

Additional file 3. Genes from *D. buzzatii* (*Dbuz*), *D. melanogaster* (*Dmel*) and *D. pseudoobscura* (*Dpse*) used in the analyses with their accession number in Genbank or Flybase and their location on the chromosome. In the case of *D. buzzatii*, the original references where the sequences were first published are also supplied. The last column contains the UniProt Knowledgebase accession numbers used for annotating repetitive blocks.

Gene symbol	Tr ^a	Tr used ^b	Intron size ^c	<i>Dbuz</i>			<i>Dmel</i>		<i>Dpse</i>		UniProt Knowledge base acc.
				Genbank acc.	Location	Reference	Flybase acc.	Location	Genbank acc.	Location	
Hox genes											
<i>abd-A</i>	2	PB	17448	AY900631	2, <i>G4g</i>	Negre et al. 2005	FBgn0000014	3R, <i>89E2</i>	AADE01000036	2, <i>Cont1892</i>	P29555
<i>lab</i>	1	PA	14389	AY900631	2, <i>G4g-5a</i>	Negre et al. 2005	FBgn0002522	3R, <i>84A1</i>	AADE01000437	2, <i>Cont1676</i>	P10105
<i>pb</i>	4	PA	30536	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0051481	3R, <i>84A5</i>	AADE01000149	2, <i>Cont1677</i>	P31264
Hox-derived genes											
<i>bcd</i>	3	PD	1126	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0000166	3R, <i>84A5</i>	AADE01000149	2, <i>Cont1677</i>	P09081
<i>zen</i>	1	PA	64	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0004053	3R, <i>84A5</i>	AADE01000149	2, <i>Cont1677</i>	P09089
<i>zen2</i>	1	PA	67	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0004054	3R, <i>84A5</i>	AADE01000149	2, <i>Cont1677</i>	P09090
Non-Hox genes											
<i>Adhr</i>	1	PA	0	AF260699	3, <i>G1a</i>	Betran and Ashburner 2000	FBgn0000056	2L, <i>35B3</i>	AADE01001152	4, <i>Cont1559</i>	P91615
<i>α-Est2</i>	1	PA	528	AF216210	2, <i>F5e-6a</i>	Robin et al. 2000	FBgn0015570	3R, <i>84D9</i>	AADE01001585	2, <i>Cont390</i>	Q961N0
<i>α-Est3</i>	1	PA	835	AF216211	2, <i>F5e-6a</i>	Robin et al. 2000	FBgn0015571	3R, <i>84D9</i>	AADE01001585	2, <i>Cont390</i>	Q8SZW5
<i>Ccp3 – Ccp84Ac</i>	1	PA	64	AY900631	2, <i>G4g</i>	Negre et al. 2005	FBgn0004781	3R, <i>84A3--4</i>	AADE01000149	2, <i>Cont1677</i>	O97060
<i>Ccp6 – Ccp84Ae</i>	1	PA	118	AY900631	2, <i>G4g</i>	Negre et al. 2005	FBgn0004779	3R, <i>84A3</i>	AADE01000437	2, <i>Cont1676</i>	O97062
<i>Ccp7 – Ccp84Af</i>	1	PA	58	AY900631	2, <i>G4g</i>	Negre et al. 2005	FBgn0004778	3R, <i>84A3</i>	AADE01000437	2, <i>Cont1676</i>	O97063
<i>Ccp8 – Ccp84Ag</i>	1	PA	215	AY900631	2, <i>G4g</i>	Negre et al. 2005	FBgn0004777	3R, <i>84A2</i>	AADE01000437	2, <i>Cont1676</i>	O97064
<i>CG1288</i>	1	PA	55	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0037484	3R, <i>84D2</i>	AADE01000014	2, <i>Cont268</i>	Q9VI77
<i>CG13617</i>	1	PA	229	AY551073	2, <i>E4g-5a</i>	Puig, Caceres, and Ruiz 2004	FBgn0039201	3R, <i>96A5</i>	AADE01000267	2, <i>Cont3813</i>	Q9VC70
<i>CG14290</i>	1	PB	509	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0038662	3R, <i>91D5</i>	AADE01000175	4, <i>Cont2631</i>	Q7KSC4
<i>CG14609</i>	1	PA	400	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0037483	3R, <i>84D2</i>	AADE01000014	2, <i>Cont268</i>	Q8T3W3
<i>CG14899</i>	1	PA	140	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0038438	3R, <i>89B18</i>	AADE01000429	2, <i>Cont2827</i>	Q9VEU2
<i>CG17836</i>	4	PA	5563	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0038661	3R, <i>91D4--5</i>	AADE01000175 AADE01000328	4, <i>Cont2631</i> 4, <i>Cont3752</i>	Q9VE23
<i>CG2520</i>	1	PA	9607	AY900632	2, <i>F1c</i>	Negre et al. 2005	FBgn0026210	3R, <i>84D1--2</i>	AADE01000014	2, <i>Cont268</i>	Q9VI75
<i>CG31363</i>	4	PE	12711	AY900631	2, <i>G4g</i>	Negre et al. 2005	FBgn0051363	3R, <i>86E11--13</i>	AADE01002495 AADE01000322 AADE01006966 AADE01000546	2, <i>Cont1082</i> 2, <i>Cont2641</i> 2, <i>Cont5794</i> 2, <i>Cont1019</i>	Q9VGL6
<i>Lsp1β</i>	1	PA	68	AY561258	2, <i>D3c</i>	Gonzalez, Casals, and Ruiz 2004	FBgn0002563	2L, <i>21E2</i>	AADE01000478	4, <i>Cont3520</i>	Q6Q431
<i>Lsp1γ</i>	1	PA	65	AY561259	3, <i>C2g</i>	Gonzalez, Casals, and Ruiz 2004	FBgn0002564	3L, <i>61A6</i>	AADE01000061	4, <i>Cont1620</i>	Q6Q429

^a Number of transcripts that exist in all the three species.

^b The longest transcript shared by all the three species and that was used for the analyses.

^c Total intron size in base pairs (sum of all introns in the gene, including those outside the limits of the coding sequence, in *D. melanogaster*).

