

Gene Cluster ID	Atp6v1C
Drosophila melanogaster gene	Vha44 (CG8048)
Predicted function	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C

Atp6v1C CLEAR element conservation:

	5'	5' / intron1
Dmel\	ACACTGGCGCCA TCACGTGA ATCTCGTGGGCA	ATCGAAAAACTG TCATGTGA CTGTCAACGAAC
Mdom\	CTGTGTCACTTA TCACGTGA ATTTCTTCTCCT	GATTCAAGCTTG TCATGTGA CTGTCATCTGTC
Ccap\	CTGCGAATGTTA TCACGTGA ATTTTGTACAT	GACCCAAAGCTG TCATGTGA CTGTCATCAAAC
Gmor\	-----	CGATCC AAACTGTCATGTGA CTTATC ATCAAA
Llon\	-----	TAGG AAAAATTTTCATGTGA AA TCATTCTT
Ppap\	-----	AAGATGGT ATGTTCATGTGA CG CAAAAACT
Agam\	CACCGAAAGCAATC ACGTGA ATACGTTGTAAT	AAAAGGTAGATGTCATGTGA CTTATCAGTTCA
Aara\	CACCGAAAGCAATC ACGTGA ATACGTTGTAAT	AAAAGGTAGATGTCATGTGA CTTATCAGTTCA
Aqan\	CACCGAAAGCAATC ACGTGA ATACGTTGTAAT	AAAAGGTAGATGTCATGTGA CTTATCAGTTCA
Achr\	CTCCGAAAGCAATC ACGTGA ATACGTTGTAAT	AAAAGGTAGATGTCATGTGA CTTATCAGTTCA
Aepi\	CTCCGAAAGCAATC ACGTGA ATACGTTGTAAT	AAAAGGTGGATGTCATGTGA CTTTATCAATT
Amin\	CTCCGAAAGCAATC ACGTGA ATACTGTAATGG	AAAAGGTAGACTGTCATGTGA CTTATCAGTTCA
Afun\	CTCCGAAAGCAATC ACGTGA ATACTGTAATGG	AAAAGGTAGATGTCATGTGA CTTATCAGTTCA
Aste\	CTCCGAAAGCAATC ACGTGA ATACTGTAAT	AAAAGGTAGATGTCATGTGA CTTATCAGTTCA
Adir\	GCT CGGAAGCAATC ACGTGA ATACTTGTAA ATG	AAAAGGTAGATGTCATGTGA CTTATCAGTTCA
Anil\	AGAGGCTT GCAATC ACGTGA ATAATATGGTAA	AAAAGGTAGATGTCATGTGA CTTATC ATTCA
Aalb\	ACGAGCCGCA ATC ACGTGA ATTCTGTGA AG	GAAAAGATAGTGT TCATGTGA CTTATC ACACAA
Adar\	CTTCTCGTGCAATC ACGTGA ATCGAATTCGGA	GAAAAGATAGTGT TCATGTGA CTTATC ACACAA
Aaeg\	ACG CAGAGACTGTC ACGTGA ATTGCAATGGTG	AGGAAAAATATGT TCATGTGA CTTATC ATTCAA
Cqui\	ACA CTTTGCTGT TCACGTGA ATTGCAATGGTG	TTGAAAAATATGT TCATGTGA CTTATC AGTCAA
Mdes\	TTATGTGA AA ATGTCACATGA CACAA ATGTTAA	CAATGGTAAATGT TCATGTGA CTCACA CAAAA
Bmor\	GGCAGCACAGCGTC ACGTGA ATTTTCATGGCG	
Msex\	GGCAGCACCGCATC ACGTGA ATTTTCATGGCG	
Dple\	GGCAGCACGGCATC ACGTGA ATTTTCATGGCG	
Hmel\	GGCAGCACAGCATC ACGTGA ATTTTCATGGCG	
Pxyl\	GGCAGCACCGCATC ACGTGA ATTTTCATGGCG	
Tcas\	CAACATGTC AAAATC ACGTGA AACCCCACTAAG	TGTCAATTGAAG TCACATGA TCGCCTTCAATT
Dpon\	ACAATTTGACAG TCACGTGA AAATAACGATCA	AGTTTAGTGTA GTCA TGTGA CTGTAGTAATGT
Amel\	AACT GTTC CAAAATC AGCTGA TATTCCTTTGG	ACACAATTGT TCATC ACGTGA ---GTC ACAATG
Bter\	TCTGGTTCCAAATC AGC CGATGAT TCTTTAG	ACACAATTGT TCATC ACGTGA ---GTC ACAATG
Bimp\	TCTGGTTCCAAATC AGCTGA TGATTCTTTAG	ACACAATTGT TCATC ACGTGA ---GTC ACAATG
Aflo\	AACGGTT CCAAATC AGCTGA TGATATCTCTT	ACACAATTGT TCATC ACGTGA ---GTC ACAATG
Mrot\	GACAA TTCC AAAGT CAGCTGA TGATACTGTGGC	ACACAATTGT TCATC ACGTGA ---GTC ACAATG
Lhum\	CAGGTTCTTG AGTC AGCTGA TAAAGTITTAAT	ACACAATTGT TCATC ACGTGA GAA GTCA CGATG
Cflo\	TCGAGTTACGAATC AGCTGA CAATATTCAGTT	ACATAATTGT TCATC ACGTGA GGAG TCACA ATG
Aech\	TCTTGATATGCCGTC AGCTGA CGGCACGGCGTT	CGCAATCTGT TCATC ACGTGA GCGG TCACA ATG
Acep\	TGCGTCGCGGAGTC AGCTGA CGGCACGGCGTT	CGCAATCTGT TCATC ACGTGA GTGG TCACA ATG
Pbar\	TGTACGCGGAGTC AGCTGA CAGCACAGCGTT	CACAATCTGT TCATC ACGTGA GTGG TCACA ATG
Sinv\	TGTC GCACGGAGTC AGCTGA CGGCACGGCGG	CACAATCTGT TCATC ACGTGA GCGG TCACA ATG
Hsal\	TCGGGCG CAAGTC AGCTGA CGGCACAGTGTG	ACACAATTGT TCATC ACGTGA GAGG TCACA ATG
Ngir\	-----	GCAGCACTCTTAT CACGTGA GAGCGTG CGTAT
Nlon\	-----	GCAGCACTCTTAT CACGTGA GAGCGTG CGTAT
Nvit\	-----	GCAGCACTCTTAT CACGTGA GAGCGTG CGTAT
Phum\	GGCAGCACATA ATC ACGTGA ATCGTTATGGTA	
Lful\	GTCA CCACCGAGTC ACGTGA ATTGTAATGAAT	
Dpul\	GGAA ACCTTAGAT TCACATGA CAGTCTCCTGAC	

Atp6v1C gene structure comparison:

Legend: UTR region, CDS region, INTRON

Order	Species	Atp6v1C orthologs	5'exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon3'	Extended CLEAR region	Position	bps from TSS
-------	---------	-------------------	--------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	--------	-----------------------	----------	--------------

Diptera (Drosophilidae)	<i>Drosophila_melanogaster</i>	DmelVha44-RA	148	3069	8	135	90	154	5873	187	64	99	62	256	62	98	59	139	70	99	1153	acactggcgccaTCACGTGAatctgtgggca atcgaaaactgTCATGTGAactgcaacgaac	5' intron1	-40 565						
Diptera (Muscidae)	<i>Musca_domestica</i>	MdomAtp6v1C	ND	ND	11	132	68	154	13312	187	164	99	68	256	71	98	70	145	2264	99	ND	ctgtgtcacttaTCACGTGAaatttctctctc gattcaagctgTCATGTGAactgcatctgtc	5' intron1?	ND ND						
Diptera (Tephritidae)	<i>Ceratitis_capitata</i>	CcapAtp6v1C	192	14748	11	132	ND	154	18961	187	2027	99	93	256	1441	98	106	142	562	99	ND	ctcgaaatgtaTCACGTGAaattttgtacat gacccaaagctgTCATGTGAactgcatcaaac	5' intron1	-115 670						
Diptera (Glossinidae)	<i>Glossina_morsitans</i>	GmorAtp6v1C TMP008641*	138	5966	11	132	72	154	10967	187	1138	99	82	256	77	98	56	139	81	99	239	cgatccaaactgTCATGTGAacttatcatcaaa	intron1	531						
Diptera (Psychodidae)	<i>Lutzomyia_longipalpis</i>	LlonAtp6v1C			ND	286			ND	187	85	99	1215	493					471	96	ND	taggaaaattTCATGTGAaattcatttctt	intron1	ND						
	<i>Phlebotomus_papatasi</i>	PpapAtp6v1C			ND	286 (132+154)			ND	187	ND	99	2288	493 (256+98+139)					ND	96	ND	aagatggtatgTCATGTGAactgcaaaaact	intron1	ND						
Diptera (Culicidae)	<i>Anopheles_gambiae</i>	AgamAtp6v1C AGAP005845	279	869	28	132	102	154	4455	187	74	99	71	256	75	98	72	136	77	96	345	caccgaaagcaaTCACGTGAatactgttgtaat aaaaggtagatgTCATGTGAacttacagttica	5'UTR intron1	25 658						
	<i>Anopheles_arabiensis</i>	AaraAtp6v1C			ND	132	101	154	4446	187	76	99	74	256	70	98	72	136	72	96	ND	caccgaaagcaaTCACGTGAatactgttgtaat aaaaggtagatgTCATGTGAacttacagttica	5'UTR? intron1?	>-1160 >-518						
	<i>Anopheles_quadriannulatus</i>	AqanAtp6v1C			ND	132	101	154	4415	187	80	99	73	256	71	98	72	136	74	96	ND	caccgaaagcaaTCACGTGAatactgttgtaat aaaaggtagatgTCATGTGAacttacagttica	5'UTR? intron1?	>-1155 >-514						
	<i>Anopheles_christyi</i>	AchrAtp6v1C			ND	132	85	154	4227	187	81	99	93	256	74	98	75	136	69	96	ND	ctcggaaagcaaTCACGTGAatactgttgtaat aaaaggtagatgTCATGTGAacttacagttica	5'UTR? intron1?	>-1166 >-514						
	<i>Anopheles_epiroticus</i>	AepiAtp6v1C			ND	135	69	154	4993	187	73	99	80	256	67	98	76	136	62	96	ND	ctcggaaagcaaTCACGTGAatactgttgtaat aaaaggtgtagTCATGTGAactttatcaatt	5'UTR? intron1?	>-1135 >-504						
	<i>Anopheles_minimus</i>	AminAtp6v1C			ND	132	84	154	4210	187	65	99	75	256	72	98	69	136	63	96	ND	ctcggaaagcaaTCACGTGAatactgtaatgg aaaggtagactgTCATGTGAacttacagttica	5'UTR? intron1?	>-1137 >-485						
	<i>Anopheles_funestus</i>	AfunAtp6v1C			ND	132	93	154	4498	187	73	99	79	256	87	98	74	136	72	96	ND	ctcggaaagcaaTCACGTGAatactgtaatgg aaaaggtagactgTCATGTGAacttacagttica	5'UTR? intron1?	>-1108 >-469						
	<i>Anopheles_stephensi</i>	AsteAtp6v1C			ND	132	77	154	ND	187	ND	99	65	256	73	98	90	136	ND	96	ND	ctcggaaagcaaTCACGTGAatactgttgtaat aaaaggtagactgTCATGTGAacttacagttica	5'UTR? intron1?	ND						
	<i>Anopheles_dirus</i>	AdirAtp6v1C			ND	132	68	154	3912	187	68	99	73	256	57	98	347	136	71	96	ND	gtctggaaagcaaTCACGTGAatactgtgtaatg aaaaggtagactgTCATGTGAacttacagttica	5'UTR? intron1?	>-1111 >-478						
	<i>Anopheles_nili</i>	AnilAtp6v1C			ND	132	78	154	ND	187	86	99	69	256	66	98	73	136	67	96		agagccttgcaaTCACGTGAataatattggttaa aaaaggtagactgTCATGTGAacttacatttca	5'UTR? intron1?	>-1108 >-500						
	<i>Anopheles_albimanus</i>	AalbAtp6v1C			ND	132	79	154	3054	187	76	99	75	256	67	98	76	136	77	96	ND	acgagccgcgcaaTCACGTGAattctgtggaag gaaaagatagtgTCATGTGAacttacacacaa	5'UTR? intron1?	>-923 >-337						
	<i>Anopheles_darlingi</i>	AdarAtp6v1C			ND	132	78	154	3189	187	73	99	70	256	73	98	71	136	77	96	ND	cttcctgtcaaTCACGTGAatcgaaatggga gaaaagatagtgTCATGTGAacttacacacaa	5'UTR? intron1?	>-946 >-367						
	<i>Aedes_aegypti</i>	AaegAtp6v1C AAEL005173	215	7535	60	132	65	154	9933	187	56	99	55	256	62	98	66	136	10450	96	593	acgcagagactgTCACGTGAaattgcaattggtg aggaaaaaatgTCATGTGAacttacatcnaa	5'UTR intron1	7 632						
	<i>Culex_pipiens_qui.</i>	CquiAtp6v1C CPU002067			58	132	55	154	2686	187	63	99	220	256 (124+132)	56	98	64	136	64	96	ND	acacttggctgTCACGTGAaattgcaattggtg ttgaaaaatgTCATGTGAacttacatcnaa	5'UTR? intron1?	ND ND						
Diptera (Cecidomyiidae)	<i>Mayetiola_destructor</i>	MdesAtp6v1C	121	3731	186	132	ND	154	14525	187	100	99	76	124	96	132	173	98	128	136	721	96	ND	ttatgtgaaatgTCACATGAcacaaatgttaa caatggtgaaatgTCATGTGAactcaacacaaa	5'UTR? intron1	-66 949				
Lepidoptera	<i>Bombyx_mori</i>	BmorAtp6v1C			250	129	1316	102	744	52	7172	98	1143	89	556	99	1370	124	853	132	1681	98	1369	139	940	99	31	ggcagcacagcTCACGTGAattttcatggcg atcgttaacctTCACATGAagaacagctgtg	5' intron1	-286 1267
	<i>Manduca sexta</i>	MsexAtp6v1C			ND	129	1581	102	328	52	ND	98	266	89	267	99	99	124	ND	132	321	98	250	139	505	96	ND	ggcagcacccgaTCACGTGAattttcatggcg	5'	>-541
	<i>Danaus_plexippus</i>	DpleAtp6v1C			ND	129	996	102	427	52	3233	98	346	89	531	99	415	124	339	132	286	98	586	139	251	93	ND	ggcagcacggcaTCACGTGAattttcatggcg	5'	>-537
	<i>Heliconius_melpomene</i>	HmelAtp6v1C			ND	129	853	102	481	52	5651	98	874	89	240	99	120	124	837	132	321	98	425	139	1281	96	ND	ggcagcacagcaTCACGTGAattttcatggcg	5'	>-544
	<i>Plutella_xylostella</i>	PxylAtp6v1C			ND	129	414	102	430	52	6248	98	363	89	513	99	508	124	711	132	596	98	693	136	721	96	ND	ggcagcacccgaTCACGTGAattttcatggcg	5'	>-568
Strepsiptera	<i>Mengenilla_moldrzyki</i>	MmolAtp6v1C			ND	132	ND	154	1478	187	ND	99	59	256	994	240			52	96	ND	ND								
Coleoptera	<i>Tribolium_castaneum</i>	TeasAtp6v1C LOC661302	81	414	33	132	46	154	2617	187	50	355 (99+124+132)				54	98	75	136	49	99	572	caaatgtcaaaTCACGTGAaaccccaactaag tgtcaattgaagTCACATGAatgccttcaatt	5'UTR? intron1	1 418					
	<i>Dendroctonus_ponderosae</i>	DponAtp6v1C	78	601	46	132	871	154	ND	187	84	99	65	124	69	132	588	98	61	136	513	99	182	acaattgacagTCACGTGAaaaatacagcata agtttagttagTCATGTGAactgtagtaagt	5'UTR? intron1	-5 497				
Hymenoptera	<i>Apis_mellifera</i>	AmelAtp6v1C	170	1039	36	286			6348	187	120	99	71	209	149	145	108	145	105	96	ND	aactgttccaaaTCAGCTGAatttcccttgg acacaattgtcaTCACGTGAaggtcacaatgt	5' 5'UTR	-22 42						

[illegible]