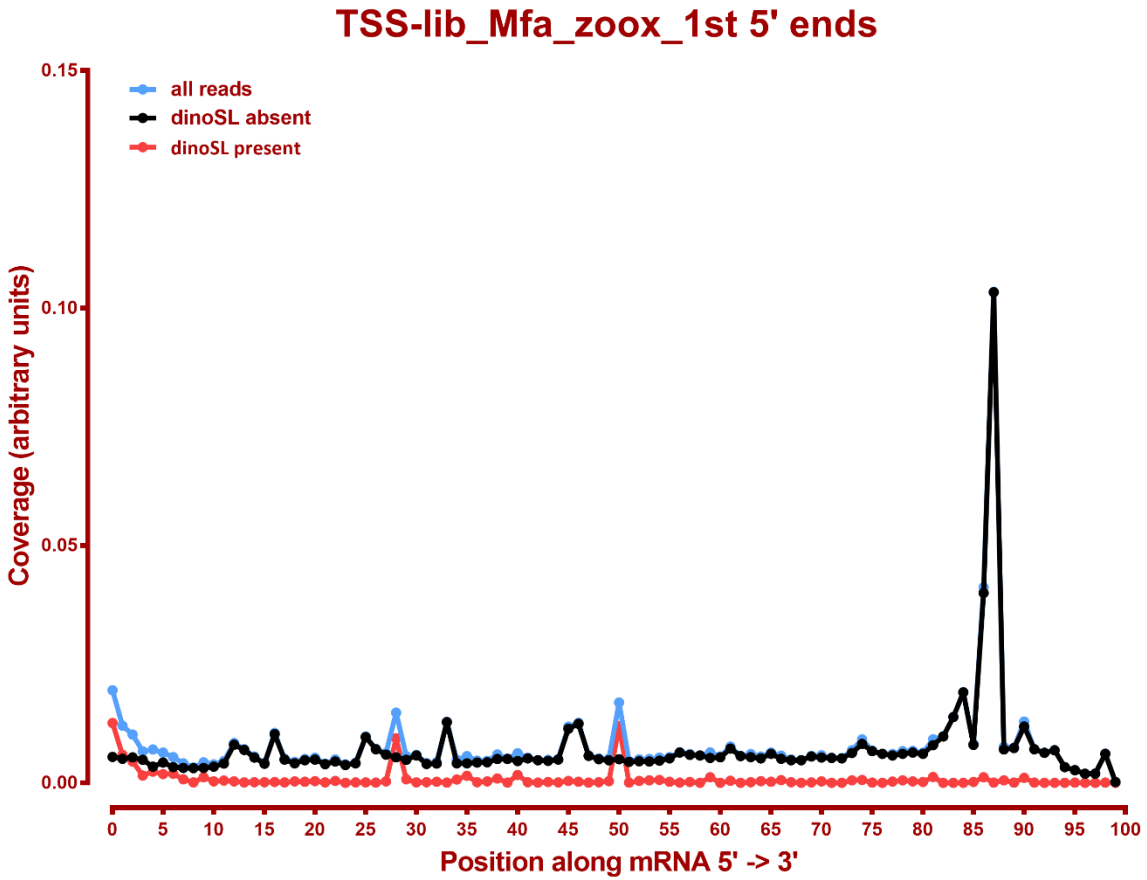
CAGE data mapping stats

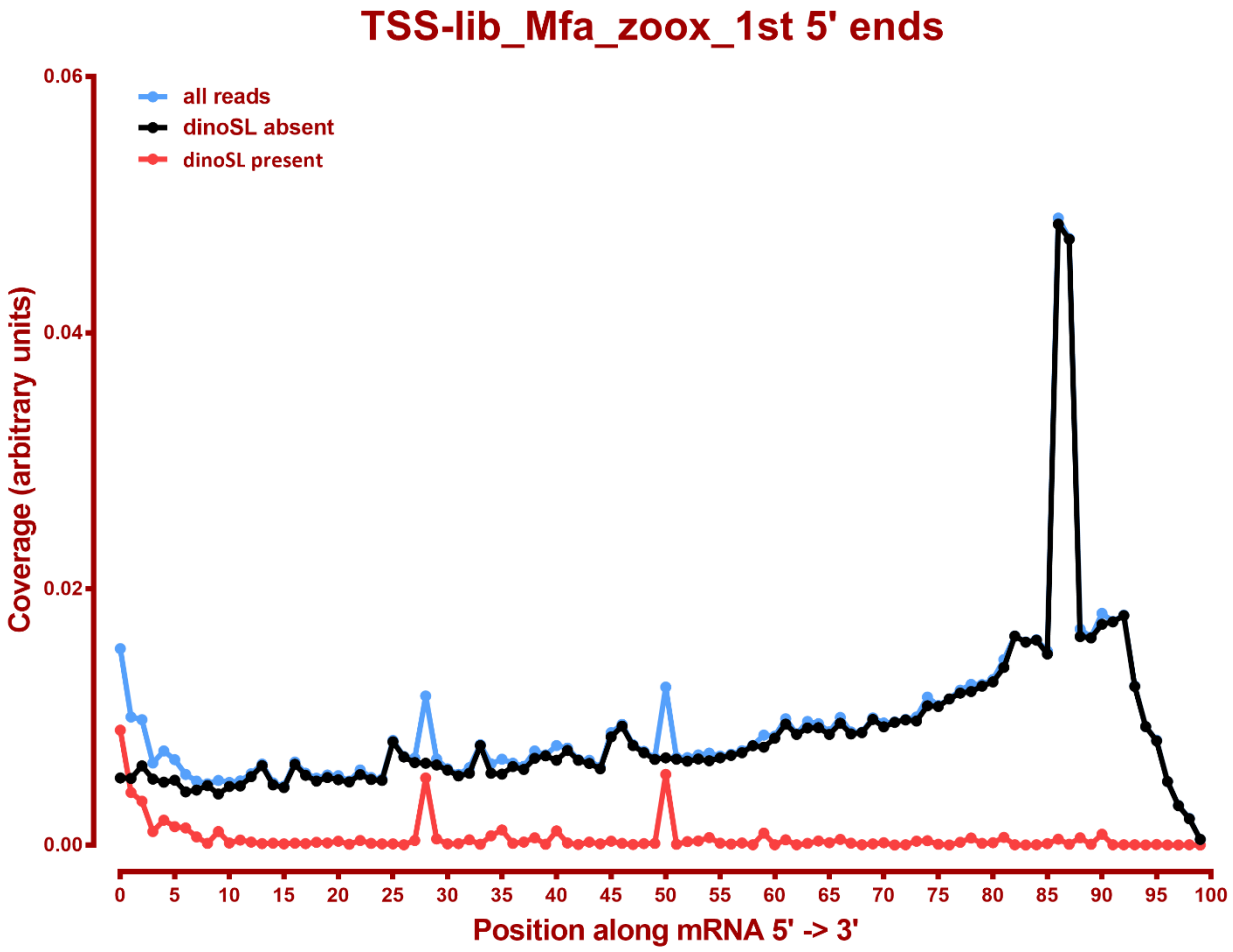
Species	Read Length	Mapping	Library	Raw fragments	Complexity	all reads					dinoSL-containing reads				
						Unique	Unique Splices	Multi	Multi Splices	Fraction mapped	Unique	Unique Splices	Multi	Multi Splices	Fraction of mapped reads
Symbiodinium_minutum_Clade_B1_Ver1.120123	**dinoSL-trimmed**	STAR-2.5.3.a	TSS-lib_Mfa_zoox_1st	60,003,621	0.03	30,990,749	16,874,672	4,285,003	2,153,971	0.45	4,815,868	3,808,760	215,485	494,018	0.17
Symbiodinium_minutum_Clade_B1_Ver1.120123	**dinoSL-trimmed**	STAR-2.5.3.a	TSS-lib_Mfa_zoox_2nd	219,915,043	0.03	86,176,976	77,415,406	14,119,397	7,505,162	0.42	10,598,950	8,389,561	541,480	876,380	0.11

Read distribution

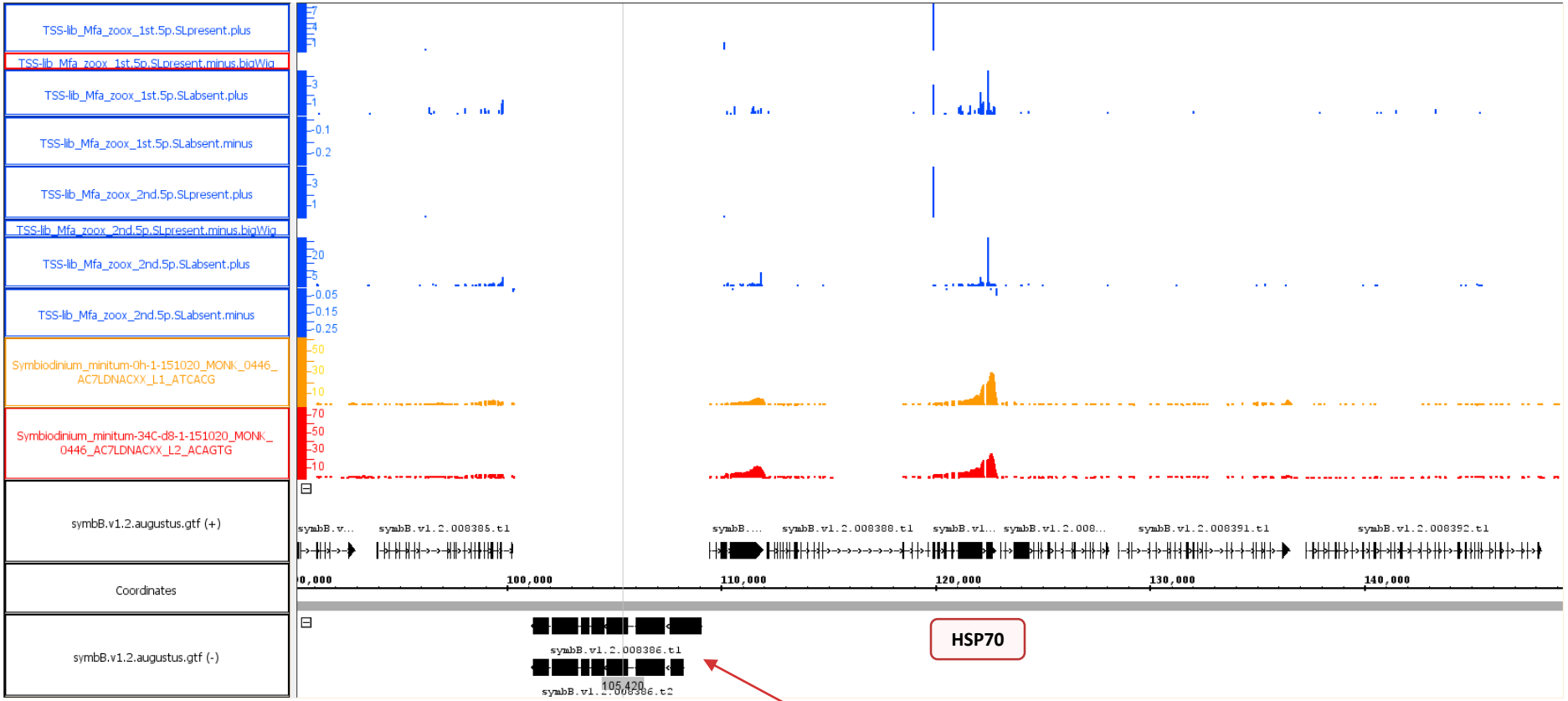
#	Exonic:	Intergenic:	Intronic:
TSS-lib_Mfa_zoox_1st.SL-only	0.60	0.25	0.14
TSS-lib_Mfa_zoox_1st	0.69	0.25	0.06
TSS-lib_Mfa_zoox_2nd.SL-only	0.62	0.23	0.14
TSS-lib_Mfa_zoox_2nd	0.77	0.15	0.09

Read distribution



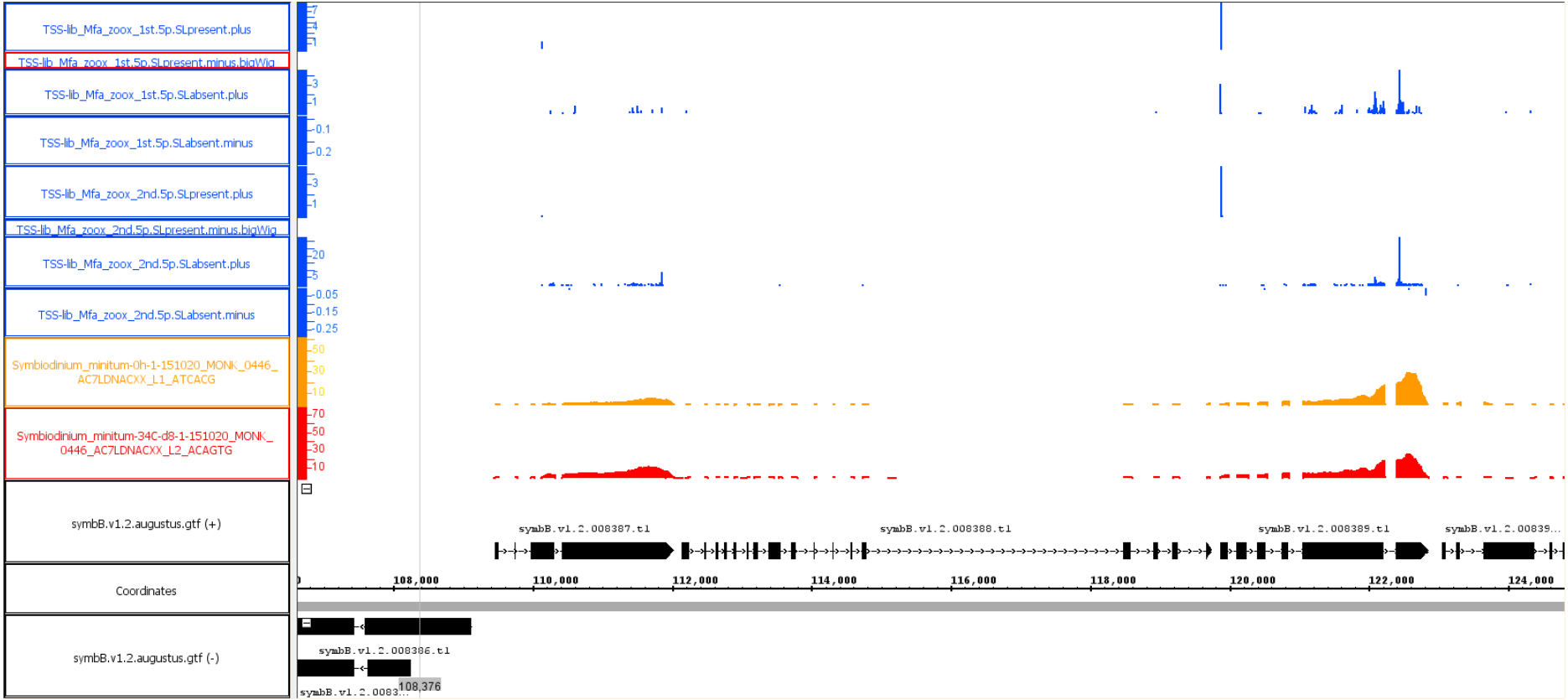

TSS-lib_Mfa_zoox_1st 5' ends
all reads
dinoSL absent
dinoSL present
Coverage (arbitrary units)
0.06
0.04
0.02
0.00
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
Position along mRNA 5' -> 3'

HSP70 context:



This is a reverse transcriptase gene, i.e. a repeat

HSP70 zoom-in



HSP70 summary

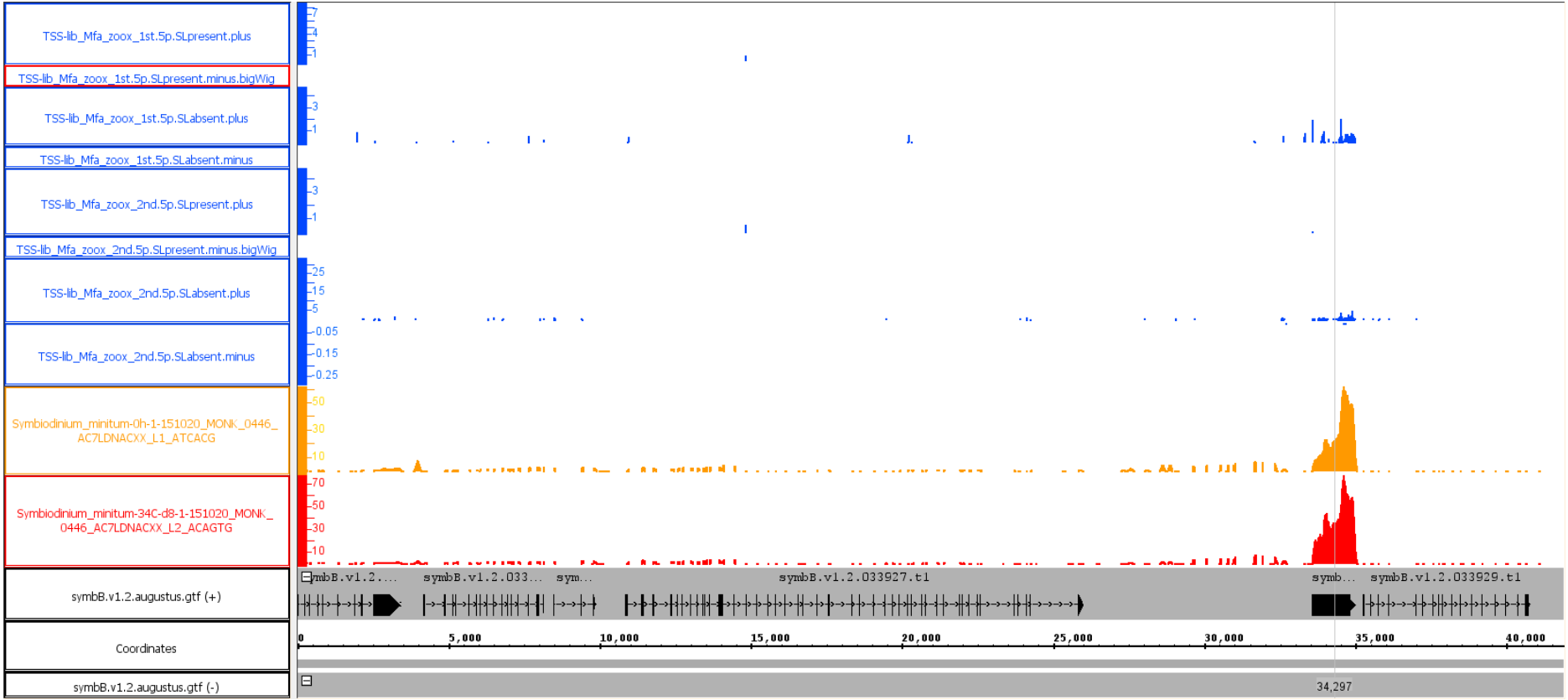
- This HSP70 copy appears to be in a multicistronic array, with transcription beginning fro mthe upstream genes

- There is a clear dinoSL CAGE peak at the HSP70 beginning

- There is no clear such peak at the beginning of the upstream gene though that gene is expressed

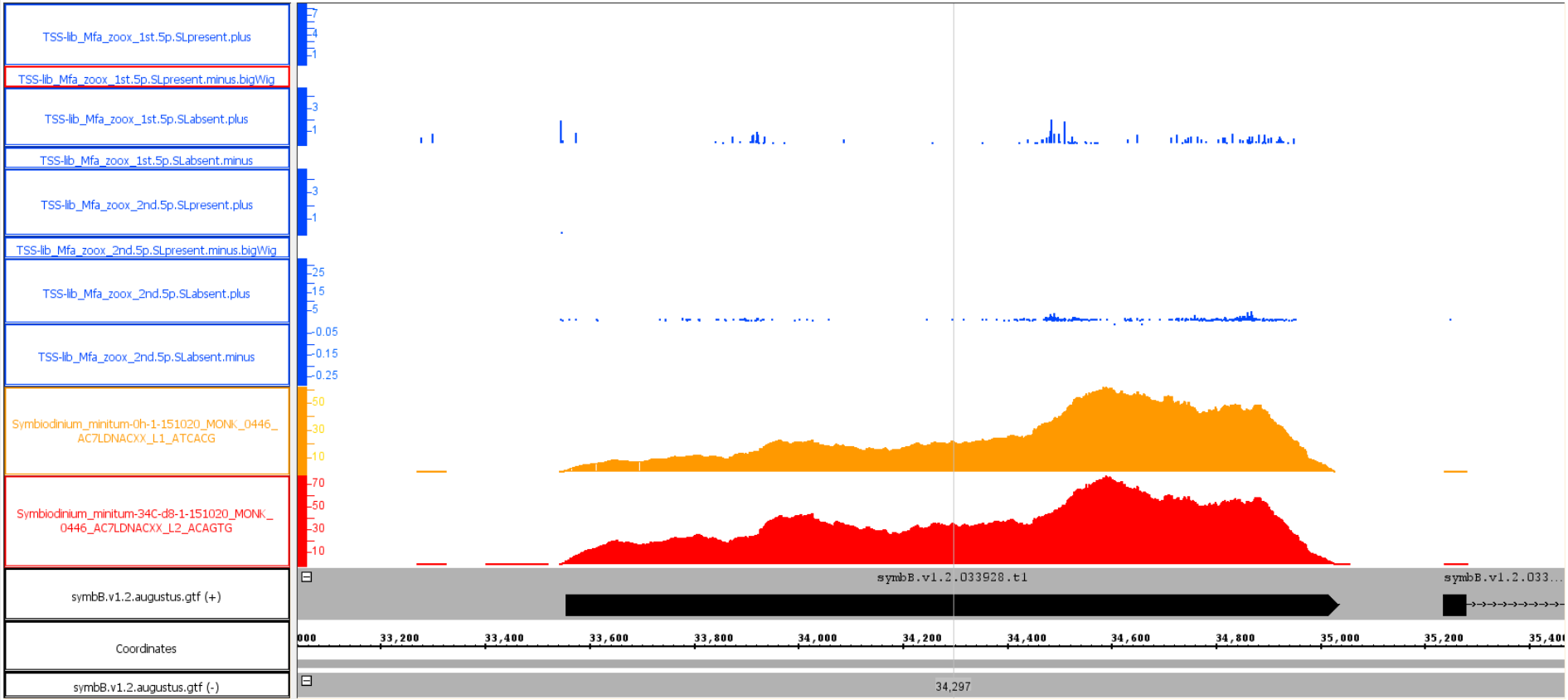
- Note, however, that the mapping here was done with default STAR settings, i.e. multimappers were limited to 20 places in the genome, and because of the big repeat that separate this array from the one to the left of it, this is probably why there is no apparent CAGE peak there. I need to realign with unlimited multimappers to check that.

GAPDH context:



GAPDH

GAPDH zoom-in



GAPDH summary

- This GAPDH copy is on a very short contig, so it's not clear what follows after it, but it does appear to be the first gene in the array

- There is no dinoSL CAGE peak (and no strong CAGE peak without a dinoSL) at the beginning of the gene, but this is not a case of a repetitive element causing alignment