ERGA Assembly Report

v24.04.03_beta

Tags: ERGA-BGE

TxID	60953						
ToLID	wrPlaCost1						
Species	Placobdella costata						
Class	Clitellata						
Order	Rhynchobdellida						

Genome Traits	Expected	Observed
Haploid size (bp)	486,921,211	488,180,838
Haploid Number	9 (source: ancestor)	18
Ploidy	2 (source: ancestor)	2
Sample Sex	Н	Н

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.7.Q49

The following metrics were automatically flagged as below EBP recommended standards or different from expected:

- . Observed Haploid Number is different from Expected
- . BUSCO single copy value is less than 90% for pri
- . Assembly length loss > 3% for pri
- . More than 1000 gaps/Gbp for pri

Curator notes

- . Interventions/Gb: 159
- . Contamination notes: "Total length of scaffolds removed: 7,620,786 (1.5%); Scaffolds removed: 12 (12.1%); Largest scaffold removed: (3,924,856); FCS-GX contaminant species (number of scaffolds; total length of scaffolds): Yokenella regensburgei, g-proteobacteria (4; 4,053,046); Myxococcus qinghaiensis, d-proteobacteria (1; 3,257,129); Chryseotalea sanaruensis, CFB group bacteria (1; 56,177); Polyangium spumosum, d-proteobacteria (1; 44,272); Citrobacter youngae, g-proteobacteria (1; 43,784); Chryseolinea soli, CFB group bacteria (1; 43,047); Chryseolinea sp., CFB group bacteria (1; 32,751); Chryseotalea sp. WA131a, CFB group bacteria (1; 28,284)"
- . Other observations: "Hi-C from a different individual (wrPlaCost2), so assembly is not Hi-C phased; Expected haploid number from GoaT is an ancestral estimate, likely incorrect in this case"

Quality metrics table

Metrics	Pre-curation pri	Curated pri				
Total bp	504,255,463	488,180,838				
GC %	32.14	31.77				
Gaps/Gbp	1,253.33	1,349.91				
Total gap bp	126,400	131,800				
Scaffolds	99	28				
Scaffold N50	32,061,669	31,539,400				
Scaffold L50	6	6				
Scaffold L90	15	15				
Contigs	731	687				
Contig N50	1,889,692	1,764,884				
Contig L50	64	64				
Contig L90	337	339				
QV	49.5	49.5				
Kmer compl.	98.85	97.56				
BUSCO sing.	89.0%	89.7%				
BUSCO dupl.	1.0%	0.7%				
BUSCO frag.	4.8%	4.3%				
BUSCO miss.	5.2%	5.3%				

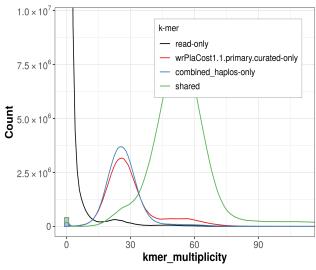
Warning: BUSCO versions or lineage datasets are not the same across results

HiC contact map of curated assembly

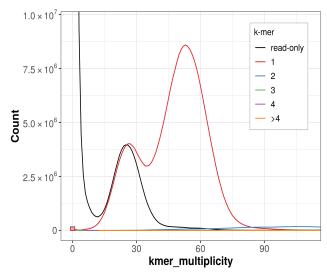
	The state of the s	September 1	## 	Will be seen and the seen and t		1000年					***	No.	THE PARTY OF	~清津清水	・ は 対 数 対 数 数 数 1	Carlo State	一 の ではない はない	Company (Arthur)
300 3 00 0 33 (10 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0			100 March 18 - 18 - 18 (19 m) 18 (19	12 10 10 10 10 10 10 10 10 10 10 10 10 10		1		100 mm	25 Mary 1		100	11.00		***************************************	- 日本語語	- Company of the Company	がは神経の	11 (11 (12 (12 (12 (12 (12 (12 (12 (12 (
	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0			200		100			#					The second second	The state of	変	A STREET, S.	CAMPA
131 1111	de la companya de la			111 111		1	i dan		1			11.1	441	4	. 1.34R.	13 (4794).	1,000	1.044
200 - 100 -					1000		40)						101 101 101 101 101	1. 接種	1122	29.00	11000	・おはまま
J·秦祥等 尹东北。	THE RESIDENCE	1 000 PT 01			×	V	Tage		m#	/建一/	4	推	312	4	-	- 1888.	10000	10 146
14444	i k etale		##1	11	1 N. W. W.		The		#	導	镰		111	事	#	-	4	第
(######	P PROBLE		PER I	A PROPERTY.					投资	響		- 18	ale:	*	養	100	ii.	水
300 X 102 11	21.000	i m ata.	141			1811				4	#	-12	1	-	#	-	STEEL STEEL	
- 1 1 1 1 1 1 1 1 1 1	C R se assiste	A MARKET STATE	-dam 6-1	144		(45)	101-	輸到	186	The same of the sa	:#		神	1	-	Ł	1	jii.
職員主権を対	Ha E annal	EMPERTY.	ame t	1,11 8/8	10	罐料	- Victorial	10	預點	#		-	雅	*	18	養	STIP.	- AEC
30 1 33 H	12 C AND ((Alberta)	mai i	anie.		i Mirit		Mag.	III.		-		114.	建	種	憲	100	で素
Jan Latin	15,450	1000 to 1	and a		100	191 :	***		112	100	温	11	M		世	*	羅	4
See a district	in with the	vaster is to the	district.	100	-	en in	7940	No.	esa:	*	-	1000		1	100	*	34	Á
THE RESERVE	THE MORE	CONTRACTOR OF THE	appear to	THE ROOM	-	mak eğet Hakkes e	(alter	施 拉	and note:	100 m	Table.	非 開於	948 *1**	連	1	1	-排 11:	18
11000 1 100 15	and the second	Transferred to	Appearance of the Control of the Con	11000		CONTRACT.	, etc.	400	ARTS.	CMC.	ATTE	1000	TATE		-	T		10
1300 Car	23,6630	18000	Appears	- syme		print.	Today .	1990 P	() and a		200	1.00E	parties.	1000	100	孝	8	

pri [LINK]

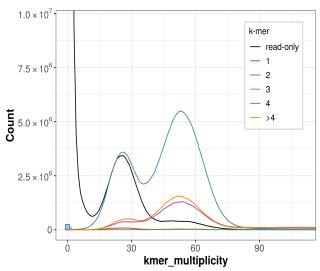
K-mer spectra of curated assembly



Distribution of k-mer counts coloured by their presence in reads/assemblies

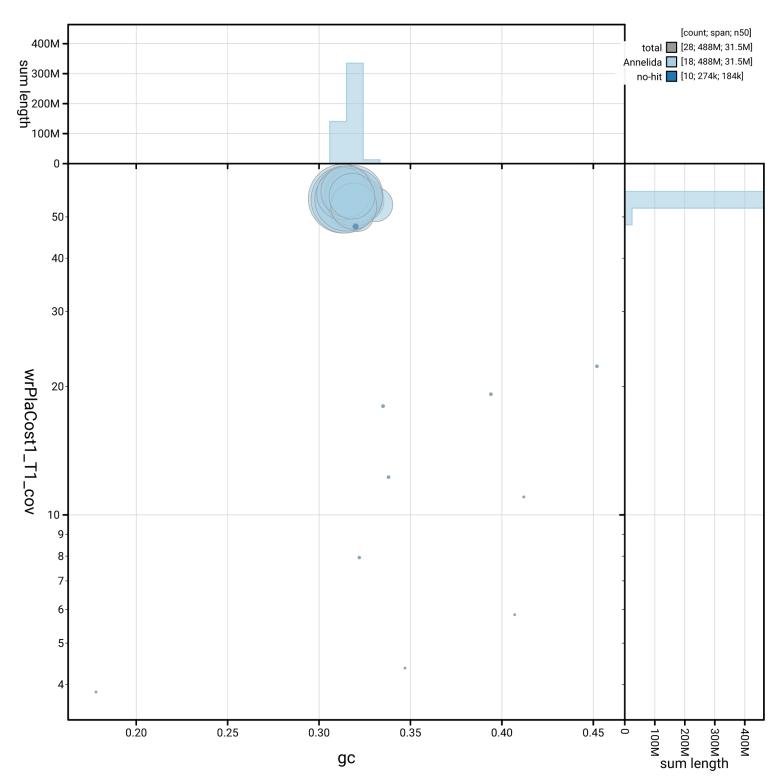


Distribution of k-mer counts per copy numbers found in asm



Distribution of k-mer counts per copy numbers found in asm

Post-curation contamination screening



pri. Bubble plot circles are scaled by sequence length, positioned by coverage and GC
proportion, and coloured by taxonomy. Histograms show total assembly length
distribution on each axis.

Data profile

Data	PacBio HiFi	Arima v2
Coverage	54x	204x

Assembly pipeline

Curation pipeline

Submitter: Thomas Mathers

Affiliation: WSI

Date and time: 2024-09-26 11:59:31 CEST