Supporting Online Material for

A natural deletion of the *HoxC* cluster in elasmobranch fishes

Benjamin L. King*, J. Andrew Gillis, Heather R. Carlisle and Randall D. Dahn

*To whom correspondence should be addressed. E-mail: bking@mdibl.org

This PDF file includes:

Materials and Methods

Isolation of DNA and RNA

Genomic DNA was extracted from a single *L. erinacea* Stage 32 embryo (Marine Biological Laboratory, Woods Hole, MA) by C. Mattingly using the Gentra Puregene kit (Qiagen). Total RNA was extracted using Tri Reagent (Ambion) from the following pooled tissue samples: six *L. erinacea* Stage 20-29 embryos (Marine Biological Laboratory, Woods Hole, MA), six *S. canicula* Stage 24-30 (Blue Reef Aquarium, Hastings, UK), and two *C. milii* stage 32 embryos (collected by J. A. G.).

DNA and RNA Sequencing

A 500bp paired-end genomic DNA library was constructed using standard Illumina protocols. 806,309,979 125 bp paired-end sequence reads were generated on an Illumina Genome Analyzer IIx (Delaware Biotechnology Institute, University of Delaware) using eleven flow cells. Embryonic RNA libraries for *L. erinacea*, *S. canicula* and *C. milii* were generated using standard Illumina protocols for single-end sequencing. 49,509,611, 50,760,902 and 43,835,550 76 bp sequencing reads were generated for *L. erinacea*, *S. canicula* and *C. milii*, respectively, on an Illumina Genome Analyzer II (Centre for Applied Genomics, Hospital for Sick Children, Toronto, Canada) using two flow cell lanes. Base calling was determined using Illumina pipeline software according to manufacturer's specification (Illumina).

Genome and Transcriptome Assemblies

Genome and transcriptome assemblies were made using CLCBio Genome Workbench version 4 (http://www.clcbio.com). The transcriptome assemblies generated 103,996 (N50 = 757 bp), 107,231 (N50 = 695 bp) and 92,334 (N50 = 576 bp) contiguous sequences for *L. erinacea*, *S. canicula* and *C. milii*, respectively. 49,10,973, 50,748,002 and 43,813,111 trimmed reads were used in the assemblies for *L. erinacea*, *S. canicula* and *C. milii*, respectively. There were 11,927,042, 14,836,404 and 16,660,753 singletons for *L. erinacea*, *S. canicula* and *C. milii*, respectively. The *L. erinacea* genome assembly used 727,225,088 trimmed reads and had 2,962,365 contigs (N50 = 665 bp) with an overall length of 1,555,444,314 bp and 52,104,024 singletons.

Gene Annotation

Hox coding sequences and microRNAs were predicted using reciprocal sequence similarity searches. Nucleotide and protein sequences for each *C. milii Hox* gene were aligned against assembled genomic and transcriptomic DNA sequences for each species using Decypher Tera-BLASTN and Tera-BLASTP searches, respectively (TimeLogic Decypher BLAST version 8.0.1). Sequences aligning to particular *C. milii Hox* sequences were validated via reciprocal blastn and blastx searches (NCBI BLAST) against sequences in GenBank and non-redundant protein databases. *C. milii*-human conserved sequence elements were found using NCBI BLAST version 2.2.23+.

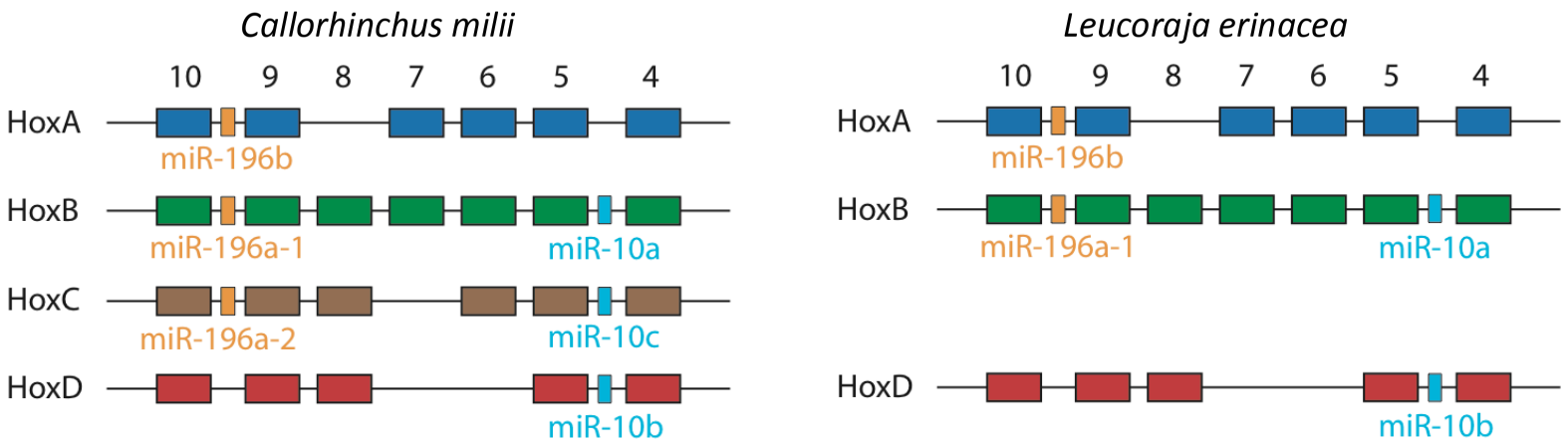
Statistical Analyses

Analysis of gene expression levels by read coverage was done using the Kruskal-Wallis test using values in Table S1 to determine whether the distribution of read coverage was similar among *HoxA, HoxB, HoxC* and *HoxD* genes for each species. For genes represented by reads instead of contigs, a cumulative length of 125 bp and an average coverage of zero were used in the calculation of the average cumulative length and coverage. Analysis of *L. erinacea* genome read coverage (Table S1) for *Hox* genes was performed using the Wilcoxon rank-sum test to determine whether the read coverage among *HoxA, HoxB* and *HoxD* genes was significantly greater than the coverage for *HoxC*. Analysis of the presence and absence of *C. milii*-human CSEs (Table S2) were calculated using Fisher's Exact Test using a contingency table where the proportion of detected *HoxC* elements (zero to 25) is compared to the proportion for *HoxA, HoxB* and *HoxD* elements (126 to 249). Analyses were performed in the R statistical computing environment (http://www.r-project.org), version 2.13.1, using the R/coin package, version 1.0-19.

Supplemental Figure 1: *HoxC*-encoded microRNAs are not present in the skate genome. (a) The genome of *C. milii* encodes three miRNAs each of the *Hox*-specific miR-10 and miR-196 families. A 26-fold coverage draft of the skate genome reveals the presence of only four miRNAs: *pre-miR-196b* physically linked to the *HoxA10* gene, *pre-miR-196a-1* and *pre-miR-10a* linked to the *HoxB10* and *HoxB5* genes, respectively, and *pre-miR-10b* linked to the *HoxD4* gene. No sequences encoding the *HoxC*-embedded *pre-miR-10c* or *pre-miR-196a-2* genes were detected (p = 6.51 x 10^{-7} by the Wilcoxon Rank Sum Test). (b) An unrooted phylogram of all *Hox*-specific miRNA sequences in the genomes of *C. milii* and *L. erinacea* confirms that the *HoxC*-specific *miR-10c* and *miR-196a-2* genes are absent in the skate, further supporting the notion that the *HoxC* cluster has been deleted in elasmobranch fishes. *C. milii* miRNAs are prefaced by 'Cm,' and *L. erinacea* miRNAs by 'Le'. *C. milii* miRNAs which have no ortholog in *L. erinacea* are indicated by red text.

Fig. S1

A


Callorhinchus milii
10 9 8 7 6 5 4
HoxA
miR-196b
HoxB
miR-196a-1
miR-10a
HoxC
miR-196a-2
miR-10c
HoxD
miR-10b
Leucoraja erinacea
10 9 8 7 6 5 4
HoxA
miR-196b
HoxB
miR-196a-1
miR-10a
HoxD
miR-10b

B

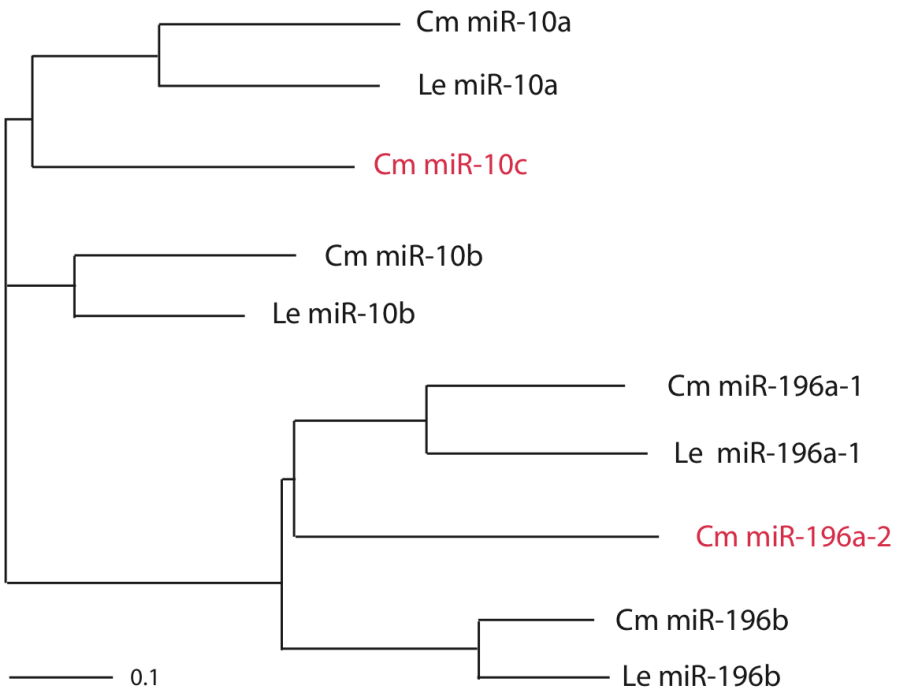

Cm miR-10a
Le miR-10a
Cm miR-10c
Cm miR-10b
Le miR-10b
Cm miR-196a-1
Le miR-196a-1
Cm miR-196a-2
Cm miR-196b
Le miR-196b
0.1

Table S1: Transcriptome representation of *HoxA, B, C* and *D* clusters for *C. milii, L. erinacea* and *S. canicula* and genome representation of *HoxA, B, C* and *D* clusters for *L. erinacea*. For each *Hox* gene, the protein length, cumulative length of contigs, number of reads that map to the contigs and the average coverage is listed. Unique identifiers are listed for *L. erinacea* genomic contigs. The *L. erinacea* genomic contigs for pre-miR-196b, pre-miR-196a-1, miR-10a and miR-10b were LER_WGS_1_CONTIG_2688791, LER_WGS_1_CONTIG_88050, LER_WGS_1_CONTIG_46923 and LER_WGS_1_CONTIG_2672030, respectively. Singlet only = no contig coverage; single read sequence coverage only. N/A = not applicable; no coverage found.

Supplemental Table 1

	C. milii				*L. erinacea (St28)*			*S. canicula*			*L. erinacea* Genomic			
GENE	**LENGTH (aa)**	**CUMULATIVE CONTIG LENGTH (bp)**	**TOTAL READ COUNT**	**AVERAGE COVERAGE**	**CUMULATIVE CONTIG LENGTH (bp)**	**TOTAL READ COUNT**	**AVERAGE COVERAGE**	**CUMULATIVE CONTIG LENGTH (bp)**	**TOTAL READ COUNT**	**AVERAGE COVERAGE**	**GENOME CONTIG ID**	**CUMULATIVE CONTIG LENGTH (bp)**	**TOTAL READ COUNT**	**AVERAGE COVERAGE**
HOXA1	326	254	16	4.79	Singlet only	Singlet only	< 1	Singlet only	Singlet only	< 1	LER_WGS_1_CONTIG_42404 LER_WGS_1_CONTIG_71793	3360	856	31.85
HOXA2	363	1492	909	46.30	1358	632	35.37	1439	583	30.79	LER_WGS_1_CONTIG_62711	5627	1080	23.99
HOXA3	409	1782	1102	47.00	1922	1733	68.53	2684	705	19.96	LER_WGS_1_CONTIG_55082 LER_WGS_1_CONTIG_35378 LER_WGS_1_CONTIG_28652	6685	1331	24.89
HOXA4	247	1624	572	26.77	1991	686	26.19	413	80	14.72	LER_WGS_1_CONTIG_37027	4271	845	24.73
HOXA5	281	1546	350	17.21	2378	802	25.63	1186	311	19.93	LER_WGS_1_CONTIG_2669412 LER_WGS_1_CONTIG_36107	10013	2074	25.89
HOXA6	229	203	13	4.87	1270	163	9.75	610	70	8.72	LER_WGS_1_CONTIG_36107	5783	1191	25.74
HOXA7	212	221	34	11.69	867	308	27.00	530	64	9.18	LER_WGS_1_CONTIG_1620669 LER_WGS_1_CONTIG_57367	8092	1476	22.80
HOXA9	259	1044	473	34.43	1961	1080	41.86	1940	784	30.71	LER_WGS_1_CONTIG_39561	7645	1564	25.57
HOXA10	344	1973	1262	48.61	1515	789	39.58	1850	677	27.81	LER_WGS_1_CONTIG_1651197 LER_WGS_1_CONTIG_36067	8964	1414	19.72
HOXA11	293	2309	1046	34.43	2696	1397	39.38	816	1362	126.85	LER_WGS_1_CONTIG_54663	7551	1503	24.88
HOXA13	296	1221	525	32.68	1076	275	19.42	1021	132	9.83	LER_WGS_1_CONTIG_35349	8650	1587	22.93
HOXB1	318	270	17	4.79	Singlet only	Singlet only	< 1	Singlet only	Singlet only	< 1	LER_WGS_1_CONTIG_1644638 LER_WGS_1_CONTIG_2666672	4178	863	25.82
HOXB2	362	1640	488	22.61	1089	477	33.29	1000	393	29.87	LER_WGS_1_CONTIG_2702210 LER_WGS_1_CONTIG_480802	2117	368	21.73
HOXB3	394	1743	1205	52.54	1544	823	40.51	1520	838	41.90	LER_WGS_1_CONTIG_58055	2466	449	22.76
HOXB4	243	606	134	16.81	535	72	10.23	1728	272	11.96	LER_WGS_1_CONTIG_46923	4238	807	23.80
HOXB5	260	847	183	16.42	1037	286	20.96	1021	251	18.68	LER_WGS_1_CONTIG_690334 LER_WGS_1_CONTIG_563542	1944	353	22.70
HOXB6	230	1070	218	15.48	1431	330	17.53	435	40	6.99	LER_WGS_1_CONTIG_40659	4828	925	23.95
HOXB7	211	856	377	33.47	1307	712	41.40	1096	452	31.34	LER_WGS_1_CONTIG_1633482 LER_WGS_1_CONTIG_80692	2969	564	23.75
HOXB8	236	1058	425	30.53	1619	785	36.85	1545	409	20.12	LER_WGS_1_CONTIG_65401 LER_WGS_1_CONTIG_2787112	1594	332	26.04
HOXB9	250	715	218	23.17	1160	304	19.92	1294	356	20.91	LER_WGS_1_CONTIG_1659208 LER_WGS_1_CONTIG_2676692	8019	1888	29.43
HOXB10	337	2737	1605	44.57	1134	386	25.87	1381	306	16.84	LER_WGS_1_CONTIG_93316 LER_WGS_1_CONTIG_121676	1710	329	24.05
HOXB13	308	1546	2000	98.32	767	211	20.91	1413	271	14.58	LER_WGS_1_CONTIG_72124	1544	323	26.15
HOXC1	291	Singlet only	7	< 1	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC3	408	1330	186	10.63	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC4	256	606	92	11.54	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC5	231	1662	452	20.67	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC6	219	1162	213	13.93	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC8	244	1270	228	13.64	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC9	262	1149	204	13.49	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC10	357	1312	187	10.83	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC11	300	743	119	12.17	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC12	270	1108	684	46.92	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXC13	318	1397	975	53.04	N/A	N/A	0.00	N/A	N/A	0.00	N/A	N/A	N/A	0.00
HOXD1	315	1445	410	21.56	948	58	4.65	798	73	6.95	LER_WGS_1_CONTIG_47922	2997	824	34.37
HOXD2	356	1379	475	26.18	1217	112	6.99	1267	253	15.18	LER_WGS_1_CONTIG_47792	2877	578	25.11
HOXD3	394	1603	628	29.77	1764	657	28.31	3562	1879	40.09	LER_WGS_1_CONTIG_61764 LER_WGS_1_CONTIG_42923	5501	1006	22.86
HOXD4	238	1131	442	29.70	1952	924	35.98	1866	468	19.06	LER_WGS_1_CONTIG_56573 LER_WGS_1_CONTIG_2672030	8294	1767	26.63
HOXD5	249	1264	229	13.77	1259	617	37.25	802	212	20.09	LER_WGS_1_CONTIG_904538	3566	692	24.26
HOXD8	239	1174	174	11.26	1414	410	22.04	852	160	14.27	LER_WGS_1_CONTIG_1618096	4177	836	25.02
HOXD9	265	1768	696	29.92	923	412	33.92	1216	341	21.31	LER_WGS_1_CONTIG_75394	2314	525	28.36
HOXD10	338	1317	275	15.87	1306	426	24.79	2324	479	15.66	LER_WGS_1_CONTIG_34676	7975	1473	23.09
HOXD11	296	1159	222	14.56	1460	163	8.48	1177	83	5.36	LER_WGS_1_CONTIG_45979 LER_WGS_1_CONTIG_975768	12189	2410	24.71
HOXD12	269	1799	1400	59.14	1443	699	36.81	1405	574	31.05	LER_WGS_1_CONTIG_45979	8323	1702	25.56
HOXD13	313	1653	612	28.14	1295	226	13.26	1092	83	5.78	LER_WGS_1_CONTIG_33336	2068	454	27.44
HOXD14	266	331	51	11.71	Singlet only	Singlet only	< 1	484	38	5.97	LER_WGS_1_CONTIG_1606040	3561	828	29.06

Table S2: Genomic representation sequence elements conversed between orthologous *C. milii* and human *HoxA, B, C* and *D* clusters *(S1)* within the *L. erinacea* genome. The presence or absence of each element is listed. For the elements that were found, the unique contig identifier, length, number of reads that map to the contig, and the average coverage is listed. The presence (Y) and absence (N) of elements in the *L. erinacea HoxA* cluster (GenBank FJ944024) and the *S. canicula HoxA* (GenBank FQ03265), *HoxB* (GenBank FQ032659), *and HoxD* (GenBank FQ032660) clusters are also shown. N/A = not applicable; no coverage found.

Supplemental Table 2

HoxA

CSE ID	*C. milii HoxA* COORDINATES (bp) in GenBank FJ824598	*L. erinacea* CONTIG ID	CONTIG LENGTH (bp)	AVERAGE COVERAGE	FOUND IN *L. erinacea HoxA* (GenBank FJ944024)	FOUND IN *S. canicula HoxA* (GenBank FQ03265)
EH-HoxA_CNE1	3789-3903	LER_WGS_1_CONTIG_1623245	3476	24.8	N	N
EH-HoxA_CNE2	3938-4208	LER_WGS_1_CONTIG_1623245	3476	24.8	N	N
EH-HoxA_CNE3	8718-8763	N/A	N/A	0	N	N
EH-HoxA_CNE4	34105-34232	LER_WGS_1_CONTIG_58188	948	22.53	N	N
EH-HoxA_CNE5	35244-35293	N/A	N/A	0	N	N
EH-HoxA_CNE6	35317-35387	N/A	N/A	0	N	N
EH-HoxA_CNE7	35398-35705	LER_WGS_1_CONTIG_262779	1268	12.13	N	N
EH-HoxA_CNE8	35711-35802	N/A	N/A	0	N	N
EH-HoxA_CNE9	35860-36014	LER_WGS_1_CONTIG_262779	1268	12.13	N	N
EH-HoxA_CNE10	45157-45223	N/A	N/A	0	N	N
EH-HoxA_CNE11	63289-63368	N/A	N/A	0	N	N
EH-HoxA_CNE12	64734-64974	LER_WGS_1_CONTIG_713392	2827	25.12	N	N
EH-HoxA_CNE13	65942-66057	N/A	N/A	0	N	N
EH-HoxA_CNE14	66143-66324	N/A	N/A	0	N	N
EH-HoxA_CNE15	72079-72259	LER_WGS_1_CONTIG_2670831	4220	22.6	N	N
EH-HoxA_CNE16	90972-91021	N/A	N/A	0	N	N
EH-HoxA_CNE17	93833-93881	N/A	N/A	0	N	N
EH-HoxA_CNE18	109401-109484	N/A	N/A	0	N	N
EH-HoxA_CNE19	115763-115833	LER_WGS_1_CONTIG_1610348	3344	24.67	Y	Y
EH-HoxA_CNE20	121666-121766	LER_WGS_1_CONTIG_35349	8650	20.4	Y	Y
EH-HoxA_CNE21	131975-132127	LER_WGS_1_CONTIG_54663	7551	21.67	Y	Y
EH-HoxA_CNE22	132800-133000	LER_WGS_1_CONTIG_54663	7551	21.67	Y	Y
EH-HoxA_CNE23	136539-136634	LER_WGS_1_CONTIG_54663	7551	21.67	Y	Y
EH-HoxA_CNE24	137148-137204	LER_WGS_1_CONTIG_54663	7551	21.67	Y	Y
EH-HoxA_CNE25	145442-145559	LER_WGS_1_CONTIG_36067	7183	16.43	Y	Y
EH-HoxA_CNE26	146401-146468	LER_WGS_1_CONTIG_69988	1440	16.16	Y	Y
EH-HoxA_CNE27	149144-149309	LER_WGS_1_CONTIG_1651197	1781	17.79	Y	Y
EH-HoxA_CNE28	153746-153796	N/A	N/A	0	N	N
EH-HoxA_CNE29	155724-155841	LER_WGS_1_CONTIG_39561	7645	20.9	Y	N
EH-HoxA_CNE30	155907-155959	N/A	N/A	0	N	N
EH-HoxA_CNE31	157784-157860	N/A	N/A	0	N	N
EH-HoxA_CNE32	157869-157945	N/A	N/A	0	N	Y
EH-HoxA_CNE33	158465-158514	N/A	N/A	0	N	N
EH-HoxA_CNE34	158525-158580	N/A	N/A	0	N	N
EH-HoxA_CNE35	159715-159782	LER_WGS_1_CONTIG_39561	7645	20.9	Y	N
EH-HoxA_CNE36	159953-160122	LER_WGS_1_CONTIG_39561	7645	20.9	Y	Y
EH-HoxA_CNE37	161878-161935	N/A	N/A	0	N	N
EH-HoxA_CNE38	165391-165487	LER_WGS_1_CONTIG_1620669	4244	22.87	Y	Y
EH-HoxA_CNE39	165992-166087	LER_WGS_1_CONTIG_1620669	4244	22.87	Y	Y
EH-HoxA_CNE40	170188-170386	LER_WGS_1_CONTIG_57367	3848	17.06	Y	Y
EH-HoxA_CNE41	175586-175801	LER_WGS_1_CONTIG_36107	5783	22.29	Y	Y
EH-HoxA_CNE42	178283-178347	N/A	N/A	0	N	N
EH-HoxA_CNE43	178466-178712	LER_WGS_1_CONTIG_36107	5783	22.29	Y	Y
EH-HoxA_CNE44	180092-180205	N/A	N/A	0	N	N
EH-HoxA_CNE45	180237-180314	LER_WGS_1_CONTIG_2669412	4230	23.36	Y	Y
EH-HoxA_CNE46	180337-180413	LER_WGS_1_CONTIG_2669412	4230	23.36	Y	Y
EH-HoxA_CNE47	180518-180658	LER_WGS_1_CONTIG_2669412	4230	23.36	Y	Y
EH-HoxA_CNE48	180659-180735	LER_WGS_1_CONTIG_2669412	4230	23.36	Y	Y
EH-HoxA_CNE49	181418-181760	LER_WGS_1_CONTIG_2669412	4230	23.36	Y	Y
EH-HoxA_CNE50	182786-182833	N/A	N/A	0	N	N
EH-HoxA_CNE51	185772-185847	N/A	N/A	0	N	N
EH-HoxA_CNE52	186235-186524	LER_WGS_1_CONTIG_37734	4962	20.93	Y	Y
EH-HoxA_CNE53	188671-188883	LER_WGS_1_CONTIG_37734	4962	20.93	Y	Y
EH-HoxA_CNE54	189397-189466	N/A	N/A	0	N	N
EH-HoxA_CNE55	191586-191656	LER_WGS_1_CONTIG_37027	4271	21.82	Y	Y
EH-HoxA_CNE56	192131-192324	LER_WGS_1_CONTIG_37027	4271	21.82	Y	Y
EH-HoxA_CNE57	193576-193624	N/A	N/A	0	N	N
EH-HoxA_CNE58	194873-195023	LER_WGS_1_CONTIG_37027	4271	21.82	Y	Y
EH-HoxA_CNE59	195307-195422	N/A	N/A	0	N	Y
EH-HoxA_CNE60	196682-196790	LER_WGS_1_CONTIG_33092	5789	22.12	Y	Y
EH-HoxA_CNE61	197088-197182	LER_WGS_1_CONTIG_33092	5789	22.12	Y	Y
EH-HoxA_CNE62	197229-197286	LER_WGS_1_CONTIG_33092	5789	22.12	Y	Y
EH-HoxA_CNE63	201894-201941	N/A	N/A	0	N	N
EH-HoxA_CNE64	202054-202171	N/A	N/A	0	N	N
EH-HoxA_CNE65	202554-202702	LER_WGS_1_CONTIG_2704880	5703	20.57	Y	Y
EH-HoxA_CNE66	202786-202880	LER_WGS_1_CONTIG_2704880	5703	20.57	Y	Y
EH-HoxA_CNE67	205622-205848	LER_WGS_1_CONTIG_2704880	5703	20.57	Y	Y
EH-HoxA_CNE68	206385-206438	LER_WGS_1_CONTIG_2704880	5703	20.57	Y	Y
EH-HoxA_CNE69	206718-206883	LER_WGS_1_CONTIG_2704880	5703	20.57	Y	Y
EH-HoxA_CNE70	208694-208885	LER_WGS_1_CONTIG_55082	3522	25.2	Y	Y
EH-HoxA_CNE71	209810-209902	N/A	N/A	0	N	Y
EH-HoxA_CNE72	209939-210056	N/A	N/A	0	N	Y
EH-HoxA_CNE73	214124-214190	N/A	N/A	0	Y	Y
EH-HoxA_CNE74	214271-214426	LER_WGS_1_CONTIG_35378	2830	19.85	Y	Y
EH-HoxA_CNE75	217183-217308	LER_WGS_1_CONTIG_62711	5627	20.9	Y	Y
EH-HoxA_CNE76	217935-218077	LER_WGS_1_CONTIG_62711	5627	20.9	Y	Y
EH-HoxA_CNE77	218203-218444	LER_WGS_1_CONTIG_62711	5627	20.9	Y	Y
EH-HoxA_CNE78	219000-219163	LER_WGS_1_CONTIG_62711	5627	20.9	Y	Y
EH-HoxA_CNE79	219292-219357	LER_WGS_1_CONTIG_62711	5627	20.9	Y	Y
EH-HoxA_CNE80	220367-220454	N/A	N/A	0	N	N
EH-HoxA_CNE81	221662-221716	N/A	N/A	0	N	N
EH-HoxA_CNE82	221742-221842	N/A	N/A	0	N	N
EH-HoxA_CNE83	224375-224488	N/A	N/A	0	N	N
EH-HoxA_CNE84	224524-224583	N/A	N/A	0	N	N
EH-HoxA_CNE85	226394-226468	N/A	N/A	0	N	Y
EH-HoxA_CNE86	226883-226931	N/A	N/A	0	N	N
EH-HoxA_CNE87	228910-229095	N/A	N/A	0	N	N
EH-HoxA_CNE88	229212-229314	LER_WGS_1_CONTIG_402806	997	21.41	Y	Y
EH-HoxA_CNE89	237432-237486	N/A	N/A	0	N	N
EH-HoxA_CNE90	246767-246825	N/A	N/A	0	N	N
EH-HoxA_CNE91	249213-249258	N/A	N/A	0	N	N
EH-HoxA_CNE92	261949-262088	N/A	N/A	0	N	N
EH-HoxA_CNE93	262173-262272	N/A	N/A	0	N	N
EH-HoxA_CNE94	269893-269962	N/A	N/A	0	N	N
EH-HoxA_CNE95	272230-272386	LER_WGS_1_CONTIG_60616	1709	14.71	N	N
EH-HoxA_CNE96	280009-280058	N/A	N/A	0	N	N

HoxB

CSE ID	*C. milii HoxB* COORDINATES (bp) in GenBank FJ824599	*L. erinacea* CONTIG ID	CONTIG LENGTH (bp)	AVERAGE COVERAGE	FOUND IN *S. canicula HoxB* (GenBank FQ032659)
EH-HoxB_CNE1	35735-35785	N/A	N/A	0	N
EH-HoxB_CNE2	49671-49723	N/A	N/A	0	N
EH-HoxB_CNE3	156398-156447	N/A	N/A	0	N
EH-HoxB_CNE4	160340-160390	N/A	N/A	0	N
EH-HoxB_CNE5	181454-181516	N/A	N/A	0	N
EH-HoxB_CNE6	190077-190158	N/A	N/A	0	N
EH-HoxB_CNE7	190212-190267	N/A	N/A	0	N
EH-HoxB_CNE8	190544-190645	LER_WGS_1_CONTIG_1147140	954	21.21	Y
EH-HoxB_CNE9	190708-190870	LER_WGS_1_CONTIG_65401	1234	17.48	Y
EH-HoxB_CNE10	192593-192646	N/A	N/A	0	N
EH-HoxB_CNE11	192680-192734	N/A	N/A	0	N
EH-HoxB_CNE12	193469-193643	LER_WGS_1_CONTIG_80692	1262	18.34	Y
EH-HoxB_CNE13	197652-197901	LER_WGS_1_CONTIG_1644194	1444	14.48	Y
EH-HoxB_CNE14	202282-202354	LER_WGS_1_CONTIG_40659	4828	21.05	Y
EH-HoxB_CNE15	202360-202424	LER_WGS_1_CONTIG_40659	4828	21.05	Y
EH-HoxB_CNE16	205908-206137	LER_WGS_1_CONTIG_690334	1709	19.92	Y
EH-HoxB_CNE17	208399-208471	N/A	N/A	0	N
EH-HoxB_CNE18	209190-209432	LER_WGS_1_CONTIG_1630427	945	21.5	Y
EH-HoxB_CNE19	213652-213708	N/A	N/A	0	N
EH-HoxB_CNE20	216967-217047	LER_WGS_1_CONTIG_1641394	2611	16.66	N
EH-HoxB_CNE21	217050-217199	LER_WGS_1_CONTIG_2672030	4613	21.67	N
EH-HoxB_CNE22	219198-219267	LER_WGS_1_CONTIG_46923	4238	19.63	Y
EH-HoxB_CNE23	220073-220157	LER_WGS_1_CONTIG_46923	4238	19.63	Y
EH-HoxB_CNE24	220574-220732	LER_WGS_1_CONTIG_46923	4238	19.63	Y
EH-HoxB_CNE25	221225-221279	N/A	N/A	0	N
EH-HoxB_CNE26	221989-222076	N/A	N/A	0	N
EH-HoxB_CNE27	223237-223366	N/A	N/A	0	Y
EH-HoxB_CNE28	224065-224209	N/A	N/A	0	Y
EH-HoxB_CNE29	224602-224694	LER_WGS_1_CONTIG_2711623	1650	17.85	Y
EH-HoxB_CNE30	226768-226963	LER_WGS_1_CONTIG_65572	1605	16.76	Y
EH-HoxB_CNE31	227053-227142	N/A	N/A	0	N
EH-HoxB_CNE32	227280-227330	N/A	N/A	0	N
EH-HoxB_CNE33	227374-227539	LER_WGS_1_CONTIG_65572	1605	16.76	Y
EH-HoxB_CNE34	227660-227731	N/A	N/A	0	Y
EH-HoxB_CNE35	231321-231371	N/A	N/A	0	N
EH-HoxB_CNE36	233653-233762	LER_WGS_1_CONTIG_2704880	5703	20.57	Y
EH-HoxB_CNE37	237771-237939	LER_WGS_1_CONTIG_73733	1382	19.76	Y
EH-HoxB_CNE38	240670-240722	N/A	N/A	0	N
EH-HoxB_CNE39	240781-240828	N/A	N/A	0	N
EH-HoxB_CNE40	256939-257009	LER_WGS_1_CONTIG_1644638	1401	23.53	N
EH-HoxB_CNE41	260429-260541	LER_WGS_1_CONTIG_2666672	2777	19.89	Y
EH-HoxB_CNE42	263890-264014	LER_WGS_1_CONTIG_69688	1326	16.85	Y
EH-HoxB_CNE43	269862-269906	N/A	N/A	0	N
EH-HoxB_CNE44	284996-285068	N/A	N/A	0	N
EH-HoxB_CNE45	296206-296253	N/A	N/A	0	N
EH-HoxB_CNE46	297114-297213	N/A	N/A	0	N
EH-HoxB_CNE47	300029-300182	LER_WGS_1_CONTIG_27393	269	12.73	N
EH-HoxB_CNE48	334393-334451	N/A	N/A	0	N

HoxC

CSE ID	*C. milii HoxC* COORDINATES (bp) in GenBank FJ824600	*L. erinacea* CONTIG ID	CONTIG LENGTH (bp)	AVERAGE COVERAGE
EH-HoxC_CNE1	30953-31010	N/A	N/A	0
EH-HoxC_CNE2	48887-48943	N/A	N/A	0
EH-HoxC_CNE3	97125-97193	N/A	N/A	0
EH-HoxC_CNE4	101967-102016	N/A	N/A	0
EH-HoxC_CNE5	104687-104771	N/A	N/A	0
EH-HoxC_CNE6	114522-114575	N/A	N/A	0
EH-HoxC_CNE7	115490-115542	N/A	N/A	0
EH-HoxC_CNE8	122769-123003	N/A	N/A	0
EH-HoxC_CNE9	123694-123831	N/A	N/A	0
EH-HoxC_CNE10	126451-126507	N/A	N/A	0
EH-HoxC_CNE11	134516-134647	N/A	N/A	0
EH-HoxC_CNE12	134683-134740	N/A	N/A	0
EH-HoxC_CNE13	138991-139158	N/A	N/A	0
EH-HoxC_CNE14	139914-139967	N/A	N/A	0
EH-HoxC_CNE15	145543-145596	N/A	N/A	0
EH-HoxC_CNE16	161707-161892	N/A	N/A	0
EH-HoxC_CNE17	168967-169107	N/A	N/A	0
EH-HoxC_CNE18	176850-177056	N/A	N/A	0
EH-HoxC_CNE19	191470-191670	N/A	N/A	0
EH-HoxC_CNE20	194758-194909	N/A	N/A	0
EH-HoxC_CNE21	195610-195660	N/A	N/A	0
EH-HoxC_CNE22	195671-195737	N/A	N/A	0
EH-HoxC_CNE23	209429-209487	N/A	N/A	0
EH-HoxC_CNE24	213152-213285	N/A	N/A	0
EH-HoxC_CNE25	214945-215002	N/A	N/A	0

HoxD

CSE ID	*C. milii HoxD* COORDINATES (bp) in GenBank FJ824601	*L. erinacea* CONTIG ID	CONTIG LENGTH (bp)	AVERAGE COVERAGE	FOUND IN *S. canicula HoxD* (GenBank FQ032660)
EH-HoxD_CNE1	6275-6648	LER_WGS_1_CONTIG_47219	1820	33.13	N
EH-HoxD_CNE2	6779-7199	LER_WGS_1_CONTIG_47219	1820	33.13	N
EH-HoxD_CNE3	10510-10579	N/A	N/A	0	N
EH-HoxD_CNE4	11784-12148	LER_WGS_1_CONTIG_59346	1137	18.36	N
EH-HoxD_CNE5	14933-14981	N/A	N/A	0	N
EH-HoxD_CNE6	15893-16191	LER_WGS_1_CONTIG_1028740	1594	25.95	N
EH-HoxD_CNE7	16320-16738	LER_WGS_1_CONTIG_1028740	1594	25.95	N
EH-HoxD_CNE8	16882-17035	LER_WGS_1_CONTIG_1028740	1594	25.95	N
EH-HoxD_CNE9	26375-26424	N/A	N/A	0	N
EH-HoxD_CNE10	26601-26837	LER_WGS_1_CONTIG_209778	1602	20.67	N
EH-HoxD_CNE11	30541-30659	LER_WGS_1_CONTIG_1633014	1322	24.25	N
EH-HoxD_CNE12	33977-34025	N/A	N/A	0	N
EH-HoxD_CNE13	42667-43130	LER_WGS_1_CONTIG_642692	3072	23.61	N

EH-HoxD_CNE14	43219-43455	LER_WGS_1_CONTIG_642692	3072	23.61	N
EH-HoxD_CNE15	46629-46981	LER_WGS_1_CONTIG_65194	914	27.5	N
EH-HoxD_CNE16	50816-50865	N/A	N/A	0	N
EH-HoxD_CNE17	51032-51078	N/A	N/A	0	N
EH-HoxD_CNE18	51379-51586	N/A	N/A	0	N
EH-HoxD_CNE19	52504-52551	N/A	N/A	0	N
EH-HoxD_CNE20	55044-55102	N/A	N/A	0	N
EH-HoxD_CNE21	57941-58103	LER_WGS_1_CONTIG_1597922	4490	28.44	N
EH-HoxD_CNE22	58592-59012	LER_WGS_1_CONTIG_1597922	4490	28.44	N
EH-HoxD_CNE23	63146-63205	N/A	N/A	0	N
EH-HoxD_CNE24	66012-66056	N/A	N/A	0	N
EH-HoxD_CNE25	70244-70297	N/A	N/A	0	N
EH-HoxD_CNE26	80102-80467	LER_WGS_1_CONTIG_79769	1895	23.15	N
EH-HoxD_CNE27	80947-81145	LER_WGS_1_CONTIG_79769	1895	23.15	N
EH-HoxD_CNE28	88503-89096	LER_WGS_1_CONTIG_70875	1822	19.33	N
EH-HoxD_CNE29	89174-89642	LER_WGS_1_CONTIG_2671970	1716	29.53	N
EH-HoxD_CNE30	89773-89826	N/A	N/A	0	N
EH-HoxD_CNE31	90170-90333	LER_WGS_1_CONTIG_2671970	1716	29.53	N
EH-HoxD_CNE32	104547-104596	N/A	N/A	0	N
EH-HoxD_CNE33	106717-107570	LER_WGS_1_CONTIG_1640383	1712	20.88	N
EH-HoxD_CNE34	107596-107650	N/A	N/A	0	N
EH-HoxD_CNE35	110021-110224	LER_WGS_1_CONTIG_690815	1583	12.3	N
EH-HoxD_CNE36	113868-113944	N/A	N/A	0	N
EH-HoxD_CNE37	114062-114152	N/A	N/A	0	N
EH-HoxD_CNE38	124175-124313	LER_WGS_1_CONTIG_653839	4125	21.78	N
EH-HoxD_CNE39	124618-124680	N/A	N/A	0	N
EH-HoxD_CNE40	124749-124820	N/A	N/A	0	N
EH-HoxD_CNE41	125433-125484	N/A	N/A	0	N
EH-HoxD_CNE42	128694-129070	LER_WGS_1_CONTIG_86891	975	14.89	N
EH-HoxD_CNE43	149246-149742	LER_WGS_1_CONTIG_1614443	6851	20.65	Y
EH-HoxD_CNE44	149998-150558	LER_WGS_1_CONTIG_1614443	6851	20.65	Y
EH-HoxD_CNE45	150587-150801	LER_WGS_1_CONTIG_1614443	6851	20.65	Y
EH-HoxD_CNE46	150825-151115	LER_WGS_1_CONTIG_1614443	6851	20.65	Y
EH-HoxD_CNE47	151801-151948	LER_WGS_1_CONTIG_1614443	6851	20.65	Y
EH-HoxD_CNE48	152712-153315	LER_WGS_1_CONTIG_1614443	6851	20.65	Y
EH-HoxD_CNE49	156657-156740	N/A	N/A	0	Y
EH-HoxD_CNE50	158212-158436	LER_WGS_1_CONTIG_35761	4347	22.34	Y
EH-HoxD_CNE51	158467-158600	LER_WGS_1_CONTIG_35761	4347	22.34	Y
EH-HoxD_CNE52	176042-176090	N/A	N/A	0	N
EH-HoxD_CNE53	184726-184813	LER_WGS_1_CONTIG_45979	8323	22.75	Y
EH-HoxD_CNE54	187625-187702	N/A	N/A	0	N
EH-HoxD_CNE55	188475-188580	N/A	N/A	0	N
EH-HoxD_CNE56	190523-190613	N/A	N/A	0	N
EH-HoxD_CNE57	190911-190996	N/A	N/A	0	Y
EH-HoxD_CNE58	197246-197342	N/A	N/A	0	N
EH-HoxD_CNE59	197353-197431	N/A	N/A	0	N
EH-HoxD_CNE60	198205-198381	N/A	N/A	0	N
EH-HoxD_CNE61	204979-205063	N/A	N/A	0	N
EH-HoxD_CNE62	205709-205791	N/A	N/A	0	N
EH-HoxD_CNE63	206425-206532	LER_WGS_1_CONTIG_75394	2314	23.37	N
EH-HoxD_CNE64	209283-209352	N/A	N/A	0	N
EH-HoxD_CNE65	209380-209447	N/A	N/A	0	N
EH-HoxD_CNE66	212739-212828	N/A	N/A	0	N
EH-HoxD_CNE67	215437-215503	LER_WGS_1_CONTIG_1618096	4177	21.42	Y
EH-HoxD_CNE68	218975-219097	LER_WGS_1_CONTIG_65489	2323	21.43	Y
EH-HoxD_CNE69	219106-219178	N/A	N/A	0	N
EH-HoxD_CNE70	224102-224151	N/A	N/A	0	N
EH-HoxD_CNE71	227740-227916	LER_WGS_1_CONTIG_57213	3255	23.52	Y
EH-HoxD_CNE72	231818-231985	LER_WGS_1_CONTIG_2672030	4613	21.67	Y
EH-HoxD_CNE73	232681-232748	N/A	N/A	0	N
EH-HoxD_CNE74	233725-233775	N/A	N/A	0	N
EH-HoxD_CNE75	235032-235164	LER_WGS_1_CONTIG_2672030	4613	21.67	Y
EH-HoxD_CNE76	235754-235807	LER_WGS_1_CONTIG_56573	3681	24.87	Y
EH-HoxD_CNE77	238639-238981	LER_WGS_1_CONTIG_56573	3681	24.87	Y
EH-HoxD_CNE78	241070-241219	LER_WGS_1_CONTIG_1171989	1371	19.71	Y
EH-HoxD_CNE79	241331-241413	LER_WGS_1_CONTIG_1171989	1371	19.71	Y
EH-HoxD_CNE80	241608-241669	LER_WGS_1_CONTIG_1171989	1371	19.71	Y
EH-HoxD_CNE81	242635-242712	LER_WGS_1_CONTIG_34637	2167	19.88	Y
EH-HoxD_CNE82	242907-243046	N/A	N/A	0	N
EH-HoxD_CNE83	244576-244692	N/A	N/A	0	Y
EH-HoxD_CNE84	244853-244980	LER_WGS_1_CONTIG_2689503	1723	23.47	Y
EH-HoxD_CNE85	245026-245118	N/A		0	N
EH-HoxD_CNE86	246273-246443	LER_WGS_1_CONTIG_1623373	1497	22.09	Y
EH-HoxD_CNE87	247193-247267	N/A	N/A	0	Y
EH-HoxD_CNE88	250419-250500	N/A	N/A	0	N
EH-HoxD_CNE89	250514-250632	N/A	N/A	0	Y
EH-HoxD_CNE90	251218-251337	LER_WGS_1_CONTIG_79523	832	17.16	Y
EH-HoxD_CNE91	256792-256927	N/A	N/A	0	Y
EH-HoxD_CNE92	256956-257086	LER_WGS_1_CONTIG_42923	2851	23.4	Y
EH-HoxD_CNE93	264147-264259	LER_WGS_1_CONTIG_2679883	1426	22.08	Y
EH-HoxD_CNE94	268069-268124	N/A	N/A	0	N
EH-HoxD_CNE95	271794-271846	N/A	N/A	0	N
EH-HoxD_CNE96	295972-296071	N/A	N/A	0	N
EH-HoxD_CNE97	296137-296443	LER_WGS_1_CONTIG_51285	1880	16.6	N
EH-HoxD_CNE98	299405-299670	N/A	N/A	0	N
EH-HoxD_CNE99	303709-303909	LER_WGS_1_CONTIG_1617891	2974	19.1	N
EH-HoxD_CNE100	309969-310108	N/A	N/A	0	N
EH-HoxD_CNE101	313224-313613	LER_WGS_1_CONTIG_41109	2077	26.98	N
EH-HoxD_CNE102	313806-314055	LER_WGS_1_CONTIG_41109	2077	26.98	N
EH-HoxD_CNE103	314075-314283	N/A	N/A	0	N
EH-HoxD_CNE104	317278-317324	N/A	N/A	0	N
EH-HoxD_CNE105	321086-321252	LER_WGS_1_CONTIG_1617352	4761	28.84	N