Figure S1. Sequence differences between ST and SR in transcripts. There is no relationship between differential expression between ST and SR and sequence differences between ST and SR. Percent sequence difference is calculated as the total number of differences divided by the length of the transcript times 100. Transcripts with significant differential expression are marked in red. All transcripts meeting the minimum coverage criteria for detecting sequence differences are shown here, including those with no sequence differences.

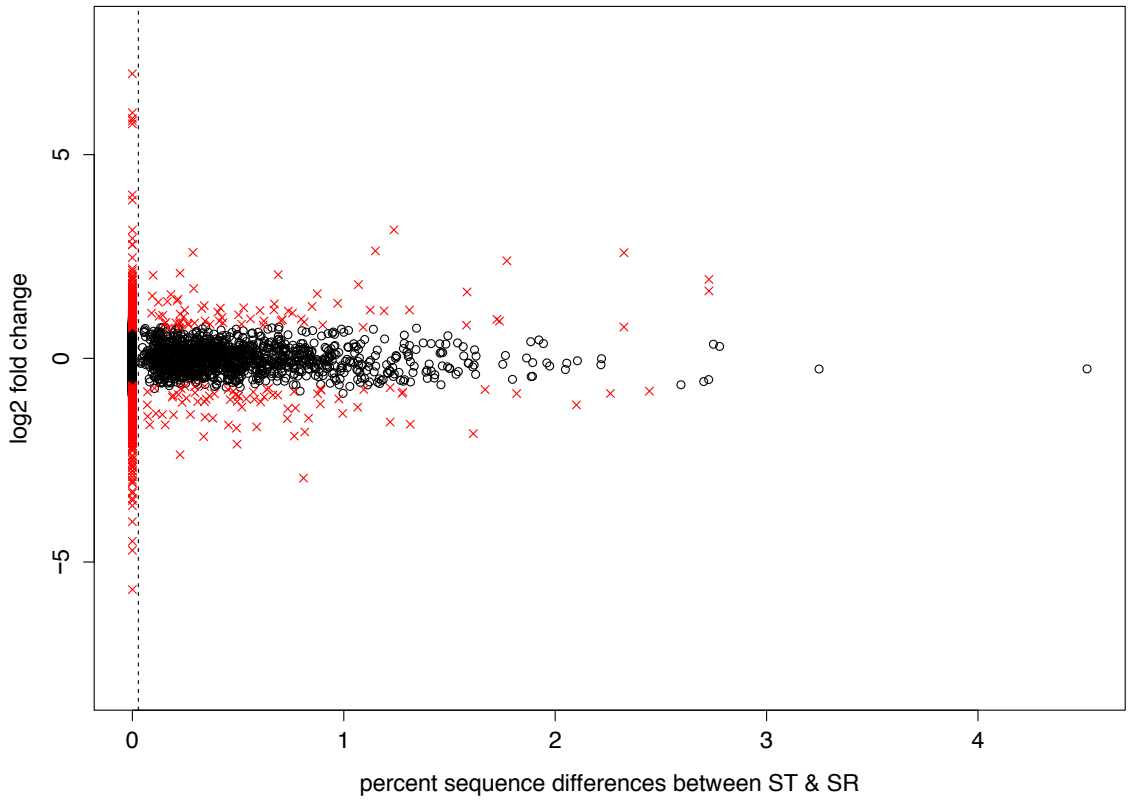


Figure S2. The transcripts that make up *X-importin-α*2. The *D. neotestacea* transcripts have been aligned to the *D. melanogaster importin-α*2 sequence (Pen-PA), which is covered by the red CDS annotation arrow. TR7043 is the autosomal copy of *importin-α*2. TR37105, TR2814, and TR10603 from the combined transcriptome assembly make up the majority of *X-importin-α*2. The end of the transcript has been extracted from both an SR only transcriptome assembly and an ST only transcriptome assembly. A single large transcript covering the end of the transcript and the 3' UTR in the SR only assembly, and two overlapping transcripts covering the same distance were pulled from the ST only assembly.

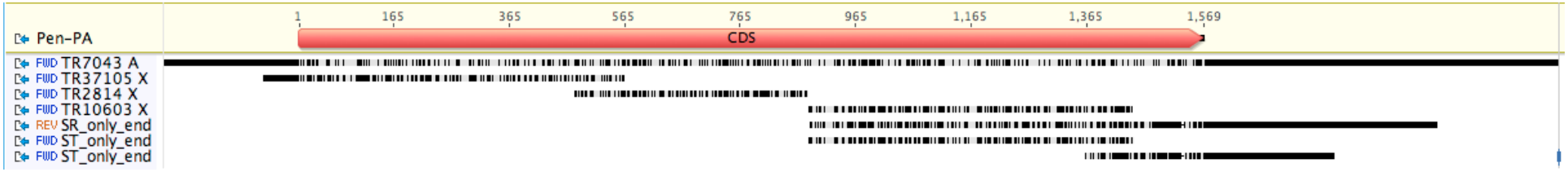
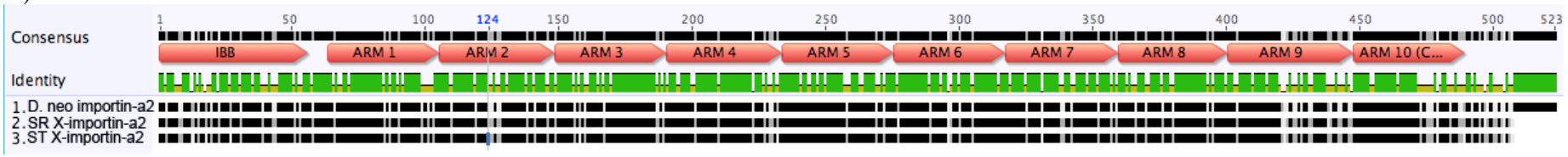


Figure S3. An alignment of the protein sequences of *importin-α2* and *X-importin-α2* in SR and ST. The protein domains are labeled, and the amino acids are coding according to similarity. Black represents identical residues, while white residues are the most different. The C-terminal end of the protein appears to be the most diverged between the autosomal and X-linked copies. The IBB domain binds to *importin-ß*, ARM repeats 1 – 9 bind to the NLS signals of cargo, and ARM repeat 10 binds to the export factor. Panel A shows the entire protein sequence, and Panel B shows the protein sequence alignment zoomed in on the diverged C-terminus. The * denotes the stop codon.

A)



B)

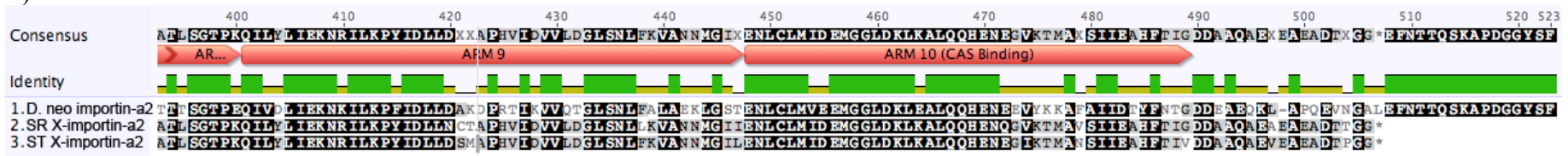
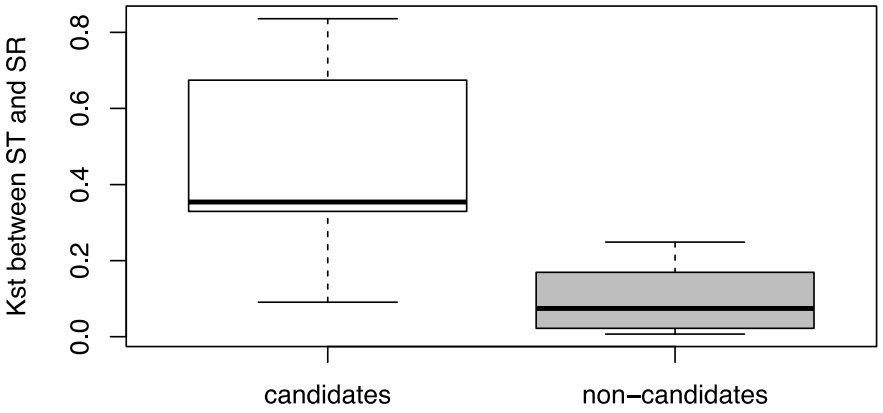


Figure S4. Box plot of divergence between SR and ST across loci between the top candidates and the non-candidate X-linked loci. Panel A shows K_{ST} and Panel B shows S_{nn}.

A)



B)

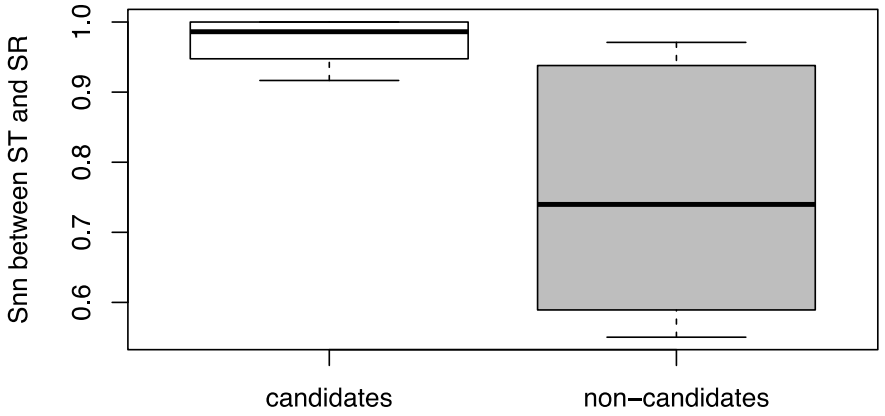


Figure S5. The segregating sites of the sequenced transcripts. Each row is a sequence from a wild-caught male, and is labelled SR or ST. Each column is a polymorphic site. Dark grey indicates the major allele, light grey is the minor allele, and white is the second minor allele when present. Sites are labeled S for synonymous, N for nonsynonymous, I for intron, U for UTR, and G for gap. Each panel is from a different locus, including A) TR261, B) TR11103, C) TR24932, D) TR37304, E) TR23125, and F) TR50351.

A) TR261 – top candidate

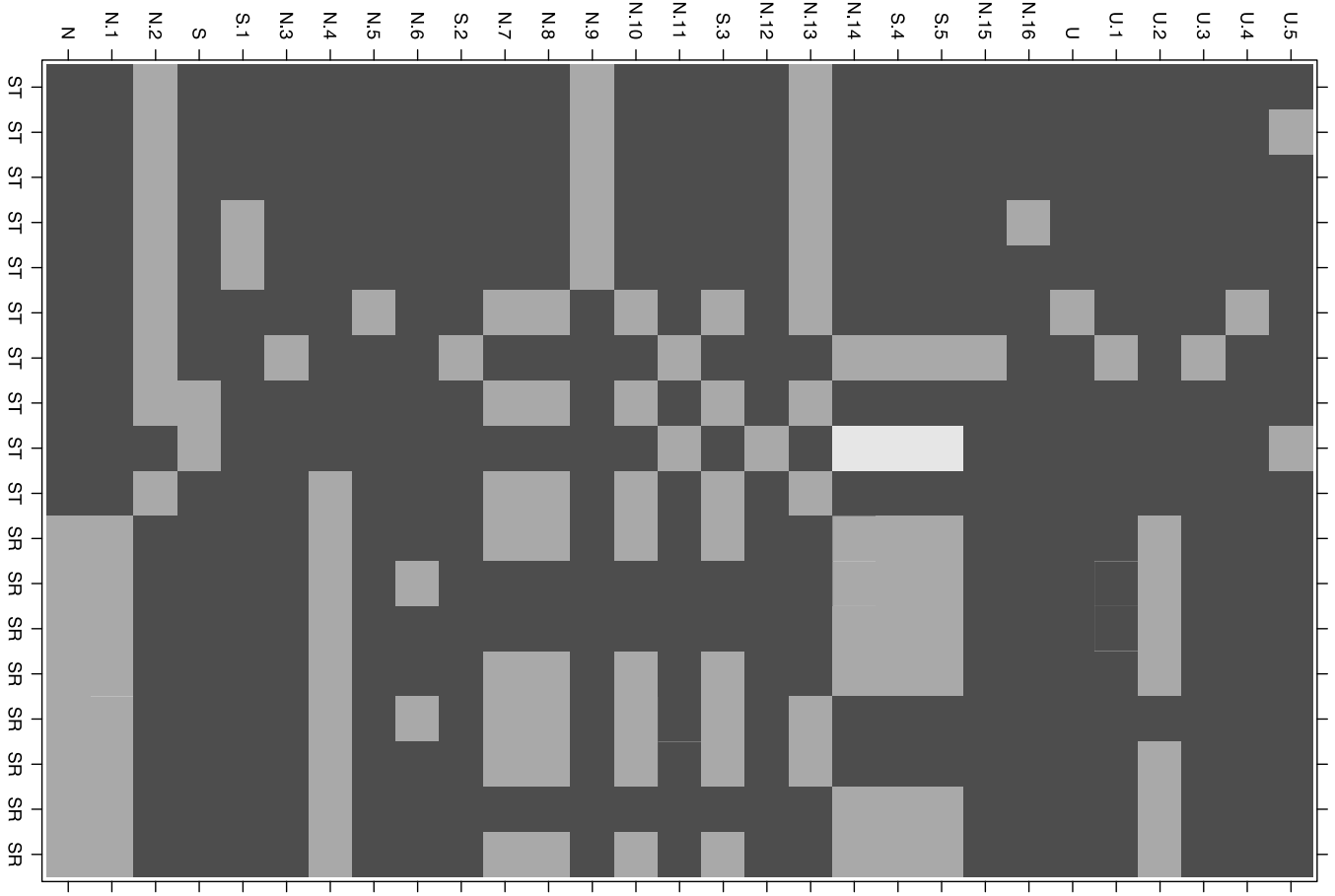


Figure S5. Continued.

B) TR11103 – top candidate

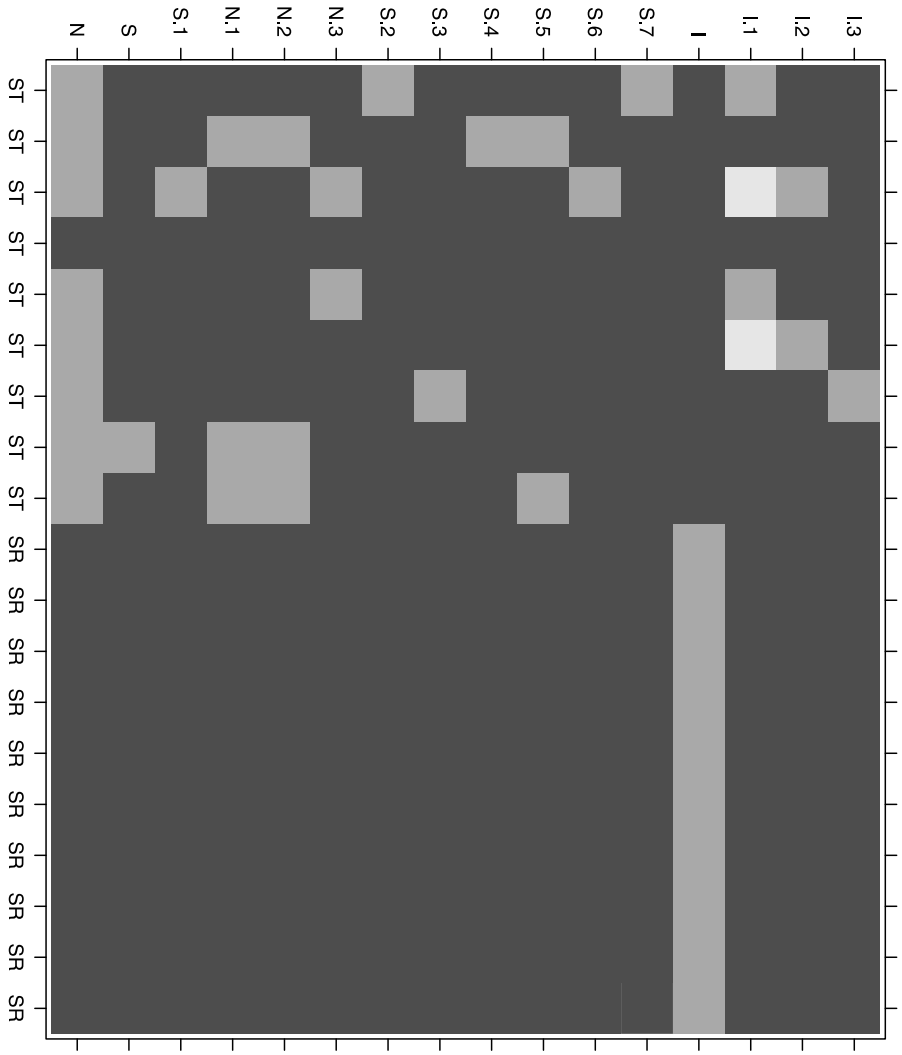


Figure S5. Continued

C) TR24932 – top candidate

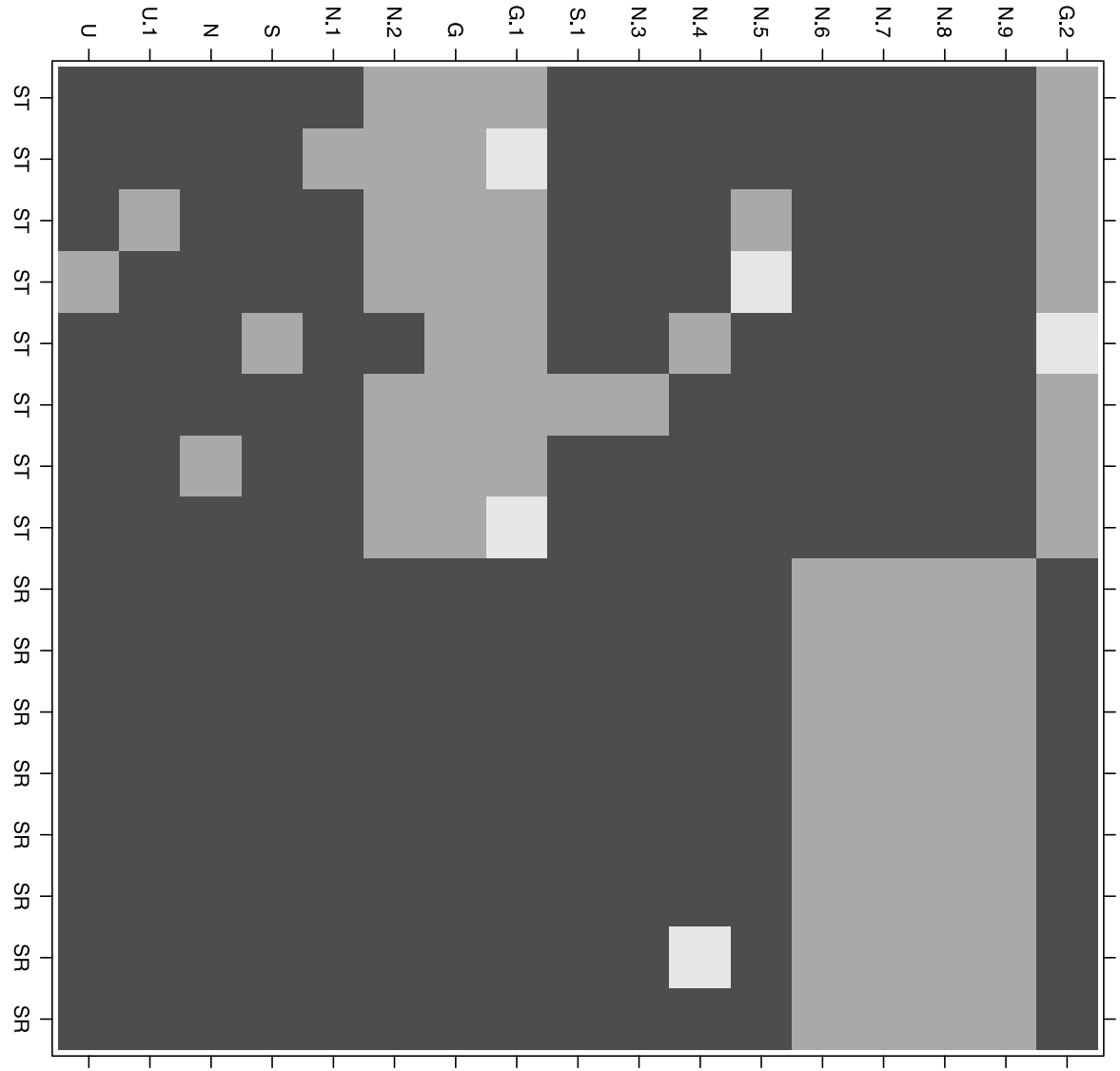


Figure S5. Continued

D) TR37304 – top candidate

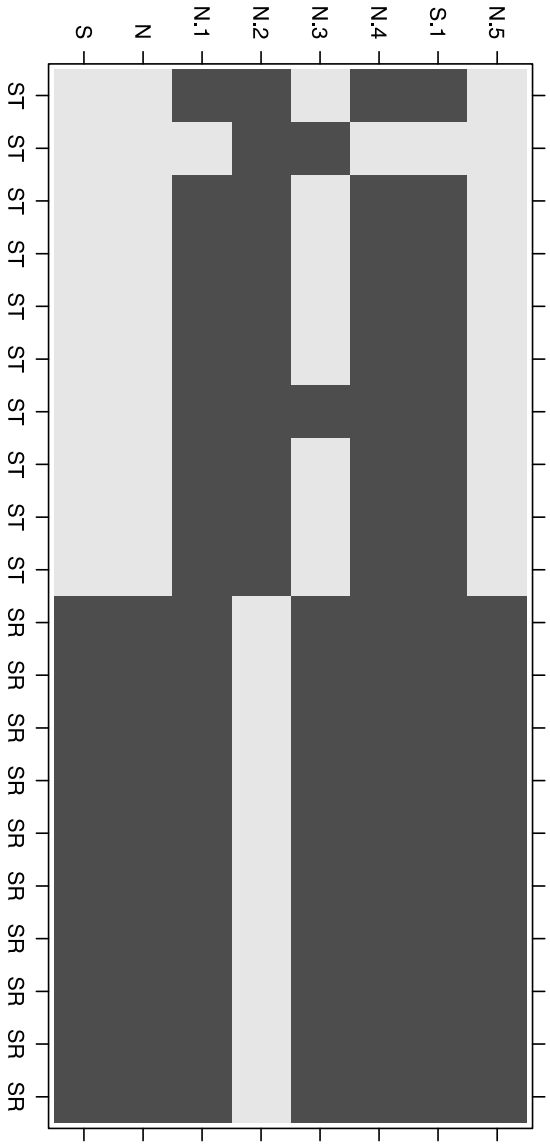


Figure S5. Continued

E) TR23125

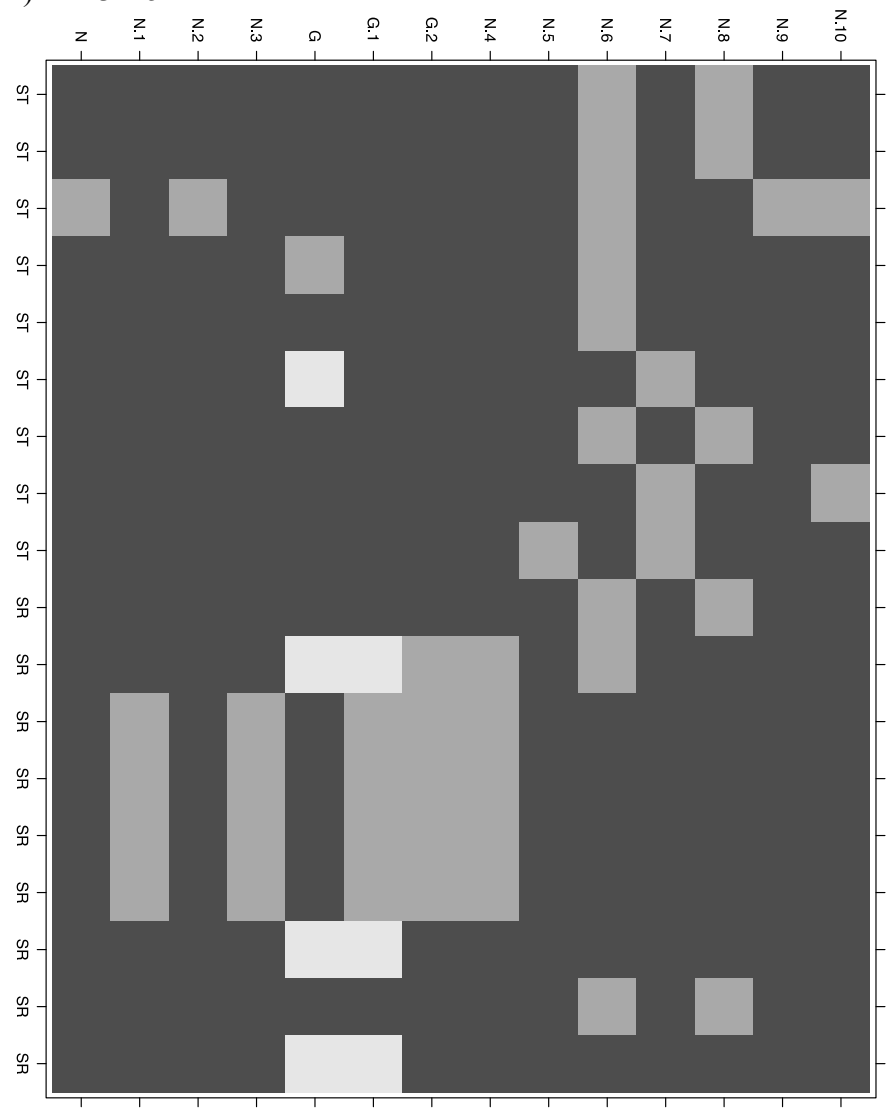


Figure S5. Continued

F) TR50351

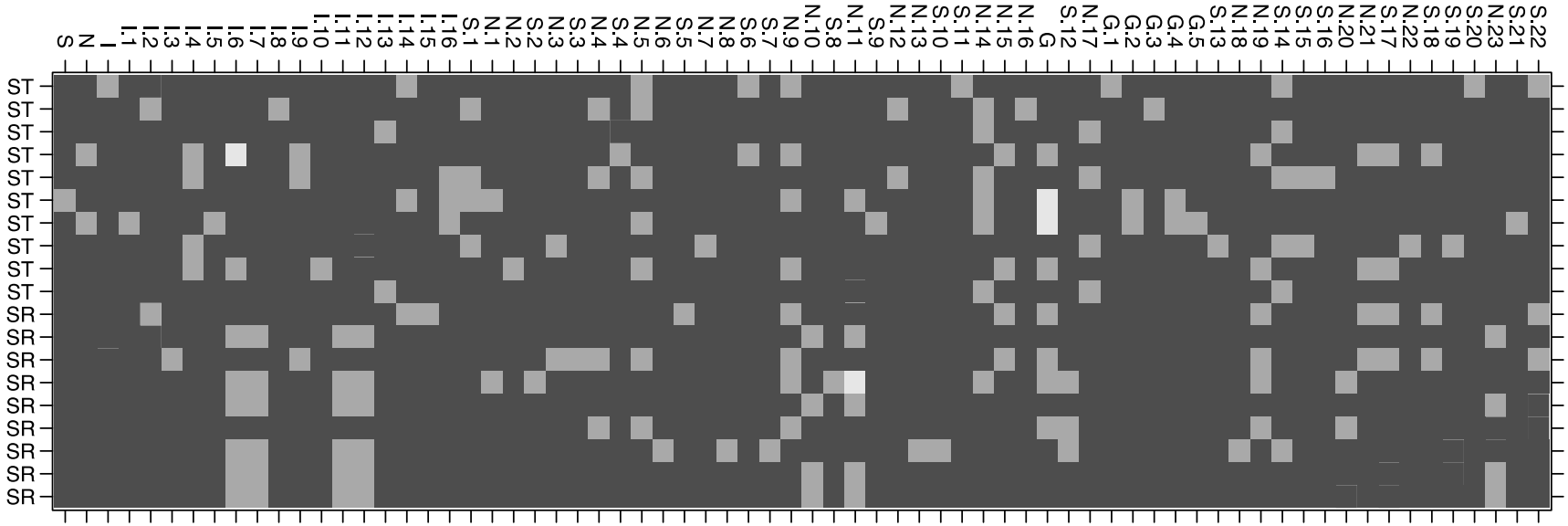


Figure S6. Part of the protein sequence of TR 261, showing a premature stop codon (*) at residue 152 of the ST sample R12. The normal end of the transcript is at site 170.

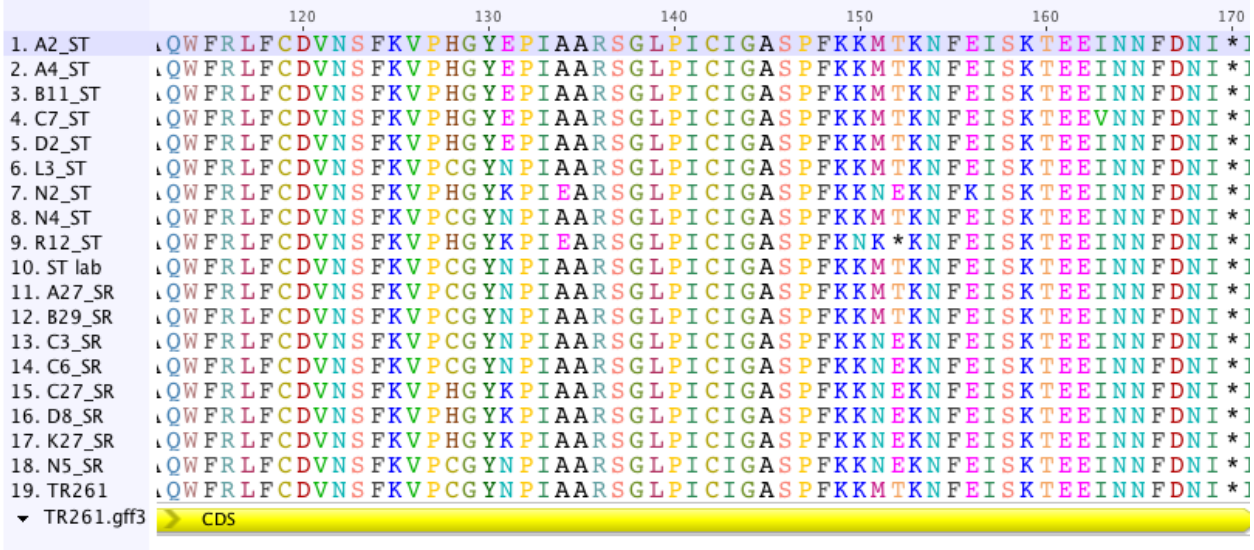
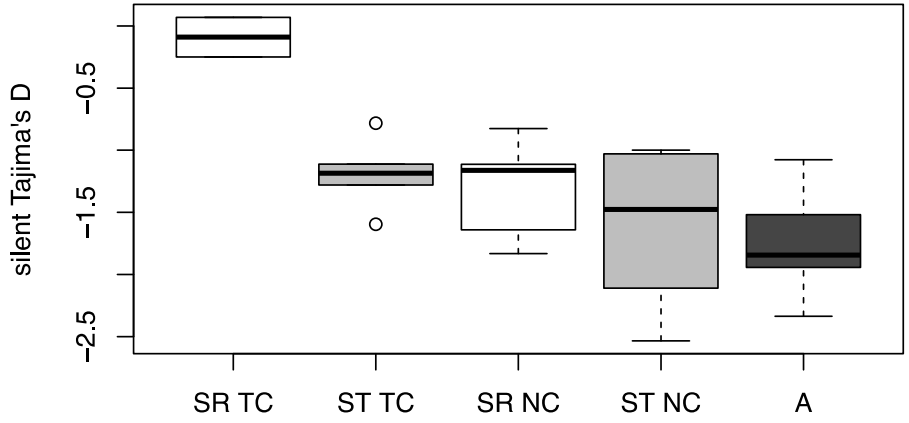


Figure S7. Box plot of Tajima's D for the top candidates (TC) and non-candidates (NC) on ST and SR and the autosomes (A). Panel A on the left was calculated using only non-coding and synonymous variation. Panel B on the right was calculated using only nonsynonymous variation.

A)



B)

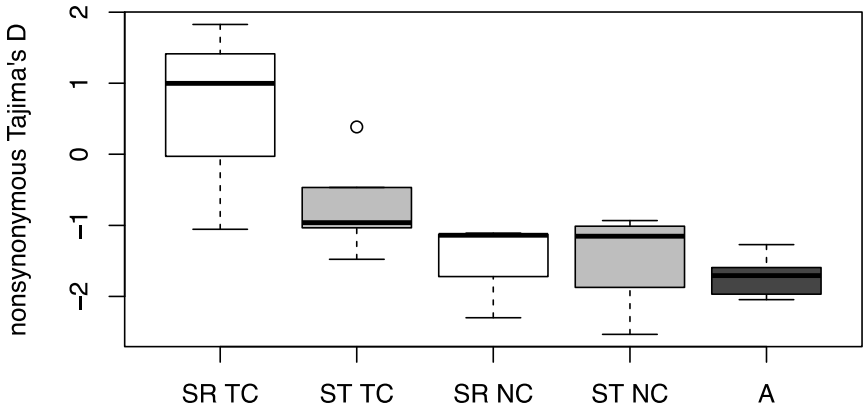


Figure S8. An alignment of the genomic sequence of *importin-α*2 in *D. melanogaster,* the transcript sequence of *X-importin-α*2 in ST *D. neotestacea*, and the genomic sequence of *X-importin-α*2in *D. neotestacea.* The *D. melanogaster* sequence contains the entire open reading frame, from start to stop. Note the *X-importin-α*2 transcript ends short of the *D. melanogaster* transcript. The ORFs of the genomic sequence of *X-importin-α*2 are annotated as yellow arrows along the transcript, and the intron is marked in blue. There are two introns in *D. melanogaster* genomic sequence that are not present in the *D. neotestacea X-importin-α*2 genomic sequence, which are marked with red arrows.

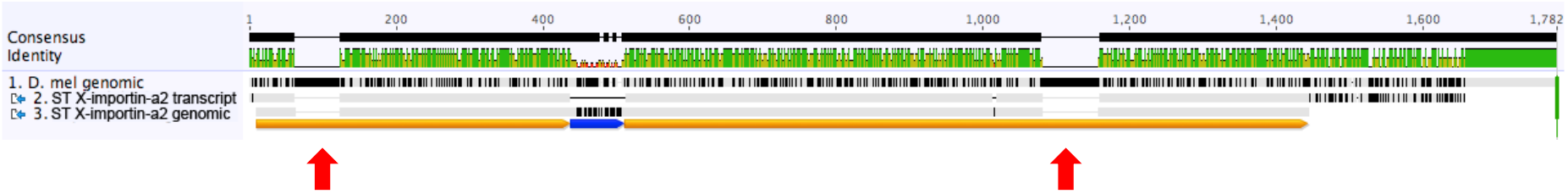


Table S1. The original genomic locations of the transcripts with sequence differences between ST and SR. The locations were assigned via homology with *D. melanogaster* and *D. virilis.* All transcripts with sequence differences were reassigned to the X-chromosome for downstream differential expression analyses.

Chromosome	Total N transcripts	N transcripts with sequence difference SR vs ST	Proportion with a sequence difference SR vs ST
X (A)	2748	1067	0.3883
B	3372	12	0.0036
C	3488	9	0.0026
D	3444	4	0.0012
E	4319	7	0.0016
F	119	0	0
mtDNA	29	0	0
Unknown	7948	250	0.0315

Table S2. Forward and reverse PCR primers.

Transcript	**Forward 5' – 3'**	**Reverse 5' – 3'**
TR261	GCAAATACACGAGCACCAACA	TGTCAGGAGAAGGTGAGATGG
TR1778	ACAAAACCCCTCGTCATCCC	GCGGCTCTTTGCTTTTTGGA
TR5481	TGAGAGCGACGCATTGACTT	GATTTCTGAAGCCCCGACCA
TR6297	CCCATTCCAACTGGCAAAGC	AGCTGACAGAACGTGCAGAA
TR11103	CGATCCCCAGCATCATTCGA	CACCAGTCCAAAGTTGCAGC
TR23125	GACAAGGAAGGCGAACAGGA	TGCTATTGATGCGCTCCGAT
TR24932	CCTTGTCGCCGTGAGTGTTA	TGGGGAACAATTGAAAGCGC
TR37304	CCGGGTGCAATAATTGAGCG	GCGGAAAATGGCACTGGTTT
TR50351	TTGAAGTGGCAATCCCGGAA	GATCAACGTTACTGCGTCGC
X-importin-α2	AACATCAGGAAAGAAAAGCATAACT	CCCAGCTGGTCAGATGTTGT
importin-α2	GCTCGCCCATATACAGCTAAAG	GCTGGGCAGGTGTGTGTTAT

Table S3. Candidates for which population genetic patterns were investigated. The first six transcripts had $K_a/K_s > 1$, were expressed only in the testes, and were differentially expressed between ST and SR. The latter four transcripts had K_a/K_s values suggestive of positive selection but did not otherwise meet the candidacy criteria. K_a/K_s was calculated between ST and SR. Mean expression level, $\log_2$ fold change between ST and SR, and FDR corrected p-value were calculated in DESeq2. Loci selected as top candidates following evolutionary analyses are marked.

Top candidate?	transcript	K_a	K_s	K_a/K_s	Length	Subs	S Subs	N Subs	Testes Mean Expression	log2 Fold Change	padj	% difference	Carcass Mean Expression	annotation
Yes	TR10603	0.025	0.021	1.183	558	13	2.789	10.211	464.581	0.767	0	2.326	0.468	*X-importin-α2*
Yes	TR2814	0.018	0.017	1.048	399	7	2.072	4.928	585.588	0.911	0	1.737	1.014	*X-importin-α2*
No	TR23125	0.012	0.01	1.29	609	7	1.527	5.473	420.433	2.638	0	1.149	1.046	l(1)1Bi
Yes	TR24932	0.019	0.007	2.694	357	5	0.792	4.208	208.266	-1.22	0	0.772	1.578	none
No	TR6297	0.008	0.008	1.012	510	4	0.998	3.002	514.679	1.188	0	0.67	0.718	CG7366
No	TR5481	0.013	0.002	8.283	1137	10	0.521	9.479	426.12	-0.873	0	0.876	0.718	none
Yes	TR261	0.02	0	50	522	8	0.05	7.95	911.2	-0.613	0	0.936	6.488	CG32371
Yes	TR11103	0.008	0.004	2.124	1224	8	1.306	6.694	508.192	-0.036	0.949	0.671	11.376	CG2685
Yes	TR37304	0.006	0.004	1.565	1371	8	1.44	6.56	3140.84	-0.427	0.001	0.593	16.864	CG4198
No	TR50351	0.013	0.011	1.201	678	8	1.699	6.301	742.04	-0.346	0.1	1.261	2.655	CG15452

Table S4. The transcripts from canonical *importin-α* genes and *X-importin-α2* in *D. neotestacea.* Mean expression level, $\log_2$ fold change between ST and SR, and FDR corrected p-value was calculated in DESeq2.

Gene	Transcript	Mean Expression	$\log_2$ Fold Change	Adjusted P-value	Chromosome	Carcass Mean Expression	% difference	K_a/K_s	Substitutions	Length
importin-α1	TR22571	5769.290	0.177	0.534	D	29.387	0.000	0.000	0.000	1655
importin-α3	TR6773	2018.999	0.134	0.482	E	49.905	0.000	0.000	0.000	1542
importin-α2	TR7043	52604.270	-0.081	0.668	B	136.289	0.000	0.000	0.000	1569
X-importin-α2	TR37105	1149.714	0.450	0.013	X	1.514	1.923	0.349	12.000	624
X-importin-α2	TR10603	464.581	0.767	0.000	X	0.468	2.326	1.183	13.000	559
X-importin-α2	TR2814	585.588	0.911	0.000	X	1.014	1.737	1.048	7.000	403

Table S5. The population origin of males used in for the population genetic analysis. Not all individual samples were sequenced at every locus.

Population	Collection location	SR samples	ST samples
AB3	Jasper, AB	2	0
BC	Vancouver, BC	1	2
ID	Coeur d'Alene, ID	1	4
MB	The Pas, MB	0	1
MN	Bemidji, MN	1	2
MT1	Columbia Falls, MT	0	2
MT2	St. Regis, MT	2	3
ND	Minot, ND	0	1
OR	MacKenzie Bridge, OR	3	1
PEI	Charlottetown, PEI	0	1

Table S6. Molecular population genetic summary statistics for each locus. Statistics for *D. neotestacea* samples. S (segregating sites) for total and silent sites is presented. M (mutations), π, θ, and D for silent sites only includes synonymous sites in open reading frames and all noncoding sites. π NS is using nonsynonymous polymorphism only. ρ is population recombination rate ($2N_er$) divided by the number of nucleotides in the marker. For the autosomal markers, all statistics were calculated from the combined set of SR and ST individuals. ρ and ZnS could not be calculated for some markers due to a lack of segregating sites. Bolded D values are less than the expected D in 95% of 10,000 simulations. Da is percent divergence between ST and SR. See also (Pieper and Dyer 2016).

Region/ Marker	SR or ST	N	Total sites	S	Silent sites	M silent	S silent	π silent	π NS	θ silent	D silent	ρ	ZnS	π_a/π_s	K_a/K_s	Da
X-linked																
marf	ST	57	1063	186	670.34	206	184	0.043	0.005	0.067	-0.999	0.122	0.037	0.195	0.424	1.581
	SR	48		73	844.49	74	71	0.009	0.001	0.020	**-1.832**	0.008	0.252	0.577	0.282	
mof	ST	65	576	30	128.36	26	26	0.009	0.000	0.043	**-2.533**	0.000	0.041	0.044	0.065	0.002
	SR	47		7	128.34	6	6	0.004	0.000	0.011	**-1.641**	0.000	0.049	0.038	0.064	
pgd	ST	56	569	38	128.98	30	28	0.026	0.002	0.051	**-1.476**	0.146	0.023	0.064	0.069	-0.010
	SR	43		18	127.7	9	9	0.012	0.001	0.016	**-0.826**	0.000	0.284	0.092	0.068	
rpl	ST	57	302	11	123.33	11	11	0.005	0.000	0.019	**-2.109**	0.024	0.051	0.00	0.00	0.008
	SR	44		4	123.33	4	4	0.004	0.000	0.007	-1.162	0.066	0.004	0.00	0.000	
spk	ST	56	382	13	81.17	12	12	0.021	0.000	0.032	-1.029	0.298	0.015	0.006	0.021	0.041
	SR	45		1	81.17	1	1	0.001	0.000	0.003	-1.113	NA	NA	0.00	0.022	
Autosomal																
esc		82	370	29	78.32	29	27	0.025	0.001	0.074	**-1.956**	0.130	0.023	0.020	0.002	
gl		80	402	52	96.96	42	39	0.035	0.002	0.087	**-1.843**	0.065	0.023	0.060	0.059	
mago		78	324	14	57.15	7	7	0.010	0.001	0.025	**-1.429**	0.086	0.058	0.109	0.500	
ntid		88	527	50	123.42	44	41	0.033	0.003	0.071	**-1.575**	0.068	0.027	0.138	0.110	
sia		82	400	13	96.33	14	13	0.004	0.000	0.029	**-2.336**	0.000	0.015	0.024	0.084	
tpi		80	347	33	83.28	29	29	0.044	0.001	0.070	-1.011	0.245	0.043	0.00	0.000	
wee		84	285	14	58.14	10	10	0.015	0.001	0.034	**-1.462**	0.561	0.027	0.048	0.003	

Table S7. Data used for the HKA tests. The first set of data is for SR, and the second set is for ST. Divergence is measured between ST and SR. S is the number of segregating sites, N is the number of samples, and divergence is measured between ST and SR. The scalar was calculated as 0.75*0.15 for SR and 0.75*0.85 for ST. Top candidates are noted with TC.

SR or ST	TC?	Marker	Length	S	N	Divergence	Theta	Scalar
SR	TC	x-imp-α2	1905	8	9	43.139	0.00148	0.1125
	TC	TR261	627	10	8	12.25	0.00631	0.1125
	TC	TR11103	826	0	10	4.556	0.00001	0.1125
	TC	TR23125	343	6	10	4.167	0.00618	0.1125
	TC	TR24932	280	1	8	7.234	0.00138	0.1125
	TC	TR37304	770	0	10	5.1	0.00001	0.1125
	TC	TR50351	649	37	9	16.522	0.0207	0.1125
	no	*marf*	788	54	46	29.996	0.01779	0.1125
	no	*mof*	576	7	47	0.889	0.00275	0.1125
	no	*pgd*	565	18	42	4.356	0.0074	0.1125
	no	*rpl*	302	4	44	0.565	0.00304	0.1125
	no	*spk*	382	14	45	1.04	0.0006	0.1125
ST	TC	x-imp-α2	1905	142	12	43.139	0.02573	0.6375
	TC	TR261	627	25	10	12.25	0.01527	0.6375
	TC	TR11103	826	10	10	4.556	0.00713	0.6375
	TC	TR23125	343	8	9	4.167	0.00843	0.6375
	TC	TR24932	280	10	8	7.234	0.01717	0.6375
	TC	TR37304	770	4	10	5.1	0.00184	0.6375
	TC	TR50351	649	50	19	16.522	0.02832	0.6375
	no	*marf*	788	186	57	29.996	0.05808	0.6375
	no	*mof*	576	30	65	0.889	0.01098	0.6375
	no	*pgd*	565	38	56	4.356	0.01541	0.6375
	no	*rpl*	302	11	57	0.565	0.0079	0.6375
	no	spk	382	13	56	1.04	0.00741	0.6375

Table S8: See the associated excel file). The results of a GO terms enrichment analysis for different sets of differentially expressed transcripts. Each spreadsheet indicates a different set of DE transcripts.