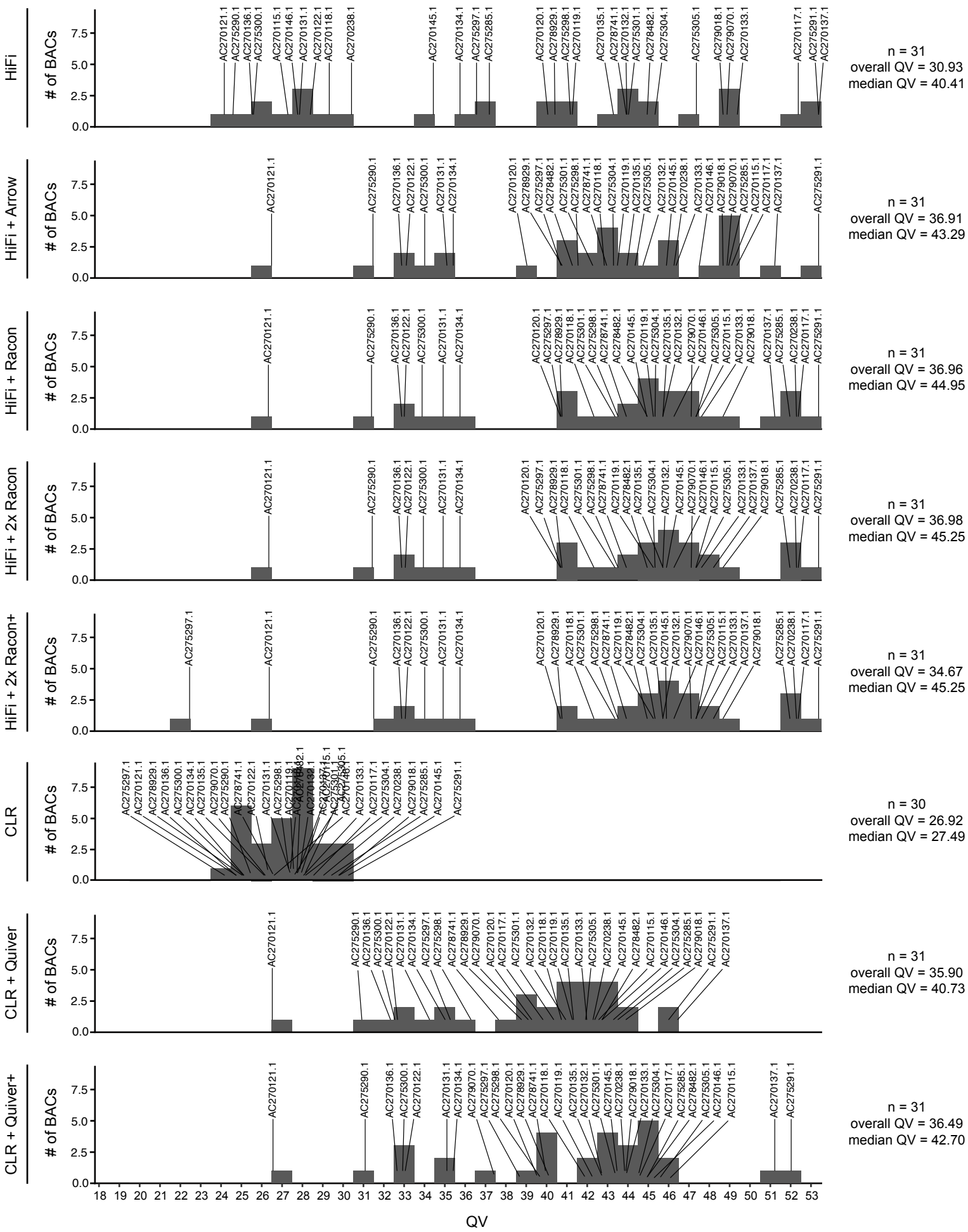


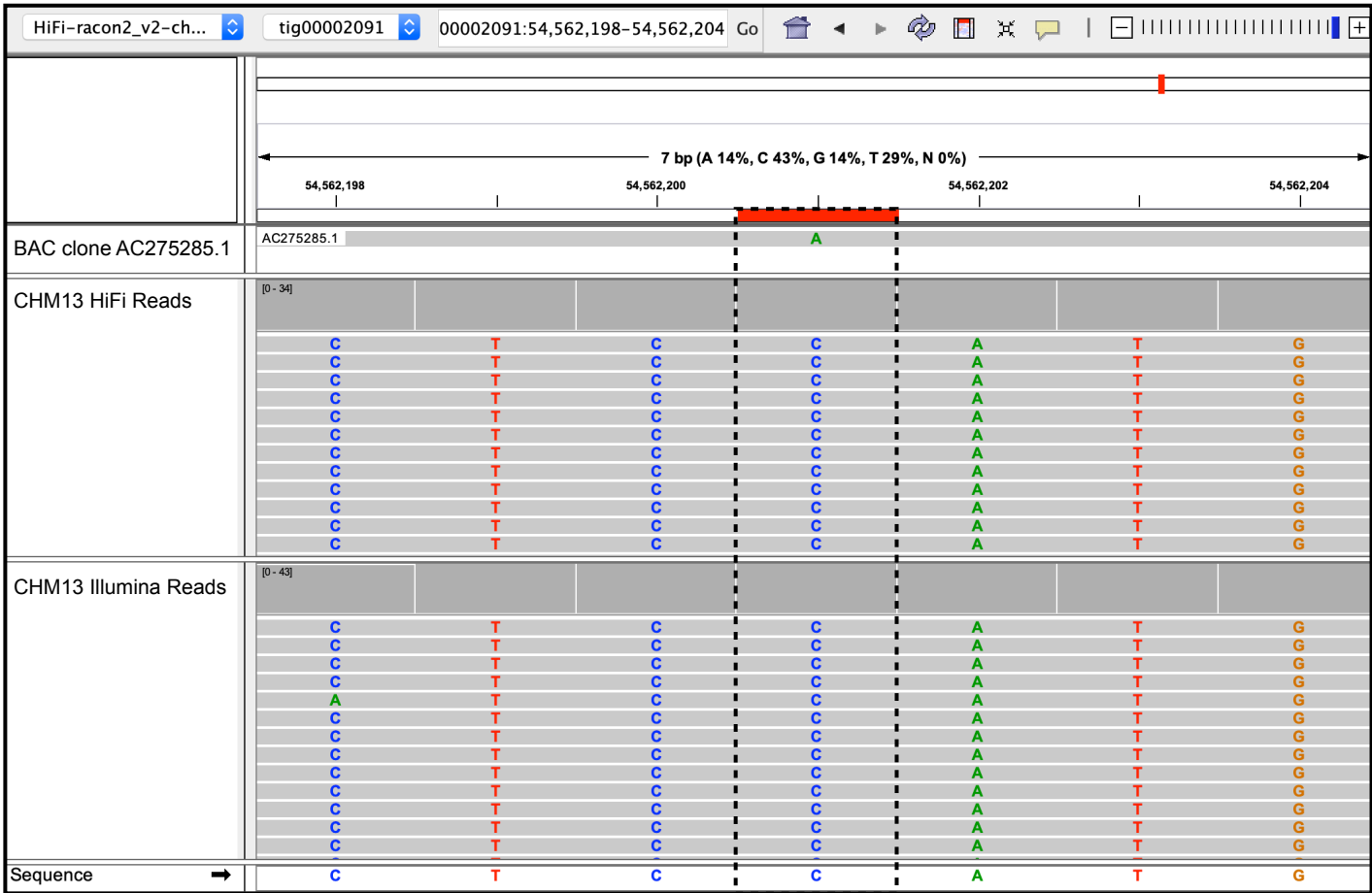
C

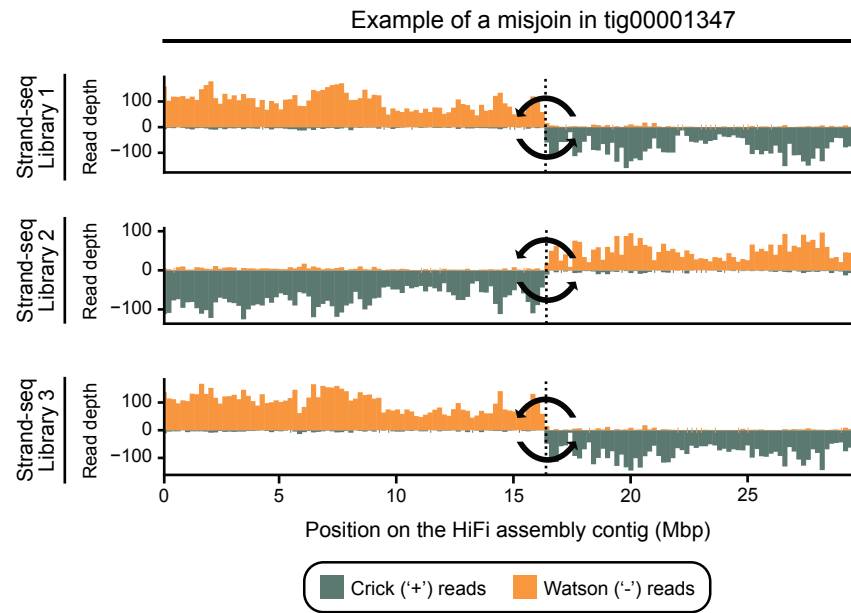
	Read QV (% accuracy) including indels		Read QV (% accuracy) excluding indels	
	Median	Mean	Median	Mean
HiFi	30.5 (99.91%)	23.4 (99.54%)	No error (100.0%)	29.6 (99.89%)
CLR	9.1 (87.67%)	8.8 (86.78%)	16.1 (97.53%)	15.4 (97.10%)

Vollger, Logsdon et al., Figure S2



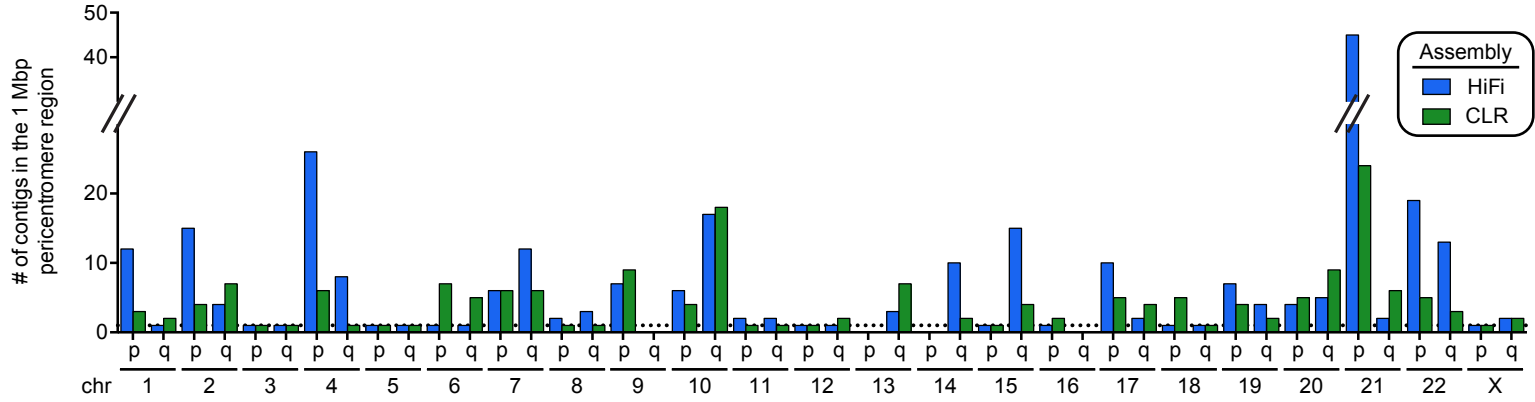
Vollger, Logsdon et al., Figure S3



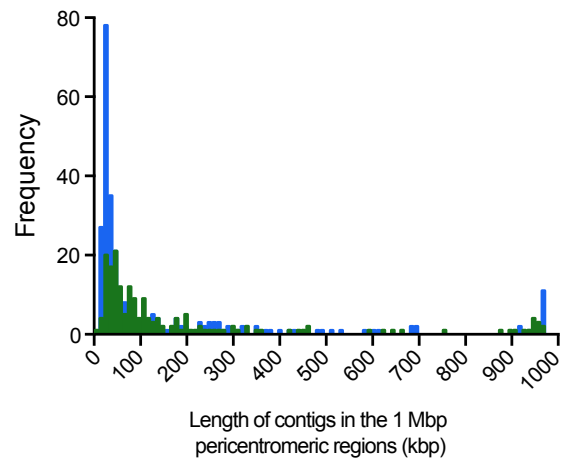


Vollger, Logsdon et al., Figure S5

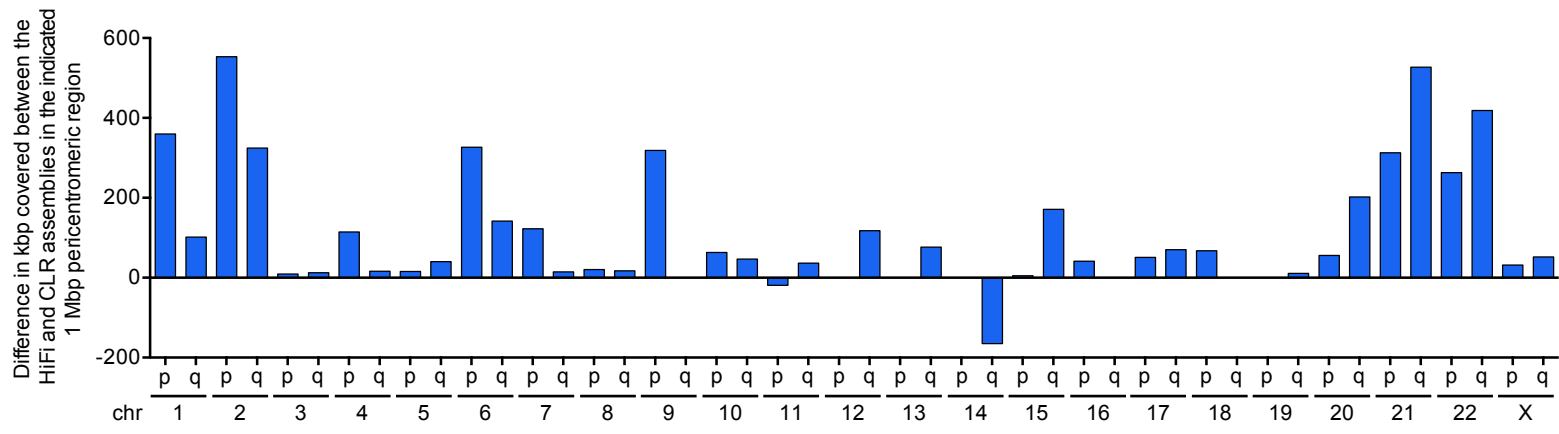
A



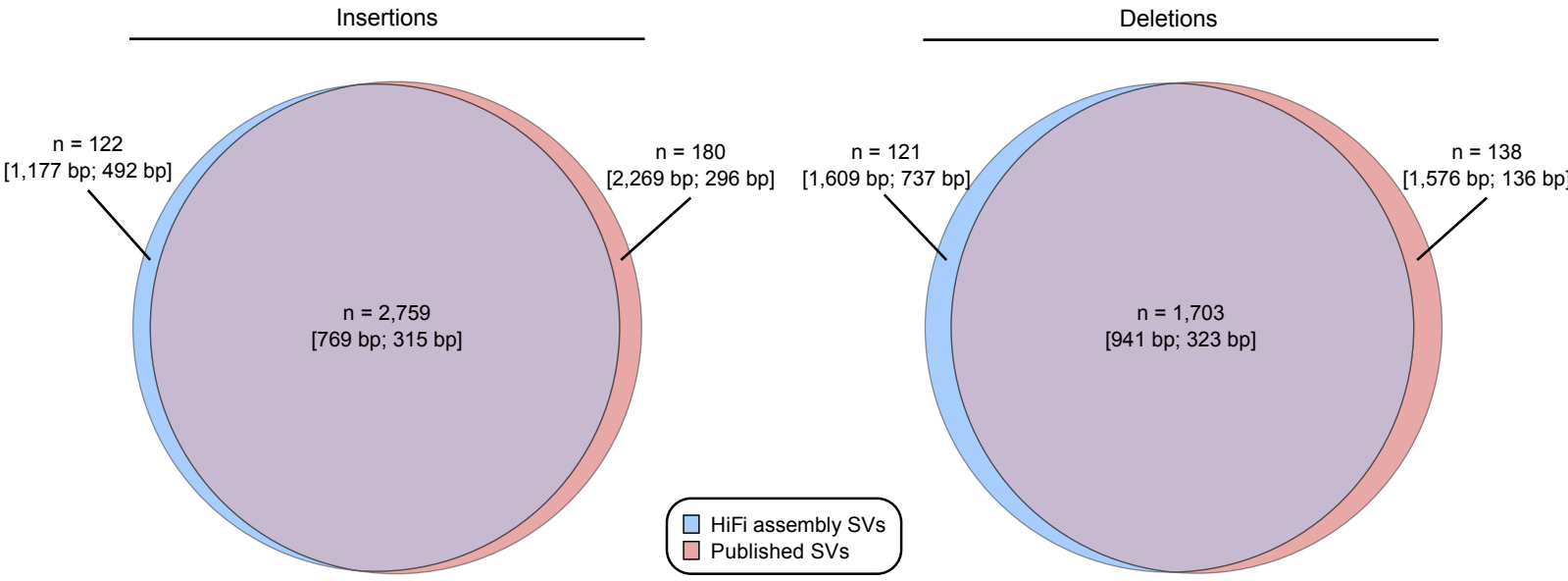
B



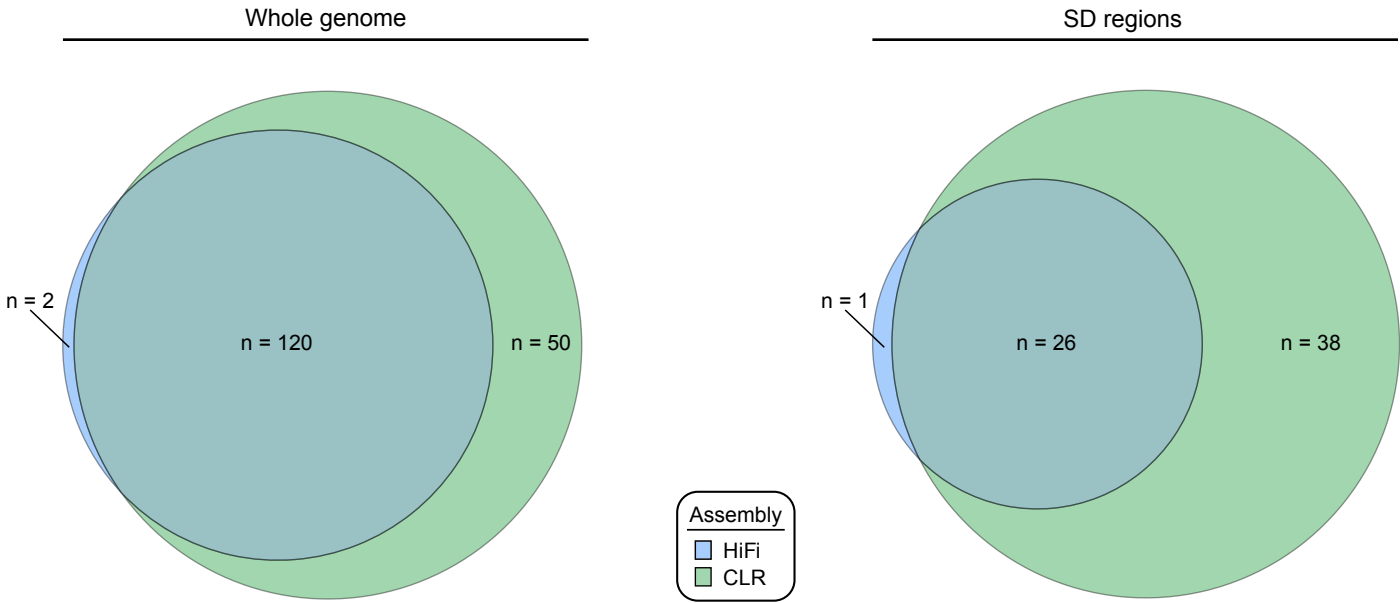
C



HiFi assembly SVs vs. published SVs



Likely gene-disrupting events in HiFi and CLR assemblies



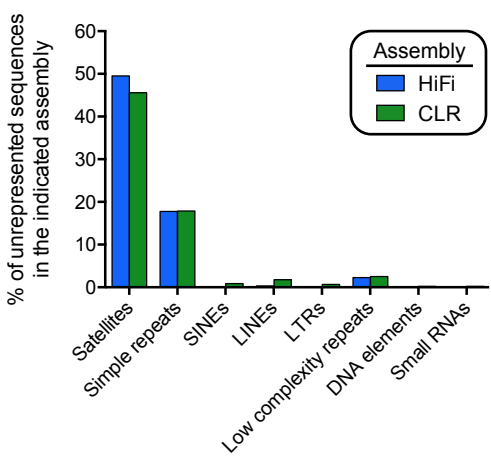


Table S1. Statistics of the HiFi and CLR genome assemblies produced by different assemblers

	Polisher	Total size (Gbp)	N50 (Mbp)	No. of contigs	Median QV	No. of CPU hours for assembly
HiFi (24-fold)						
<i>CHM13 genome</i>						
<i>Canu assembly</i>						
	None	3.03	25.51	5,296	40.41	~2,800
	Arrow	3.03	25.51	5,296	43.29	~10,000
	Racon	3.03	25.51	5,296	44.95	~2,950
	2x Racon	3.03	25.51	5,296	45.25	~3,100
	2x Racon+	3.03	25.51	5,296	45.25	~4,200
HiFi (24-fold)						
<i>CHM13 genome</i>						
<i>FALCON assembly</i>						
	None	3.00	31.91	2,115	27.65	~4,400
	Arrow	3.00	31.92	2,115	43.45	~12,400
	2x Racon	3.00	31.92	2,115	44.45	~4,700
CLR (77-fold)						
<i>CHM13 genome</i>						
<i>FALCON assembly</i>						
	None	2.88	29.26	1,916	27.49	>50,000
	Quiver	2.88	29.26	1,916	40.73	>55,000
	Quiver+	2.88	29.26	1,916	42.70	>55,000
CLR (24-fold)						
<i>CHM13 genome</i>						
<i>Canu assembly</i>						
	None	2.48	0.31	21,167	26.50	~37,300

HiFi: HiFi assembly

CLR: CLR assembly

2x Racon: Two rounds of Racon

2x Racon+: Two rounds of Racon and one round of Pilon

Quiver+: Quiver, Pilon, and FreeBayes-based indel correction

Median QV: Median QV over 31 BACs

Table S2. False joins identified by Strand-seq within *de novo* assembly contigs

	Contig	Start coordinate	End coordinate	No. of false joins
HiFi				
<i>CHM13 genome</i>				
<i>Canu assembly</i>				
	tig00000017	99,602,681	99,618,276	1
	tig00001347	16,388,858	16,426,941	1
	tig00003369	22,812,690	22,841,943	1
	tig00002385	3,056,424	3,204,103	1
	tig00002385	10,092,563	10,210,819	1
	tig00001433	2,544,909	2,612,082	1
	tig00004976	785,171	805,928	1
CLR				
<i>CHM13 genome</i>				
<i>FALCON assembly</i>				
	NTIA01000004.1	51,797,117	51,869,304	1
	NTIA01000039.1	16,068,231	16,072,573	1
	NTIA01000061.1	12,087,358	12,166,683	1
	NTIA01000067.1	4,639,556	4,640,945	1
	NTIA01000093.1	746,665	783,795	1

HiFi: HiFi assembly
CLR: CLR assembly

Table S3. Comparison of PSVs linked with SDA in HiFi and CLR assemblies

Locus	Expected no. of paralogs	Expected no. of phased bases (kbp)	No. of paralogs by SDA (HiFi, CLR)	No. of bases phased by SDA (HiFi, CLR; kbp)	Average length of SDA-phased block (HiFi, CLR; kbp)	No. of PSVs (HiFi, CLR)
<i>OPN1LW</i>	2	70	0, 1	0, 20	0, 20	0, 14
<i>NOTCH2NL</i>	5	500	9, 4	443, 345	49, 86	728, 482
<i>SRGAP2</i>	4	520	5, 5	493, 494	99, 99	1194, 821
<i>FCGR2/3</i>	3	220	3, 4	141, 140	47, 35	611, 139
<i>KANSL1</i>	2	280	4, 3	54, 150	13, 50	48, 92