Supplementary Materials 1.
Heterochromatin-based assembly extended methods and results

We mapped nanopore reads to the D. pseudoobscura genome using minimap2 (version 2.1-r316-dirty). We then used seqtk seq to extract reads that 1) did not map to euchromatin and/or 2) were unmapped.

Canu assemblies

We ran canu for each of these Y-replacement line male samples with the parameter rawErrorRate=0.300. For the individual lines, we adjusted the genomeSize parameter according to flow cytometry estimates (genomeSize=161m, 166m, and 175m for YS, YM, and YL, respectively).

Falcon assemblies

We re-headed our Nanopore reads to have PacBio-formatted headers using DASCRUBBER. This allowed us to run the Falcon assembly on our sequences. We then ran Falcon (falcon-kit 1.4.2) on the heterochromatin-enriched reads using the following parameters:

pa_daligner_option =-v -e.75 -l1000 -s100 -k14 -h256 -w8
pa_HPCdaligner_option=-e0.75 -l1000 -k18 -h35 -w8 -s100 -t32 -M24
ovlp_daligner_option =-v -h60 -e.95 -l500 -s100 -k24 -M24 -t40
ovlp_HPCdaligner_option=-v -l500 -B128 -t32

falcon_sense_option = --output_multi --min_idt 0.70 --min_cov 4 --max_n_read 200 --n_core 20
overlap_filtering_setting = --max_diff 100 --max_cov 100 --min_cov 1 --bestn 10 --n_core 24

We adjusted the genome_size parameter for Y-replacement sample:

100000 for YL, 80000 for YM, and 70000 for YS.

The Falcon assembler generates both primary and associated contigs. For this study, we only used primary contigs.

Assembly polishing

We polished both Canu and Falcon assemblies with 3 iterations of Racon (version 1.3.1) and then 1 iteration of Pilon (version 1.22) using the nanopore and Illumina data for each respective software. Briefly, we mapped long-reads to the initial Canu/Falcon assembly using minimap2 and then used those alignments for Racon. We repeated this process by using the intermediate assemblies from Racon for a total of 3 rounds. We then ran Pilon using the last assembly from Racon and the raw Illumina reads. Finally, we used BLASTN to align the polished contigs against the entire NCBI database and removed contigs with hits to bacteria. **Table S1** shows the assembly statistics for our intermediate assemblies.

Table S1. Assembly statistics of heterochromatin assemblies from Canu and Falcon.

	N50	Max scaffold size	Genome size	No. scaffolds
YS (Canu)	353564	8615011	69011105	485
YM (Canu)	241676	1288246	70418057	552
YL (Canu)	390971	5565708	95514135	522
YS (Falcon)	78606	1431440	50222737	949
YM (Falcon)	134122	1065890	50509490	810
YL (Falcon)	366010	3822829	112450760	624

To check if we assembled most of the Y's, we first mapped male and female reads back to the heterochromatin-enriched assemblies. We then called Y chromosome contigs where Log2(female/male) < -1 and male coverage >= 5x. We then used the median coverage of the rest of the heterochromatin assembly to estimate the copy number of the Y contigs [**Table S2**].

Table S2. Sizes (bp) of intermediate Canu and Falcon Y assemblies after adjusting for copy number (normalized by remaining heterochromatin assembly)

	YS	YM	YL
Canu	34,266,923	53,632,303	65,167,484
Falcon	35,714,894	48,172,177	63,399,543

Quickmerge of heterochromatin-only assemblies and identification of Y contigs

We then used QuickMerge {Chakraborty:2016ji} on the heterochromatin-enriched assemblies from these two softwares to more closely recapitulate the differences in genome size. We ran QuickMerge for these assemblies in reciprocal directions and again on their results. For each Y assembly, we chose the merged assembly of a merged Canu to Falcon and a merged Falcon to Canu (direction: CFFC). These assemblies had the fewest number of scaffolds and high N50s for each Y chromosome [**Table S3**].

Table S3. Assembly statistics of final QuickMerge heterochromatin assemblies (direction: CFFC)

	N50	Max scaffold size	Genome Size	No. scaffolds
YS	479854	8624421	69275244	441
YM	504589	2510978	70070409	425
YL	1021014	10971662	104766850	334

We then used similar cutoffs to identify Y chromosome contigs on these QuickMerged assemblies [**Table S4**]. **Figure 1** details the entire pipeline. We identified 2 misassemblies in YM and 1 misassembly in YS that we manually split [**Figure 2**]. We confirmed that the overall size differences between Y's was consistent with our Illumina and dnaPipeTE estimates by mapping WGS reads to Y contigs and the masked female genome [**Table S5**]. This method is analogous to the manner in which we originally estimated TE abundance from Illumina data using the TE library.

Table S4. Size estimation of Ys from QuickMerged assemblies. Size is calculated by normalizing by the remaining heterochromatin assembly.

	YS assembly size (Mb)	YM assembly size (Mb)	YL assembly size (Mb)
Mapping to assembly	36.454	48.818	56.438
Difference to YS		12.363	19.983

Table S5. Confirmation of Y sizes from QuickMerged assemblies using mappings to full male genomes. Size (Mb) is calculated by normalizing by the autosomes/X chromosome.

	YS assembly size	YM assembly size	YL assembly size
Mapping to male genome	38.655	43.917	59.526
Difference to YS		5.261	20.871

For the remainder of the study, we decided to use our QuickMerged assemblies because of better assembly contiguity (Table S9 in Supplementary Tables). We also used the copy number estimates from mappings to the whole genome because differences between the Ys aligned better to differences from flow cytometry and paired-end data analyses (Table 1 in Main Tables).

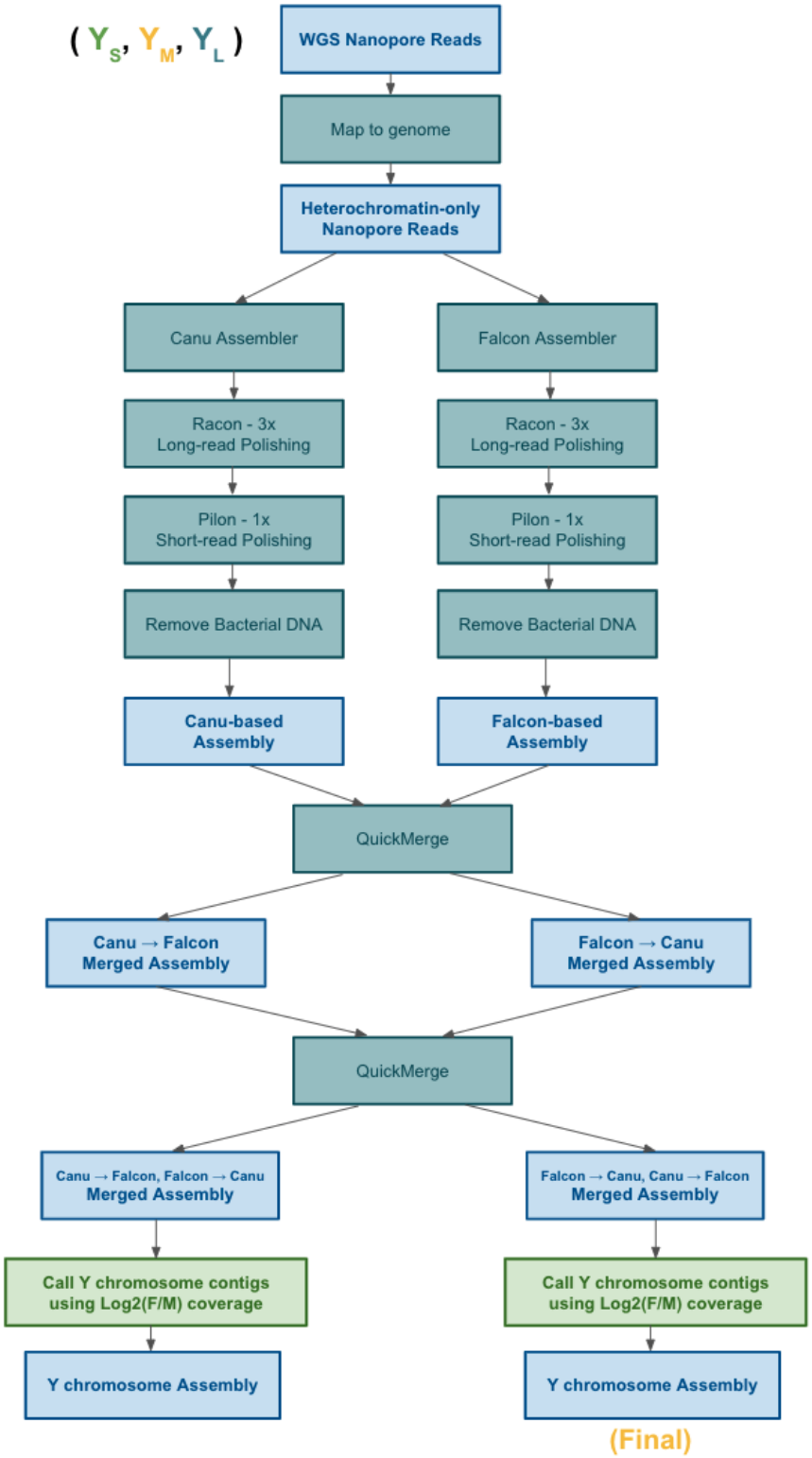

Figure 1. Diagram of Y chromosome assembly pipeline.

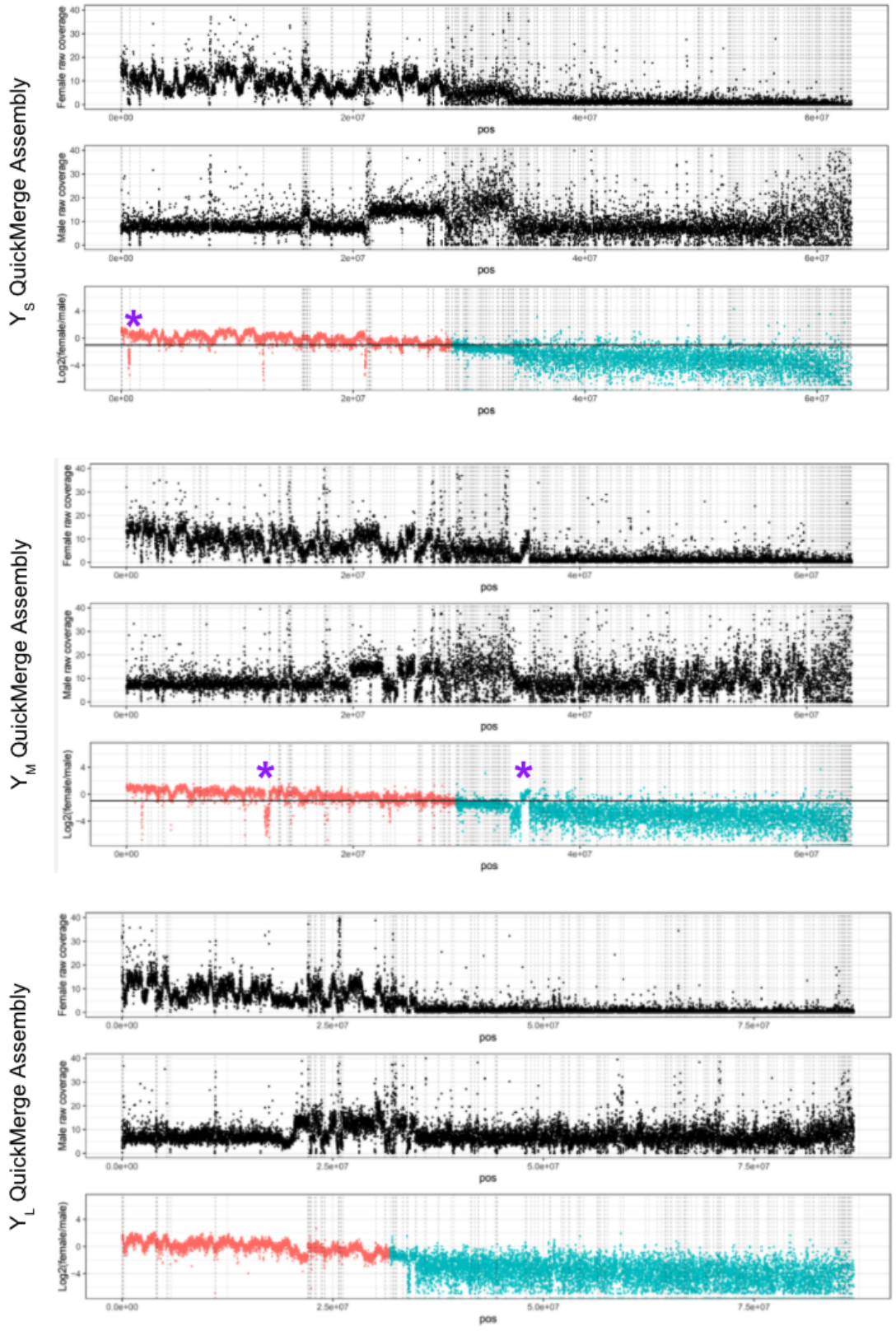

Figure 2. Female and male coverage of final heterochromatin assemblies with Y chromosomes identified. *s denote contigs that were split because of mis-assemblies.