393	1,407
contig04	20,842,862	1,368	1,405
contig05	20,988,244	1,444	1,484
contig06	18,892,189	1,268	1,290
contig07	18,210,815	1,130	1,157
contig08	19,494,550	1,304	1,347
contig09	17,113,717	1,303	1,329
contig10	18,825,551	1,193	1,179
contig11	19,023,145	1,045	1,067
contig12	17,604,455	984	1,033
contig13	14,656,638	1,106	1,158
contig14	19,188,885	875	909
contig15	15,594,272	1,017	1,064
contig16	17,579,196	1,000	1,004
contig17	12,894,463	817	804
contig18	15,789,271	989	984
contig19	14,717,538	956	959
contig20	13,895,535	823	851
contig21	18,890,200	802	840
contig22	13,918,218	798	826
contig23	13,557,329	856	884
contig24	12,409,377	746	764
contig25	16,284,938	1,037	996
Total	423,847,674	26,162	26,656

Table S6. Repeat element statistics.

Class	Count	bpMasked	%masked
LTR			
Copia	29,638	23,669,266	5.58%
Gypsy	66,720	136,212,194	32.14%
unknown	33,793	27,527,466	6.49%
TIR			
CACTA	19,922	6,611,009	1.56%
Mutator	67,237	21,682,280	5.12%
PIF_Harbinger	8,792	2,848,858	0.67%
Tc1_Mariner	3,566	732,199	0.17%
hAT	32,526	10,934,099	2.58%
nonLTR			
LINE_element	395	279,738	0.07%
unknown	58	32,646	0.01%
nonTIR			
helitron	85,531	25,401,269	5.99%
repeat_region	12,538	5,112,000	1.21%
Total	360,716	261,043,024	61.59%

Table S7. The LTR assembly index of different genomes.

Chr	From	To	Intact	Total	raw LAI	LAI
bl.hap1	1	407,856,775	0.0792	0.5122	15.46	21.1
bl.hap2	1	422,361,479	0.0831	0.516	16.09	18.6
ly.hap1	1	419,518,595	0.078	0.521	14.97	19.9
ly.hap2	1	408,711,460	0.0773	0.5155	14.99	18.8
yll.hap1	1	430,229,670	0.076	0.5279	14.39	17.8
yll.hap2	1	411,235,785	0.0734	0.5138	14.3	19.2
Yugafu_3H	1	401,464,593	0.0764	0.4707	16.22	19.3
Yugafu_3C	1	430,957,593	0.0738	0.4702	15.7	17.6
GL1	1	483,117,357	0.0959	0.5661	16.94	17.6
reference	1	423,847,674	0.0798	0.5344	14.93	17.5
CB5	1	486,424,895	0.0604	0.4881	12.37	12.5
F153	1	311,256,003	0.0106	0.3388	3.13	8.8
ACMD2	1	509,748,646	0.0236	0.4902	4.81	3.8

Table S8. BUSCO analysis of genome assembly(genome)(embryophyta_odb10).

[illegible]

Table S9. BUSCO analysis of genome annotation(protein)(embryophyta_odb10).

[illegible]

Table S10. The genome coverage and mapping ratio by HiFi, ONT and mRNAseq reads, respectively

Genome	Length	HiFi	ONT	mRNAseq (34 libraries average)
reference	423,847,674	99.98%	99.95%	88.35%
bl.hap1	417,542,093	99.75%	99.65%	×
bl.hap2	427,300,947	99.87%	99.90%	×
ly.hap1	430,190,176	99.80%	99.79%	×
ly.hap2	414,598,513	99.90%	99.93%	×
yll.hap1	450,469,036	99.84%	98.11%	×
yll.hap2	419,787,393	99.93%	99.70%	×

Table S11. The genome quality of the reference and six haplotypes was assessed by Merquy software.

Object	Chr	Error	Length	QV	pvalue
reference	contig01	145	15,805,278	63.16	4.83E-07
	contig10	43	18,825,533	69.20	1.20E-07
	contig11	209	19,023,127	62.38	5.78E-07
	contig12	61	17,604,437	67.39	1.82E-07
	contig13	32	14,656,620	69.40	1.15E-07
	contig14	0	19,188,867	inf	0
	contig15	75	15,594,254	65.97	2.53E-07
	contig16	123	17,579,178	64.34	3.68E-07
	contig17	94	12,894,445	64.16	3.84E-07
	contig18	26	15,789,253	70.62	8.67E-08
	contig19	155	14,717,520	62.56	5.54E-07
	contig02	282	18,373,124	60.93	8.08E-07
	contig20	83	13,895,517	65.03	3.14E-07
	contig21	32	18,890,182	70.50	8.92E-08
	contig22	79	13,918,200	65.25	2.99E-07
	contig23	0	13,557,311	inf	0
	contig24	33	12,409,359	68.54	1.40E-07
	contig25	94	16,284,920	65.17	3.04E-07
	contig03	64	19,297,830	67.58	1.75E-07
	contig04	0	20,842,844	inf	0
	contig05	73	20,988,226	67.37	1.83E-07
	contig06	32	18,892,171	70.50	8.91E-08
	contig07	39	18,210,797	69.48	1.13E-07
	contig08	75	19,494,532	66.94	2.02E-07
	contig09	65	17,113,699	66.99	2.00E-07
six haplotypes	bl.hap1	5,959	417,542,093	61.24	7.51E-07
	bl.hap2	3,574	427,300,947	63.56	4.40E-07
	ly.hap1	4,390	430,190,176	62.70	5.371E-07
	ly.hap2	1,584	414,598,513	66.97	2.01E-07
	yll.hap1	6,759	450,469,036	61.03	7.90E-07
	yll.hap2	2,369	419,787,393	65.27	2.97E-07

Table S12. The genome quality value of the reference and six haplotypes was assessed by CRAQ.

Object	Chr	Covered.Rate	Low-confident.Rate	Avg.CRH	Avg.CSH	Avg.CRE	R-AQI	Avg.CSE	S-AQI
reference	contig01	100.00%	0.00%	19.73	3.26	0.19	98.12	0.00	100.00
	contig10	99.92%	0.26%	23.43	3.93	0.11	98.94	0.00	100.00
	contig11	99.98%	0.00%	16.91	3.01	0.16	98.43	0.00	100.00
	contig12	100.00%	0.06%	22.67	3.23	0.17	98.31	0.00	100.00
	contig13	99.74%	0.23%	28.98	5.32	0.14	98.64	0.00	100.00
	contig14	99.98%	0.14%	25.45	4.14	0.10	98.96	0.00	100.00
	contig15	100.00%	0.01%	19.79	2.98	0.13	98.73	0.00	100.00
	contig16	99.97%	0.26%	20.78	2.75	0.11	98.87	0.00	100.00
	contig17	99.97%	0.00%	23.05	3.11	0.16	98.46	0.00	100.00
	contig18	99.98%	0.01%	24.58	3.28	0.13	98.74	0.00	100.00
	contig19	99.87%	0.00%	12.59	2.30	0.14	98.65	0.14	87.28
	contig02	99.94%	0.08%	18.09	3.05	0.11	98.92	0.00	100.00
	contig20	99.98%	0.00%	19.36	3.17	0.14	98.57	0.00	100.00
	contig21	99.92%	0.36%	18.89	3.20	0.11	98.95	0.05	94.84
	contig22	99.98%	0.00%	27.42	5.43	0.14	98.57	0.00	100.00
	contig23	99.99%	0.01%	28.96	4.01	0.15	98.54	0.07	92.89
	contig24	100.00%	0.00%	21.54	2.86	0.16	98.40	0.00	100.00
	contig25	100.00%	0.00%	14.77	2.12	0.12	98.78	0.00	100.00
	contig03	99.99%	0.00%	35.64	5.98	0.10	98.97	0.00	100.00
	contig04	99.96%	0.00%	25.01	2.75	0.10	99.04	0.00	100.00
	contig05	99.96%	0.01%	24.84	3.96	0.10	99.05	0.05	95.35
	contig06	100.00%	0.05%	30.74	4.76	0.16	98.42	0.05	94.84
	contig07	98.89%	1.53%	23.90	3.17	0.19	98.08	0.22	80.08
	contig08	99.97%	0.00%	29.36	4.89	0.10	98.98	0.05	95.00
	contig09	99.99%	0.00%	33.40	4.62	0.18	98.26	0.00	100.00
	average	99.92%	0.12%	23.60	3.65	0.14	98.66	0.03	97.61
six haplotypes	bl.hap2	99.91%	0.19%	20.06	3.29	0.15	98.48	0.01	99.05
	bl.hap1	99.83%	0.41%	21.35	3.46	0.23	97.77	0.02	97.55
	ly.hap1	99.92%	0.35%	29.10	4.85	0.18	98.24	0.03	97.15
	ly.hap2	99.85%	0.25%	30.08	5.00	0.18	98.22	0.07	93.69
	yll.hap1	99.76%	0.49%	22.88	3.50	0.22	97.87	0.01	99.07
	yll.hap2	99.77%	0.35%	24.89	3.67	0.22	97.79	0.04	95.67

Table S13. The length and frequency of tandem repeats in Reference genome identified by TRF.

length	frequency	percentage(%)
166	217,605	47.6%
332	52,914	11.6%
498	37,351	8.2%
156	25,321	5.5%
83	11,496	2.5%
322	8,423	1.8%
61	7,533	1.6%
312	5,981	1.3%
468	5,228	1.1%
497	4,458	1.0%
488	3,912	0.9%
other	76,517	16.8%

Table S14. The locations of telomere and centromere in genome.

Chr	telomere				centromere		
	P1 Start	P1 End	P2 Start	P2 End	start	end	length (Mb)
contig01	1	10,015	15,799,325	15,805,296	4,533,354	7,064,759	2.53
contig02	1	13,580	18,363,407	18,373,142	8,322,278	11,352,779	3.03
contig03	1	11,871	19,288,243	19,297,848	6,497,183	7,194,617	0.70
contig04	1	9,405	20,820,885	20,842,862	7,206,798	9,715,274	2.51
contig05	1	10,617	20,970,343	20,988,244	10,883,923	12,961,697	2.08
contig06	1	3,351	18,881,172	18,892,189	10,310,973	12,441,298	2.13
contig07	1	8,155	18,192,947	18,210,815	8,956,740	11,917,685	2.96
contig08	1	16,563	19,485,727	19,494,550	8,773,911	9,821,144	1.05
contig09	1	12,194	17,100,618	17,113,717	7,054,468	9,186,633	2.13
contig10	1	16,143	18,813,654	18,825,551	7,257,635	11,166,157	3.91
contig11	1	12,523	19,009,452	19,023,145	9,564,176	12,782,160	3.22
contig12	1	5,442	17,597,842	17,604,455	9,431,156	11,953,365	2.52
contig13	1	16,626	14,635,384	14,656,638	9,409,871	10,361,410	0.95
contig14	1	9,534	19,178,009	19,188,885	8,668,548	13,105,788	4.44
contig15	1	17,376	15,582,891	15,594,272	8,453,767	11,301,516	2.85
contig16	1	12,658	17,574,712	17,579,196	10,215,336	12,272,662	2.06
contig17	1	14,584	12,883,696	12,893,858	4,902,251	6,249,928	1.35
contig18	1	13,979	15,764,912	15,789,271	9,197,692	11,143,261	1.95
contig19	1	19,113	14,696,365	14,717,538	3,933,327	5,830,058	1.90
contig20	1	10,295	13,882,445	13,895,535	5,343,063	6,677,287	1.33
contig21	1	7,502	18,876,706	18,890,200	8,143,953	12,943,194	4.80
contig22	1	10,951	13,907,375	13,918,218	6,160,014	7,746,795	1.59
contig23	1	14,725	13,553,126	13,557,329	5,167,455	7,438,049	2.27
contig24	1	7,428	\	\	8,071,432	9,746,297	1.67
contig25	1	16,284	16,283,014	16,284,938	8,653,398	11,080,375	2.43

Table S15. manual check gene statistics.

chr	cds modified	genenum	pers(%)
1	134	878	15.3
2	150	1,037	14.5
3	233	1,407	16.6
4	137	1,405	9.8
5	163	1,484	11.0
6	118	1,290	9.1
7	101	1,157	8.7
8	146	1,347	10.8
9	130	1,329	9.8
10	222	1,179	18.8
11	160	1,067	15.0
12	141	1,033	13.6
13	166	1,158	14.3
14	140	909	15.4
15	140	1,064	13.2
16	120	1,004	12.0
17	164	804	20.4
18	187	984	19.0
19	201	959	21.0
20	98	851	11.5
21	101	840	12.0
22	86	826	10.4
23	100	884	11.3
24	79	764	10.3
25	175	996	17.6

Table S16. Gene model information summary.

reference	gene num	average mRNA length	average exon length	average intron length	average cds length	average utr length
Ref	26,656	4,244	230	647	230	/
CB5 (2022)	26,113	3,773	264	567	219	242
CB5 (2019)	29,415	6,014	199	984	187	161
MD2 (2016)	23,598	4,295	253	628	222	/
F153	26,730	6,968	218	1,350	218	246
3H	45,983	3,526	234	545	211	257
3C	49,334	3,517	235	543	212	256

Table S17. Primer sequences required for this research experiment.

Primers		sequences (5'-3')
AcActin	F	CTGGCCTACGTGGCACTTGACTT
	R	CACTTCTGGGCAGCGGAACCTTT
AcMYB267-CDS	F	ATGGGGAGGAGACCTTGTTGTT
	R	TCAAATACATGCAAAATCGCCA
AcMYB267-pDONR	F	AAAAAGCAGGCTCCATGGGGAGGAGACCTTGTTGTT
	R	AGAAAGCTGGGTTTCAAATACATGCAAAATCGCCA
AcCHSpro-p0800-Luc	F	CTTGAATTCGAGCTCGGTACCATTTGGATTATCGCTTGCTTTTAA
	R	TGTTTTTGGCGTCTTCCATGGCATCGCCCTCGCTACTCTATCC
AcDFRpro-p0800-Luc	F	CTTGAATTCGAGCTCGGTACCCGCTAAATCCTTAGTTCTCC
	R	TGTTTTTGGCGTCTTCCATGGCTCTCCCCCGATCTCTACAC
AcANSpro-p0800-Luc	F	CTTGAATTCGAGCTCGGTACCGTGAGCCTCATTACAATAGCC
	R	TGTTTTTGGCGTCTTCCATGGATACACAATAATACACACTCTCC
AcUFGTpro-p0800-Luc	F	CTTGAATTCGAGCTCGGTACCCCCAGGTACTAAGTAGAAGAA
	R	TGTTTTTGGCGTCTTCCATGGTTTAGGGGGGAGGGTTGCAA