Supplementary material

Energy biogenesis: one key for coordinating two genomes

Marco Sardiello[1,2], Gaetano Tripoli[1], Antonio Romito[1], Crescenzio Minervini[1], Luigi Viggiano[1], Corrado Caggese[1] and Graziano Pesole[3]

[1]Dipartimento di Anatomia Patologica e di Genetica, Sezione di Genetica, Università di Bari, via Amendola 165/A, 70126 Bari, Italy
[2]Telethon Institute of Genetics and Medicine (TIGEM), via Pietro Castellino 111, 80131, Napoli, Italy
[3]Dipartimento di Scienze Biomolecolari e Biotecnologie, Università di Milano, via Celoria 26, 20133 Milano, Italy

Table 1. ***Drosophila melanogaster*** **nuclear genes for mitochondrial proteins containing NRG sites**

Protein	Gene	NRG element sequence	Position[a] (bp)	PatSearch score
Complex I: NADH ubiquinone oxidoreductase				
acyl carrier protein	*mtacp1*	`tttacataat`	219	8.98
B8 subunit	*CG15434*	`attatataat`	183	9.83
B12 subunit	*CG10320*	`cttatgtaat`	175	8.6
B14 subunit	*CG7712*	`gttatgtcat`	234	7.39
B14.5a subunit	*CG3621*	`gttatataag`	226	8.13
B14.5b subunit	*CG12400*	`gtgacgtaac`	216	7.71
B15 subunit	*CG12859*	`gttatgtcat`	222	7.39
B16.6 subunit	*CG3446*	`gttatgcaaa`	51	5.81
B17 subunit	*CG13240*	`attatacaat`	16	6.74
		`attatgtaat`	353	10.22
B17.2 subunit	*CG3214*	`cttatataag`	175	6.58
B18 subunit	*CG5548*	`tttacgtaat`	68	9.37
B22 subunit	*CG9306*	`attatataat`	273	9.83
13 kDa-A subunit	*CG8680*	`attacgcaaa`	244	6.27
13 kDa-B subunit	*CG6463*	`gttatgtcat`	215	7.39
15 kDa subunit	*CG11455*	`gttacgtcac`	198	7.71
		`attgtgtaaa`	661	5.88
18 kDa subunit	*CG12203*	`tttacgtaat`	52	9.37
19 kDa subunit	*CG3683*	`gttatataac`	364	9.68
20 kDa subunit	*CG9172*	`gttacataaa`	57	8.9
23 kDa subunit	*ND23*	`gttatgtaag`	221	8.52

		cttatgtcat	231	5.84
24 kDa subunit	*CG5703*	gttatataat	443	9.76
30 kDa subunit	*CG12079*	attacataaa	43	8.98
39 kDa subunit	*CG6020*	attatgtaat	210	10.22
		atgacgtaaa	779	6.61
42 kDa subunit	*ND42*	gtgatgtaac	224	7.32
49 kDa subunit	*CG1970*	cttacataat	206	8.6
51 kDa subunit	*CG9140*	attatgtaat	240	10.22
75 kDa subunit	*ND75*	cttatgtaat	260	8.6
ASHI subunit	*CG3192*	tttacgtaac	316	9.29
MLRQ subunit	*CG32230*	attgcataac	184	7.05
MNLL subunit	*CG18624*	attacgtaat	205	10.61
PDSW subunit	*Pdsw*	attacataat	363	10.22
		atgacataac	371	7.39
SGDH subunit	*CG9762*	attacgcaac	310	7.44
Complex II: Succinate dehydrogenase				
flavoprotein	*Scs-fp*	gttatacaac	358	6.59
		gttatgtaat	458	10.15
iron-sulfur protein	*SdhB*	attatataag	299	8.21
		gttatgtaaa	406	8.9
cytochrome b560	*CG6666*	gttacgcaat	729	7.44
cytochrome b560, small subunit	*CG10219*	attatataac	200	9.76
Complex III: Ubiquinol-cytochrome C oxidoreductase				
cytochrome c1, heme protein	CG4769	attatgtaac	424	10.15
core protein 1	*CG3731*	attacgtaat	395	10.61
core protein 2	*CG4169*	gttacgtaaa	383	9.29
14 kDa subunit	*CG3560*	tttacgtaaa	53	8.13
		attacgtaaa	184	9.37
11 kDa subunit	*Ucrh*	tttacatcat	246	6.22
		tttatgcaac	278	5.81
iron-sulfur subunit	*RFeSP*	attacataat	251	10.22
		cttacgtaat	359	8.98
ubiquinone-binding protein QP-C	*CG7580*	attacataac	515	10.15
7.2 kDa subunit	*ox*	gttatgtaag	241	8.52
6.4 kDa subunit	*CG14482*	attgcataac	182	7.05
Complex IV: Cytochrome c oxidase				
subunit IV	CG10664	attacacaac	218	7.05
		gttacataac	387	10.07
subunit Va	*CoVa*	gttacgtaac	730	10.46
subunit Vb	*CG11015*	attgcataat	271	7.13
		gttatgtaat	379	10.15
		tttatgtaac	540	8.9

subunit VIa	*CG17280*	gttatgtaat	213	10.15
subunit VIb	*CG18809*	gttacataac	295	10.07
		gttacataag	440	8.52
subunit VIc	*cype*	cttacgtaac	209	8.91
subunit VIIa	*CG9603*	attacgtaag	214	8.98
subunit VIIc	*CG2249*	gttacataat	191	10.15
Complex V: F_0-F_1 ATP synthase				
alpha subunit	*blw*	gttacgcaat	419	7.44
		gttacataat	463	10.15
		attacataat	573	10.22
		cttacataat	794	8.6
beta subunit	*ATPsyn-β*	tttacataac	50	10.15
gamma subunit	*ATPsyn-γ*	gttacataaa	63	8.9
delta subunit	*CG2968*	gttacataat	225	10.15
		gttatatcac	552	6.93
epsilon subunit	*sun*	attacataac	186	10.15
		attatataaa	394	8.59
b subunit	*ATPsyn-b*	tttatgtaat	252	8.98
d subunit	*ATPsyn-d*	attacataat	213	10.22
	CG3321	attgcataaa	131	5.88
		attacataac	244	10.15
		attatgtaaa	719	8.98
f subunit	*CG4692*	gttacgtaat	214	10.53
		tttacataat	274	8.98
g subunit	*CG6105*	attatgtaat	210	10.22
oligomicin sensitivity-conferring protein	Oscp	gttatgtaat	252	10.15
Cf6 subunit	*ATPsyn-Cf6*	attacataac	201	10.15
lipid-binding protein	*CG1746*	gttacgtaac	481	10.46
		attacataat	752	10.22
Other oxidative phosphorylation				
cytochrome c	*Cyc-c-p*	gtgacgtaaa	535	6.54
		tttacgtaac	613	9.29
		gttacataaa	695	8.90
electron transfer protein, α chain	*walrus*	attatataaa	565	8.59
electron transfer protein, β chain	*CG7834*	tttgcgtaat	232	6.27
electron transfer protein - ubiquinone oxidoreductase	*CG12140*	attacataac	420	10.15
		attatacaat	472	6.74
SCO1	*CG8885*	tttacgtaaa	73	8.13
protoheme IX farnesyltransferase	*CG5037*	attacgcaac	235	7.44
cytochrome oxidase biogenesis protein OXA1	*CG6404*	gtgacataac	426	7.32
copper chaperone	*CG9065*	attatataaa	49	8.59
Tricarboxylic-acid pathway				
fumarase	*CG4094*	cttatgtaat	258	8.6

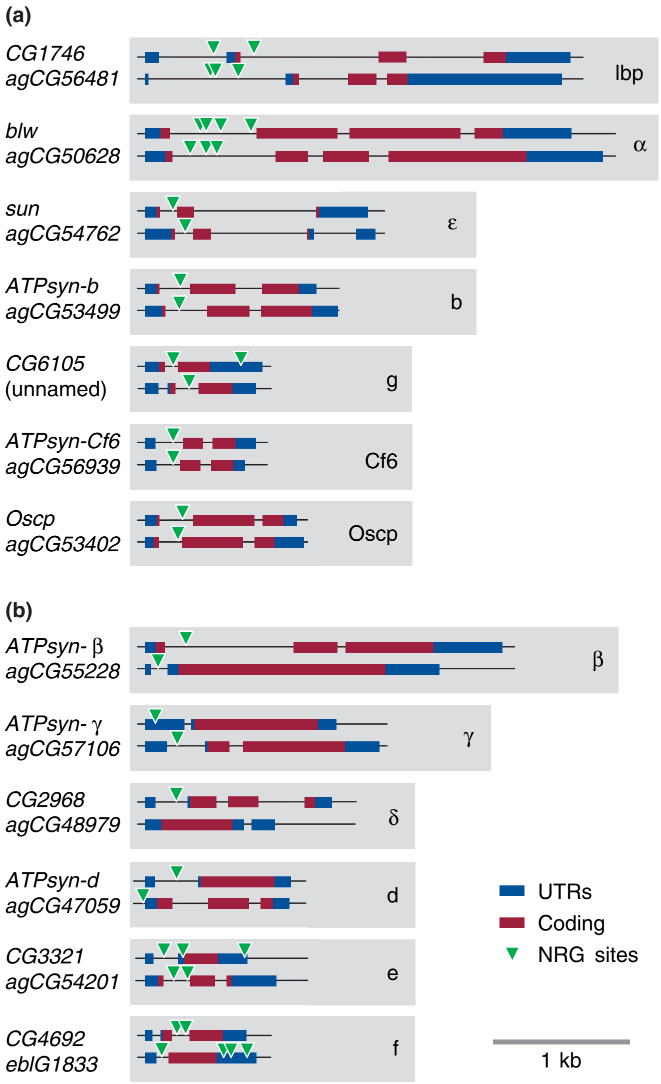
		attacacaat	387	7.13
malate dehydrogenase	*CG7998*	tttacgtaac	231	9.29
citrate synthase	*CG3861*	attacataat	586	10.22
aconitate idratase	Acon	tttacataat	231 (mRNA 2)	8.98
		ttgacgtaat	492 (mRNA 2)	6.61
isocitrate dehydrogenase α	*CG12233*	attacgcaaa	395	6.27
		tttacataac	618	8.9
isocitrate dehydrogenase β	*CG6439*	cttacataaa	417	7.36
isocitrate dehydrogenase γ	*CG5028*	attacataat	246	10.22
isocitrate dehydrogenase [NADP]	*CG7176*	cttatataaa	160	6.97
oxoglutarate dehydrogenase, E1 component	*CG11661*	cttatgtaat	351	8.6
oxoglutarate dehydrogenase, E2 component	*CG5214*	atgacataag	0	5.84
		attatgtaat	362	10.22
succinyl-CoA ligase [GDP-forming] α	Scsalpha	attatataac	228	9.76
succinyl-CoA ligase [GDP-forming] β	Sucb	tttatgtaat	448	8.98
succinyl-CoA ligase [ADP-forming] β	*CG11963*	tttacataat	260	8.98
Pyruvate metabolism				
NADP-dependent malic enzyme	*Mdh*	attatgtaag	357	8.6
pyruvate carboxylase	*CG1516*	attgcataat	434	7.13
pyruvate dehydrogenase, E1 component, α chain	*CG7010*	gttacataat	687	10.15
pyruvate dehydrogenase, E1 component, β chain	*CG11876*	attacgtaaa	412	9.37
pyruvate dehydrogenase, E3 component	*CG7430*	gttatgcaat	373	7.05
Amino acid metabolism				
2-amino-3-ketobutyrate coenzyme A ligase	*CG10361*	tttacataat	346	8.98
delta-1-pyrroline-5-carboxylate dehydrogenase	*CG7145*	attgcgtaat	239	7.51
glycine cleavage system, H protein	*ppl*	tttatgcaat	329	5.88
2-oxoisovalerate dehydrogenase, E1 component, β subunit	CG17691	gttatgtcat	35	7.39
2-oxoisovalerate dehydrogenase, E2 component	CG5599	gttatgtaat	219	10.15
		attgtataat	472	6.74
3-hydroxyisobutyrate dehydrogenase	CG15093	attacataac	182	10.15
methylcrotonyl-CoA carboxylase α chain	CG2118	attacataaa	219	8.98
methylcrotonyl-CoA carboxylase β chain	CG3267	ctgacgtaac	27	6.16
methylmalonate-semialdehyde dehydrogenase	CG17896	gttacgcaat	415	7.44
proline oxidase	slgA	gttgtataac	231	6.59
glutaryl-CoA dehydrogenase	CG9547	attatataaa	603	8.59
Nucleotide metabolism				
adenylate kinase isoenzyme 2	*Adk2*	attacataat	234	10.22
		gttatataag	374	8.13
GTP:AMP phosphotransferase	*Adk3*	tttatataac	99	8.52
Metabolism of complex lipids				
glycerol-3-phosphate dehydrogenase	*CG8256*	ctgacataat	313 (mRNA 1)	5.84
		ctgatgtaat	228 (mRNA 2)	5.84
3-ketoacyl-CoA thiolase	*Yip2*	gttatgtaaa	653	8.9

enoyl-CoA-hydratase	*CG6543*	attgcgtaac	312	7.44
acyl-CoA dehydrogenase short chain specific	*Arc42*	tttatgcaat	80	5.88
acyl-CoA dehydrogenase short chain specific	*CG4860*	tttacataac	369	8.9
acyl-CoA dehydrogenase short/branched chain specific	*CG3902*	gttacgtaac	473	10.46
3,2 trans-enoyl-CoA isomerase	*CG4594*	tttatataac	273	8.52
succinyl-CoA:3-ketoacid-CoA transferase	*CG1140*	cttacgtaac	496	8.91
Porphyrin metabolism				
cytochrome c-type heme lyase	*CG6022*	gttacgcaac	364	7.37
protoporphyrinogen oxidase	*CG5796*	gttatgtcac	1	7.32
delta 1-pyrroline-5-carboxylate synthetase	*CG7470*	tttatataag	251	6.97
protoporphyrinogen-oxidase	*CG5796*	gttatgtcac	2	7.32
NADPH:adrenodoxin oxidoreductase	dare	attatgcaaa	731	5.88
ferrochelatase	*ferrochelatase*	gtgatgtaaa	527	6.15
Sulfur metabolism				
sulfite oxidase	*CG7280*	gttacataac	263	10.07
Detoxification				
superoxide dismutase	*Sod2*	attacataaa	223	8.98
DNA synthesis				
DNA polymerase γ, subunit 2	*DNApol-γ35*	gttacgcaac	425	7.37
singol strand binding protein	*mtSSB*	tttatgtaat	509	8.98
Transcription				
RNA polymerase	*CG4644*	attatgcaac	73	7.05
transcription factor A	*TFAM*	tttacgtaaa	54	8.13
Ribosomal proteins				
L1	*mRpL1*	gttatgcaaa	17	5,81
L3	*mRpL3*	attatataag	10	8.21
L3a	*mRpL3a*	ttgatgtaac	206	6,15
L4	*mRpL4*	gttatgtaag	375	8,52
L5	*mRpL5*	tttacataat	181	8,98
		attacataaa	73	8,98
L10	*mRpL10*	gttgtgtaac	153	6,98
L11	*mRpL11*	tttgcataac	6	5.81
L13	*mRpL13*	attacatcac	326	7.39
L14	*mRpL14*	attgtataat	20	6,74
L21	*mRpL21*	gtgacataac	184	7,32
L22-24	*mRpL22-24*	attacgcaaa	32	6,27
		attacgtcag	724	6,23
L23	*mRpL23*	cttatgtaac	310	8,52
L45	*mRpL45*	tttatataat	64	8,59
L48	*mRpL48*	gtgatataac	24	6,93
L50	*mRpL50*	tttacgtcat	791	6,61
		tttacataaa	261	7,74

		tttacgtaaa	55	8,13
S6	*mRpS6*	gttacataaa	219	8.9
S7	*mRpS7*	attatataac	375	9,76
S9	*mRpS9*	tttgcataac	126	5.81
S11	*mRpS11*	gtgatgtaat	21	7.39
S14	*mRpS14*	cttacgtcac	337	6,16
S15	*bonsai*	gtgatataac	35	6.93
		tttacatcat	322	6.22
S17	*mRpS17*	gtgacataac	385	7,32
S28	*mRpS28*	tttacgtaaa	16	8,13
S30	*mRpS30*	atgacataac	15	7,39
S32	*mRpS32*	attgcataat	556	7,12
S33	*mRpS33*	gtgacataac	34	7,32
S35	*mRpS35*	tttatgtaat	382	8,98
Translation				
elongation factor Ts	*CG6412*	attatgtaat	725	10.22
elongation factor Tu	*EfTuM*	tttacgcaat	11	6.27
Protein stabilization				
10 kDa heat shock protein	*CG11267*	gttgtgtaat	225	7.05
10 kDa heat shock protein	*CG9920*	attatataat	43	9.83
60 kDa heat shock protein	*Hsp60*	attatataag	230	8.21
		gttacataaa	611	8.9
		tttacataat	739	8.98
stress-70 protein	*Hsc70-5*	attatataaa	393	8.59
		attatgtaag	576	8.6
GrpE protein homolog	*Roe1*	gtgacgtaat	216	7.78
Mitochondrial import receptor				
subunit TOM20	*CG7654*	tttatgtaaa	7	7.74
		attatataat	372	9.83
subunit TOM40	*Tom40*	atgacataac	384	7.39
subunit TOM70	*CG6756*	attacacaat	323	7.13
Mitochondrial import inner membrane translocase				
subunit TIM9 B	*Tim9b*	attgtataat	37	6.74
subunit TIM23	*Tim23*	gttacataat	297	10.15
subunit TIM44	*CG11779*	gttatataaa	497 (mRNA 1)	8.52
		attacataaa	268 (mRNA 2)	8.98
Proteolysis				
ATP-dependent CLP protease	*CG5045*	tttgcataac	120	5.81
processing peptidase, α subunit	*CG8728*	cttatgtaat	748	8.6
processing peptidase, β subunit	*CG3731*	attacgtaat	394	10.61
Carriers				
calcium-binding carrier protein	*Aralar1*	attatataaa	502	8.59
VDAC protein	*porin*	gttacgtaac	509 (mRNA 1)	10.46

		`gttatgcaac`	246 (mRNA 2)	6.98
		`gttatatcac`	393 (mRNA 2)	6.93
VDAC protein	*CG17137*	`attatataaa`	512	8.59
tricarboxylate transport protein	*CG31305*	`attgtgtaat`	226	7.13
		`gttatgtaat`	390	10.15
		`attgcataac`	753	7.05
ATP-binding cassette, sub-family B, member 6	*CG4225*	`attgcataat`	189	7.13
ATP-binding cassette, sub-family B, member 8	*CG1824*	`gttatgtcat`	253	7.39
ornithine transporter 1	*CG1628*	`tttatataat`	605	8.59
		`cttatataaa`	624	6.97
ADP,ATP carrier protein	SesB/Ant2	`gttatgtaat`	225 (mRNA 1)	10.15
		`cttacataac`	364 (mRNA 1)	8.52
		`gttatgtaaa`	247 (mRNA 2)	8.90
		`attacatcac`	445 (mRNA 2)	7.39
		`gttacataaa`	621 (mRNA 2)	8.90
phosphate carrier protein	*Mpcp*	`gttatgtaac`	264	10.07
		`gttgcataat`	720	7.05
phosphate carrier protein	*CG9090*	`attacgcaat`	233	7.51
		`gttatgtaaa`	279	8.9
		`attgtataat`	467	6.74
		`gttacatcat`	619	7.39
folate transporter	*CG8026*	`attacataag`	198	8.6
frataxin	*fh*	`attacatcat`	382	7.47
carnitine/acylcarnitine carrier protein	*CG3476*	`tttacgtaaa`	191	8.13
carnitine/acylcarnitine carrier protein	*colt*	`gtgacgtaat`	272	7.78
2-oxoglutarate/malate carrier protein	*CG1907*	`gttgtgtaat`	255	7.05
Cell death				
programmed cell death protein 8	*CG7263*	`gttacacaat`	220	7.05
Cell rescue				
thioredoxin-dependent peroxide reductase	Prx5037	`cttatataaa`	93	6.97
		`attacgtaat`	202	10.61
thioredoxin	*CG3719*	`attatgcaac`	350	7.05
Others				
apoptosis regulator bcl-2	debcl	`attacacaac`	449	7.05
mitochondrial uncoupling protein 4	Ucp4A	`gttatgtaaa`	273	8.9
complement component 1, Q subcomponent binding protein	*CG6459*	`gttatataac`	86	9.68
		`gttacataac`	349	10.07

[a] Position refers to the transcription-start site as annotated in the MitoDrome database [11].



TRENDS in Genetics

Figure 1. Comparison between *Drosophila melanogaster* and *Anopheles gambiae* genes encoding subunits of the F_0F_1-ATP synthase. The mosquito counterparts of the 13 known *Drosophila* genes were searched using the fly peptides against the *Anopheles* genome and EST collection in TBLASTN analysis, and the resulting sequences were aligned to determine the exon-intron structure of the mosquito genes. **(a)** Comparison of *Drosophila* (upper) with *Anopheles* (lower) homologous gene structures suggests that seven pairs of genes could be regarded as putative orthologs because the genes show comparable exon-intron structures, whereas six pairs of genes are pairs of putative paralogs **(b)**, in that the differences between the number and position of introns could be explained with retrotranscription events that occurred during the molecular evolution of both species. A pattern search analysis performed across the mosquito genes resulted in one or more matches to the NRG motif in all investigated genes with the exception of the ATP synthase delta subunit gene that shows no match. Coding sequences are in red, Untranlsated regions (UTRs) in blue. Green arrowheads indicate the position of NRG sites. The name of the encoded subunit is indicated for each pair of genes. Abbreviations: lbp, lipid binding protein; Oscp, oligomycin sensitivity-conferring protein.