Manuscript submitted to: Volume 1, 2013, 1-26.

AIMS Molecular Science DOI: 10.3934/molsci.2013.1.1

Received date 13 October 2013, Accepted date 27 November 2013, Published date 12 December 2013

Review

Implication of bidirectional promoters containing duplicated GGAA motifs of mitochondrial function-associated genes

Fumiaki Uchiumi [1,4] *, Makoto Fujikawa [2], Satoru Miyazaki [3] and Sei-ichi Tanuma [2,4]

[1] Department of Gene Regulation, Faculty of Pharmaceutical Sciences, Tokyo University of Science,Yamazaki 2641, Noda-shi, Chiba-ken 278-8510, Japan
[2] Department of Biochemistry, Faculty of Pharmaceutical Sciences, Tokyo University of Science, Yamazaki 2641, Noda-shi, Chiba-ken 278-8510, Japan
[3] Department of Bioinformatics, Faculty of Pharmaceutical Sciences, Tokyo University of Science, Yamazaki 2641, Noda-shi, Chiba-ken 278-8510, Japan
[4] Research Center for RNA Science, RIST, Tokyo University of Science, Yamazaki 2641, Noda-shi, Chiba-ken 278-8510, Japan

* **Correspondence:** E-mail: uchiumi@rs.noda.tus.ac.jp; Tel: +81-4-7121-3616; Fax: +81-4-7121-3608.

Abstract: Mitochondria are well known as the primary required organelle in all eukaryotic cells. They have their own mtDNA containing genes that encode tRNAs, rRNAs and a set of functional proteins required for energy (ATP) production. However, almost all (99%) of mitochondrial proteins are encoded by host nuclear genes. Therefore, expression of mitochondrial protein-encoding genes should be regulated similarly to genes that are present in the host nuclear chromosomes. Interestingly, from genomic database assisted surveillance, it was revealed that a lot of mitochondrial function associated protein-encoding genes are oppositely linked in a head-head manner. If the two head-head conjugated genes are regulated by the same transcription factor(s), their expression would be dependent on the direction of transcription machinery that contains RNA polymerase II to execute mRNA synthesis. In this article, we will focus on several examples of the mitochondrial and the partner gene sets and discuss putative functions of transcription factor binding elements in the bidirectional promoters of mitochondrial function-associated genes in chromosomes.

Keywords: bidirectional promoter; gene loop; GGAA-motif; interferon stimulated genes; mitochondria; TATA less promoter

1. Introduction

Molecular mechanisms of transcription from chromosomal DNA in eukaryotic cells have been well studied. It is widely known that transcription in eukaryotic cells is executed by three major RNA polymerases (pols) I, II, and III, which catalyze synthesis of the majority of ribosomal RNAs (rRNAs), mRNAs, and tRNAs, respectively. Among these, the molecular mechanism of transcription involving RNA pol II has been the most characterized [1]. Mitochondria, which carry their own circular genomic DNA (mtDNA), transcribe several RNAs that are required for mitochondrial functions [2]. Three distinct promoters, LSP, HSP1 and HSP2 are located in the mtDNA to express several protein encoding RNAs, tRNAs and rRNAs [3]. Thus, mitochondria have their specific transcription system that is separated from eukaryotic nuclear transcription and controlled by mitochondrial transcription factors, including mitochondrial RNA pol (POLRMT), TFB1M, TFB2M, and transcriptional activator TFAM [3]. Nevertheless, most of the mitochondrial or their function-associated proteins, which comprise 99% of mitochondrial proteins, have to be translated from mRNAs that are synthesized from nuclear genomic DNAs [4]. Mitochondrial proteome analysis showed that more than 1,000 different proteins are included in mitochondria [5]. After translation and modification, they are imported from the cytosol by two import pathways [6]. A variety of functions for these proteins, including energy metabolism (15%), protein synthesis (13%), genome maintenance and transcription (12%), and protein transport (7%), are suggested from the functional classification of the yeast mitochondrial proteome [6,7]. Those nuclear genes encoding mitochondrial function-associated proteins are likely to be regulated by a common system to form a mitochondrion as a huge complex composed of various proteins, nucleic acids, and other materials to supply energy required for cell survival.

From the analysis of promoter regions of cytochrome c (CYCS) and cytochrome oxidase subunit IV (COX9) showed that transcription factors NRF1 and NRF2 are involved in the regulation of expression of these genes [8,9]. Other transcription factors, such as estrogen-related receptor ERR, PGC-1, and PGC-1, are thought to be involved in the regulation of the mitochondrial gene expression system [2]. Previous studies of the promoter region of the human poly (ADP-ribose) glycohydrolase (PARG) gene, whose encoding protein PARG is shown to be localized in mitochondria [10], showed that it contains duplicated c-ETS binding sequences or GGAA-motifs and that it is linked with the most 5'region -upstream of the mitochondrial gene TIMM23B in a head-head configuration [11,12]. Moreover, we have a identified that a similar bidirectional promoter region between the IGHMBP2 and the MRPL21 genes has duplicated c-ETS binding- or GGAA-motifs [13]. The TIMM23B and the MRPL21 genes encode a protein component of the TIM23 complex and a large subunit of the mitochondrial ribosomal protein complex, respectively [6,14]. These findings suggest that some of the mitochondrial genes might be regulated by bidirectional promoters that contain duplicated GGAA-motifs.

In this review article, we will discuss the results of investigations to select promoter regions of the human genes encoding mitochondrial or mitochondrial function-associated proteins, and confirm that quite a few bidirectional gene pairs are included in the list. We also discuss the results of genome-informatic surveillance of the Ensembl database to find human genomic regions in which two genes are configured head-head, sharing the same 5'-flanking region, which regulates transcription in opposing directions for each gene.

2. Bidirectional promoters that are found in DNA-repair associated genes and 5'-upstream regions containing duplicated GGAA-motifs

Because we have identified the bidirectional partners for the *PARG* and the *IGHMBP2* genes, which are thought to be involved in the DNA repair synthesis system in eukaryotic cells, we searched 5'-upstream regions of several DNA-repair associated genes [15]. For example, *TP53/WRAP53* and *APEX1/OSGEP* gene pairs are linked by bidirectional promoter regions. Interestingly, it has been shown that p53, which is encoded by the *TP53* gene, and the APEX1 (APE1) proteins are reported to be involved in mitochondrial functions [16,17,18], and mitochondrial DNA repair [19,20], respectively. The palindromic c-Ets element, which is located near the transcription start site (TSS) of the *TP53* promoter [21], has an essential role in its regulation [22].

The essential role of the duplicated GGAA-motifs in the *PARG* promoter has been reported [12,23]. Additionally, the duplicated GGAA-motifs were identified in the 5'-flanking regions of the DNA-repair associated genes, including *XPB*, *Rb1*, and *ATR* [24]. A comparison of the duplicated GGAA-motifs of these gene regulatory regions tentatively determined a consensus 14-bp sequence as 5'-(A/G/C)N(A/G/C)(C/G)(C/G)**GGAA**(A/G)(C/T)(G/C/T)(A/G/C)(A/G/C)-3' or 5'-VNVSS**GGAA**RYBVV-3' in IUPAC code [24]. Thus we explored the duplicated 14-bp sequences within 2,000-bp upstream and 200-bp downstream of 47,553 genes extracted from the Ensembl data base [25]. The candidate sequences were found near the TSSs of 234 genes. Re-confirmation of the 5'-upstream regions of these 234 genes revealed that 21 pairs of protein encoding genes, including *MRPL32/PSMA2*, *NDUFB3/FAM126B*, *NDUFS3/KBTBD4*, *SDHAF2/CPSF7*, and *YRDC/C1orf122*, are linked to each other in a head-head configuration [25]. The *MRPL32* gene encodes a 39S subunit of the L32 mitochondrial ribosomal protein [14]. The *NDUFB3* and *NDUFS3* genes encode NADH dehydrogenase (ubiquinone) 1β subcomplex 3 and Fe-S protein 3, respectively [26]. These are the components of the NADH dehydrogenase or complex I, which consists of 45 subunits with a combined mass of 980-kDa [26]. SDHAF2 (succinate dehydrogenase complex assembly factor 2) is a protein that activates a succinate dehydrogenase complex subunit by flavination [27]. Activation of this enzyme complex is needed not only for the citrate cycle but also for the electron transport system. YRDC (yrdC domain containing) protein, which is suggested to have properties of a putative threonylcarbamoyl transferase [28], might also affect mitochondrial tRNA synthesis. These observations suggest the hypothesis that some of the mitochondrial function associated genes might be located near the 5'-upstream regions of other nuclear genes utilizing the same transcription regulatory elements to direct transcription in the opposite direction.

3. Survey of putative bidirectional promoter regions from Ensembl data base

As shown in Table 1, computer-based *in silico* analysis retrieved 240 bidirectional gene pairs from the Ensembl database covering approximately 50,000 human genes. The sequences between two genes were reconfirmed using the National Center for Biotechnology Information (NCBI) nucleotide database (http://www.ncbi.nlm.nhi.gov/nuccore/). Twelve gene pairs (Table 1, characters in green) were found to be separated over 500 nucleotides and were eliminated from the total number. As indicated in Table 1, at least 72 gene pairs contain mitochondria associated genes (31.6%), including 11 mitochondrial-mitochondrial gene pairs (4.82%). The remaining 156 gene pairs, which include twelve pairs of histone protein-encoding genes, are not directly associated with mitochondria but with DNA-repair, replication, transcription, translation, apoptosis, and other cellular functions.

Interestingly, surveillance of the genomic database revealed that *RPL9*, *RPS18*, and *RPS28* genes, which encode components of ribosomal protein subunits, have bidirectional partners. In addition, *EIF3I* and *EIF2B1* genes are also listed in Table 1, suggesting that some of the ribosomal protein-encoding genes are controlled by bidirectional promoters.

We have hypothesized that duplicated GGAA-motifs are frequently found in bidirectional promoters, so we therefore surveyed whether 630-bp regions covering both TSSs of gene pairs have duplicated GGAAs. It was confirmed that 54 out of 72 (75%) of mitochondrial, and 112 out of 156 (71%) of non-mitochondrial gene pairs possess duplicated GGAA-motifs. The observation suggests that the frequency of localization of GGAA duplication within 630-bp of the bidirectional promoter is not dependent on the biological functions or localizations of these gene-encoding products. Although the *IGHMBP2/MRPL21* gene pair is listed in Table 1, *TP53/WRAP53* and *PARG/TIMM23B* gene pairs were not retrieved from the Ensembl database, suggesting that other bidirectional partners of mitochondrial function related genes could be found from other nucleotide databases. Therefore, we decided to search bidirectional partners of mitochondria-associated genes from the NCBI database.

Two hundred and forty (240) bidirectional gene pairs that were extracted from the Ensembl database are shown. Genes indicated by red characters represent pseudo genes. Green colored characters indicate gene pairs that are separated over a distance of 500 nucleotides. Numbers of chromosomes are shown on the left of each gene pair. Mitochondrial or their function-associated genes are indicated by yellow background. Non-colored background represents genes whose encoding protein's functions or localizations are unknown at present (Jan.31.2013).

Table 1. Bidirectional promoter pair genes retrieved from Ensembl database

1	SSU72	AL645728.1	5	DMGDH	BHMT2	11	APIP	PDHX	16	CHTF8	CIRH1A
1	EMC1	MRTO4	5	DHFR	MSH3	11	KBTBD4	NDUFS3	16	KARS	TERF2IP
1	TMEM234	EIF3I	5	ATG12	AP3S1	11	MED19	TMX2	16	RNF166	CTU2
1	KIAA0319L	NCDN	5	HARS	HARS2	11	CPSF7	SDHAF2	16	GALNS	TRAPPC2L
1	YRDC	C1orf122	6	TDP2	ACOT13	11	C11orf48	C11orf83	16	CHMP1A	C16orf55
1	LEPRE1	C1orf50	6	HIST1H2AA	HIST1H2BA	11	C11orf68	DRAP1	17	MED31	C17orf100
1	EBNA1BP2	WDR65	6	HIST1H2BC	HIST1H2AC	11	CCDC87	CCS	17	LOC100506713	RNASEK
1	MED8	SZT2	6	HIST1H2AD	HIST1H2BF	11	MRPL21	IGHMBP2	17	ATPAF2	GID4
1	LRRC41	UQCRH	6	HIST1H2BG	HIST1H2AE	11	LOC100133315	RNF121	17	CCT6B	ZNF830
1	WDR78	MIER1	6	HIST1H2BJ	HIST1H2AG	11	C2CD3	PPME1	17	RAD51D	FNDC8
1	TMED5	CCDC18	6	HIST1H2BK	HIST1H2AH	11	TIMM8B	SDHD	17	FAM134C	TUBG1
1	SASS6	TRMT13	6	HIST1H2BL	HIST1H2AI	11	C11orf71	RBM7	17	RAMP2-AS1	RAMP2
1	WDR77	ATP5F1	6	HIST1H2AJ	HIST1H2BM	11	SRPR	FOXRED1	17	COA3	CNTD1
1	FAM212B	DDX20	6	HIST1H2AK	HIST1H2BN	12	LOH12CR2	LOH12CR1	17	PTGES3L	RUNDC1
1	HIST2H2BE	HIST2H2AC	6	HIST1H2AM	HIST1H2BO	12	YAF2	PPHLN1	17	MRPL10	LRRC46
1	LYSMD1	SCNM1	6	DOM3Z	STK19	12	CAPS2	GLIPR1L1	17	DDX5	CEP95
1	LOC100507670	ZNF687	6	TAP1	PSMB9	12	CCDC59	METTL25	17	GGA3	MRPS7
1	DCST2	DCST1	6	RXRB	SLC39A7	12	NT5DC3	GNN	17	JMJD6	METTL23
1	KRTCAP2	TRIM46	6	VPS52	RPS18	12	GPN3	FAM216A	17	TK1	AFMID
1	COPA	NCSTN	6	YIPF3	POLR1C	12	VPS29	RAD9B	17	ENTHD2	C17orf89
1	PFDN2	NIT1	6	RARS2	ORC3	12	EIF2B1	GTF2H3	17	OXLD1	CCDC137
1	INTS7	DTL	6	TSPYL4	DSE	12	DDX51	NOC4L	17	ARL16	HGS
1	NSL1	TATDN3	6	ADAT2	PEX3	12	POLE	PXMP2	17	RFNG	GPS1
1	GPATCH2	SPATA17	6	TCTE3	C6orf70	13	SKA3	MRP63	18	SPIRE1	PSMG2
1	HIST3H2A	HIST3H2BB	7	PMS2	AIMP2	13	NUFIP1	KIAA1704	18	TPGS2	KIAA1328
1	TTC13	ARV1	7	MPLKIP	C7orf10	13	MZT1	BORA	18	FLJ25715	CTDP1
1	C1orf131	GNPAT	7	PSMA2	MRPL32	14	SLC25A21	MIPOL1	19	LSM7	SPPL2B
2	CMPK2	RSAD2	7	DNAJC30	WBSCR22	14	AL139099.1	LRR1	19	SAFB2	SAFB
2	PTRHD1	CENPO	7	PEX1	RBM48	14	L3HYPDH	JKAMP	19	XAB2	PET100
2	CEBPZ	NDUFAF7	7	PTCD1	CPSF4	14	TRMT5	SLC38A6	19	NDUFA7	RPS28
2	ABCG5	ABCG8	7	MCM7	AP4M1	14	KIAA0317	FCF1	19	YIPF2	C19orf52
2	WDR92	PNO1	7	PLOD3	ZNHIT1	14	POMT2	GSTZ1	19	CCDC151	PRKCSH
2	MCEE	MPHOSPH10	7	ALKBH4	LRWD1	14	TMED8	SAMD15	19	ZNF490	ZNF791
2	CCDC142	TTC31	7	TMEM209	C7orf45	14	ALKBH1	SLIRP	19	MAN2B1	WDR83
2	AUP1	HTRA2	8	RP11-297N6.4	FDFT1	14	XRCC3	ZFYVE21	19	SLC1A6	CCDC105
2	C2orf68	USP39	8	SLC25A32	DCAF13	15	OIP5	NUSAP1	19	ANO8	GTPBP3
2	POLR1A	PTCD3	8	NUDCD1	ENY2	15	LRRC57	HAUS2	19	COPE	DDX49
2	CD8B	ANAPC1P1	8	MRPL13	MTBP	15	COPS2	GALK2	19	SUGP1	MAU2
2	MKI67IP	TSN	8	RP11-539E17.5	FAM83A	15	KIAA0101	TRIP4	19	NFKBID	HCST
2	SMPD4	MZT2B	8	TATDN1	NDUFB9	15	RASL12	KBTBD13	19	SIRT2	NFKBIB
2	CIR1	SCRN3	8	SLC45A4	RP-10J21.3	15	AAGAB	IQCH	19	PAF1	MED29
2	TYW5	C2orf47	8	BOP1	HSF1	15	MRPL46	MRPS11	19	GSK3A	LOC100132272
2	AAMP	PNKD	8	CYHR1	KIFC2	16	POLR3K	SNRNP25	19	PPP5D1	CALM3
2	THAP4	ATG4B	9	BAG1	CHMP5	16	RPUSD1	CHTF18	19	FAM71E1	EMC10
3	XPC	LSM3	9	ALG2	SEC61B	16	TSR3	GNPTG	20	NFS1	ROMO1
3	HACL1	BTD	9	NDUFA8	MORN5	16	MRPS34	EME2	20	NEURL2	CTSA
3	EPM2AIP1	MLH1	9	TRUB2	COQ4	16	HAGH	FAHD1	20	DPM1	MOCS3
3	GORASP1	TTC21A	9	DDX31	GTF3C4	16	ZNF598	NPW	21	URB1	C21orf119
3	KIAA1143	KIF15	9	SURF1	SURF2	16	NTHL1	TSC2	21	C21orf67	FAM207A
3	MBD4	IFT122	9	SDCCAG3	PMPCA	16	HCFC1R1	THOC6	22	UFD1L	CDC45
3	MRPL47	NDUFB5	9	ANAPC2	SSNA1	16	TMEM186	PMM2	22	PHF5A	ACO2
3	RTP2	PRPF31	9	TMEM203	NDOR1	16	PALB2	DCTN5	22	NHP2L1	C22orf46
4	RPL9	LIAS	10	KIN	ATP5C1	16	GTF3C1	KIAA0556	22	SMC1B	RIBC2
4	HELQ	MRPS18C	10	CDNF	HSPA14	16	BOLA2	SLX1B	X	GPR143	SHROOM2
4	UBE2D3	CISD2	10	AL355490.1	PGAM1	16	BOLA2B	SLX1A	X	AC115618.1	RBM3
5	BRD9	TRIP13	10	COX15	CUTC	16	ITFG1	PHKB	X	H2BFWT	H2BFM
5	FAM173B	CCT5	10	POLL	DPCD	16	NUDT21	OGFOD1	X	PSMD10	ATG4A
5	DHX29	SKIV2L2	10	PSD	FBXL15	16	LRRC29	TMEM208	X	RNF113A	NDUFA1
5	ERCC8	NDUFAF2	11	IGF2	IGF2-AS	16	KCTD19	LRRC36	X	IDH3G	SSR4
5	MRPS27	PTCD2	11	C11orf21	TSPAN32	16	SLC7A6OS	PRMT7	X	F8	FUNDC2

Mitochondria	metabolism/enzyme/signaling/hormone
Immune response	apoptosis/differentiation/autophagy
DNA repair/replication/recombination/mitosis/cell cycle	Golgi/peroxisome/endosome/lysosome/proteasome
transcription/RNA metabolism	membrane/ER/transport
translation/proteolysis/chaperone/proteasome	Histone proteins

4. Survey of 5'-flanking regions of human mitochondrial- or mitochondrial function associated genes

A list of human genes encoding mitochondria or their function associated proteins was obtained from the NCBI Entrez cross-database search (http://www.ncbi.nlm.nih.gov/sites/gquery).

This database holds 3,006 genes, including not only directly mitochondrial function-associated protein encoding genes, but also indirectly associated or function-unknown protein-encoding genes. First, we omitted non-coding RNA sequences and pseudo genes from the list and searched 5'-upstream regions of each gene to find 549 head-head linked gene pairs. Next, 233 gene pairs that are separated from each other but within 500-bp in distance were selected. Some of these pairs had overlapping 5'-untranslated regions. The duplicated GGAA motifs were then searched in the TSS containing 630-bp regions to select 140 gene pairs (Fig. 1, Group A). This group contained the same 32 gene pairs that were retrieved from the Ensembl data base (Fig. 1, Group B). Thus, we obtained a list of 151 gene pairs (Table 2) that are thought to be associated with mitochondria or their functions and that are head-head configured within 500-bp distances carrying at least one apparent duplicated GGAA-motif near putative TSSs. Most of these bidirectional promoters do not have apparent TATA boxes in common, suggesting that they might belong to TATA-less promoters [29].

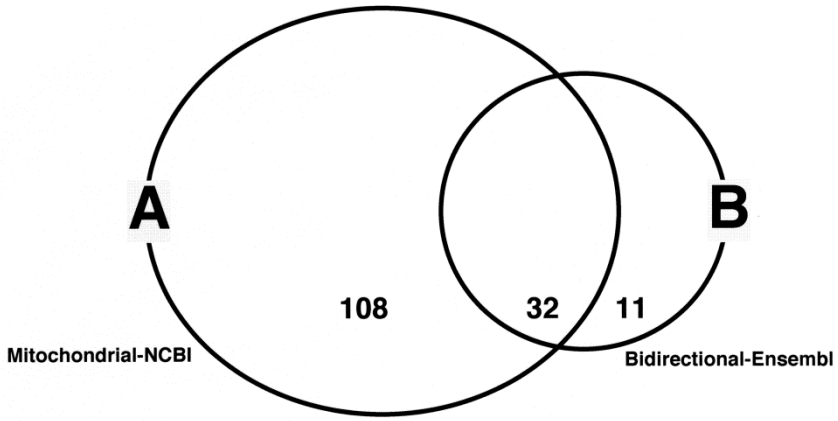

Figure 1. Mitochondrial and bidirectional genes surveyed from human DNA databases. Group A contains mitochondrial function-associated genes that were retrieved from analysis of the NCBI DNA database. The genes that belong to Group B were found from a bidirectional promoter search of the Ensembl database. The number of common genes identified from both human DNA databases was 32.

Table 2. GGAA motifs located in the 5'-upstream regions of head-head oriented pairs of human genes encoding mitochondrial proteins

Gene	Sequence
ACADVL-DLG4	TCTCT**TTCC**CTACTT**TTCC**CTTCT
ACO2-PHF5A	CCACC**GGAA**GCTTCG**GGAA**C**TTCC**GTCCG
ACOT13-TDP2	GGACT**TTCC**AGCTC**TTCC**GAAGT
ALDH6A1-LIN52	CGGCG**GGAA**AACGAGG**TTCC**ATAAG
ALKBH1-SLIRP	GCCCC**GGAA**AAAAT**TTCC**GGATCC**GGAA**CACGA, CTTTC**GGAA**ACTT**TTCC**GCTTC
ALKBH4-LRWD1	CGACC**GGAAGGAA**GC**GGAA**CCCAG
APEX1-OSGEP	CACTG**GGAA**AGACACCGC**GGAA**CTCCC, CCGTT**TTCC**TATCTCT**TTCC**CGTGG
APOPT1-BAG5	GACGC**TTCC**ACGAC**TTCC**GCAGC, TGACC**GGAA**G**GGAA**GACAA
ATG12-AP3S1	ACTAA**GGAA**AGC**GGAA**ACATT, CCGTC**TTCC**GCTGCAGT**TTCC**CCG**GGAA**CAGA**GGAA**CCTGC

ATM-NPAT	CAGCA**GGAA**CCACAATAA**GGAA**CAAGA, CCTTC**GGAA**CTGTCGTCAC**TTCC**GTCCT
ATP5A1-HAUS1	GGGCA**GGAA**G**GGAA**AGGCC, GCCAC**TTCC**CAGCTC**TTCC**CGCC**TTCC**GCGGT
ATP5C1-KIN	CAGCC**GGAA**AC**GGAA**CCGGG
ATP5F1-WDR77	TTGGG**GGAA**GA**TTCC**ACTCC, GGGGT**TTCCTTCC**GCATC, CCCAA**GGAA**AGTTGAA**GGAA**GAGTA
ATP5H-KCTD2	CCCCG**GGAA**GATAC**TTCCGGAA**CCAGC, GCTGA**GGAA**AGATC**TTCC**CGTGACCCAC**TTCC**GTTAC
ATP5J-GABPA	GCCGA**TTCC**GCG**GGAA**GGGCC, CCTCG**TTCC**GGGGCCTT**TTCC**CCCAC
AUP1-HTRA2	CAGTA**GGAA**GCAGTCACCC**GGAA**GCCTG
BAG1-CHMP5	GGCGT**TTCC**CGATTCTTT**TTCC**GGATT
BCS1L-ZNF142	CCTCC**TTCC**GAGAG**TTCC**CAGCG, CCGCA**TTCC**C**TTCC**CAG**TTCC**GCCCC
BOK-BOKAS1	AGCGG**GGAA**GCTC**GGAA**AGCGT
BOLA3-BOLA3AS1	GAGTG**GGAA**T**GGAA**AAGTA
BRCA1-NBR2	ATGCT**GGAA**ATAATTAT**TTCC**CTCCA, AATTC**TTCC**TC**TTCC**GTCTCT**TTCC**TTTTA, TTGGT**TTCC**GTGGCAAC**GGAA**AAGCGCG**GGAA**TTACA
BRE-RBKS	TCTTC**TTCC**T**GGAA**TAGTC, GCTGA**GGAAGGAA**CTGTC
BTG2-LOC730227	CCACG**GGAA**G**GGAA**CCGAC
C2orf47-TYW5	CCCTC**TTCC**AGGTCTTC**GGAA**CTTCG, GCGGT**TTCC**CACCGAC**TTCC**T**TTCC**ATACA
C7orf10-MPLKIP	CGGGG**TTCC**CC**GGAA**GCTGC**TTCC**GCTAC
C10orf2-MRPL43	GAGGC**TTCC**GG**TTCC**GGGAC, TGAGG**GGAA**GGAGAAGC**GGAA**GAGGG
CDK2-PMEL	CGAGA**TTCC**CGGC**TTCC**TGGT**TTCC**AAAGG, GCCAG**GGAA**ACGCG**GGAA**GCAGG
CDKN2A-CDKN2AAS1	AGCCA**GGAA**TAAAATAAGG**GGAA**TAGGG
CHCHD4-TMEM43	TGCTG**GGAA**ATGTAGT**TTCC**GGCTG
CHPF-TMEM198	CCTGA**GGAA**GG**GGAA**GGCGC, GGCAC**TTCC**GGGGGTCC**TTCC**CCTTT
CLRN1-CLRN1AS1	GCTGA**GGAA**G**GGAA**ACATT
CMC2-CENPN	GGCCG**TTCC**GAACGCG**TTCC**GTTG**TTCC**TCCTC
COA3-CNTD1	CTGGA**TTCC**TCGTCCCC**TTCC**AATGA
COA5-UNC50	CGGGC**TTCC**C**TTCC**CTCAA, TCACG**GGAA**CCGAC**TTCC**GCCGC
COQ4-TRUB2	TCAGT**TTCC**CCCTC**GGAA**AACAGA**GGAA**AGTGA, CCATG**TTCC**ACAGCC**GGAA**GAGGT
COQ9-CIAPIN1	CTGCG**TTCC**CATCGA**GGAA**CGGGGTG**GGAA**GAGAA
COX10-COX10AS1	TGCGG**TTCC**C**GGAA**GTCCT, CCGCC**GGAA**GTGGCGGCCC**GGAA**CTACT, GGCGG**GGAAGGAA**GATGG
COX15-CUTC	TTTTA**GGAA**GG**TTCC**TCTTCACGGA**GGAA**GAGGG
DAP3-YY1AP1	GACTG**TTCC**A**TTCC**TGGCG
DARS2-CENPL	TTGAA**TTCCTTCC**CGGTA
DBI-C2orf76	GCTCT**TTCCTTCC**GTGCC
DDX11-DDX11AS1	GAGCG**GGAA**AACA**TTCCGGAA**GTGGA
DDX19B-AARS	GTGGA**TTCC**TG**GGAA**GGCGG
DDX20-FAM212B	CCAGC**GGAA**C**GGAA**AAC**GGAA**GCACG, CTTTC**TTCC**AC**TTCC**AGGCC
DDX24-IFI27L1	GGCCA**GGAA**CCCGGG**TTCC**TATCG
DFFB-CEP104	GCCGC**TTCC**TCAGAC**GGAA**CTCGG
DHFR-MSH3	GGCTC**TTCC**CACC**TTCC**CCTTC
DHX29-SKIV2L2	TTTTT**TTCC**TGCTTTTCT**TTCC**T**TTCC**T**TTCC**GACCT
DHX38-TXNL4B	TAGCG**GGAAGGAA**ACCGA, CCTTT**TTCC**CCTCCTT**TTCC**TGCCC, ATCCA**GGAA**TCGGGCGTG**TTCC**AGGCT, CCTTC**TTCCTTCC**TGGCC
EARS2-UBFD1	TGGCT**GGAA**GCAGTCCCC**GGAA**GTGAC

ECI2-LOC100507506	CCGTT**GGAA**GACCCTCC**TTCC**CTATT
EIF2A-SERP1	AAAGA**GGAAGGAA**ACGCA
GATC-TRIAP1	GCCAA**GGAAGGAA**GAAAT
GATM-LOC145663	GTACT**GGAAGGAA**AGCAC, GGCGC**TTCC**CGACAG**TTCC**TAATT, GGCCA**GGAA**CA**TTCC**GCGCG
GLRX5-SNHG10	ACACC**GGAA**CC**GGAA**ACTTC, ACCCC**TTCC**CGGCG**TTCC**GC**GGAA**CGGGC, CAGGA**GGAA**AGTCGTC**TTCC**CTCTT
GTPBP3-ANO8	GTTGG**GGAA**GA**TTCC**TGGTG
HADHA-HADHB	TGCGG**GGAAGGAA**GT**GGAA**TCTCG
HARS-HARS2	CGGCT**TTCC**GGGACA**GGAA**CAAAA, GGCAC**TTCC**GGGAGGAGCC**GGAA**ATAAT
HSPA1A-HSPA1L	ACCCT**GGAA**TA**TTCC**CGACC
HSPD1-HSPE1	TTTCT**GGAA**AGTTCT**GGAA**CCGAG
IBA57-C1orf148	CCCGC**TTCC**TTGGGCCC**TTCC**CGCTG
IDH1-IDH1AS1	AGCCG**GGAA**GA**GGAA**AAGCT, TCTAA**TTCC**GCAGAAGGCA**GGAA**TGGGGTAAA**GGAA**AAAAG
IMMP1L-ELP4	CAATA**GGAA**CTCTG**GGAA**CGCAA
ISCA2-NPC2	CGAAG**TTCC**AAGCTCG**GGAA**AGAAG, CCCCG**TTCCTTCC**CTTTA
KIAA0391-PPP2R3C	ATTAG**TTCC**GGCTTGAG**GGAA**GCGCC, TATCT**GGAA**GCC**TTCC**AGGTC
LIPT1-TSGA10	TGCCT**GGAA**CCTGG**TTCC**CGCCC
LMNA-MEX3A	TTTCT**TTCC**ATTA**TTCC**AGATA, GTGGT**GGAA**G**GGAA**AAGAG
LRTOMT-NUMA1	ACCCG**GGAA**AA**GGAA**GTTTG
MARCH5-CPEB3	GCTGC**TTCC**T**GGAA**AGCGG, ATTTT**TTCC**CCCTGGA**GGAAGGAA**ACGGG
MLLT11-CDC42SE1	CCCCC**GGAA**TCTCACG**TTCC**CTTTA
MRP63-SKA3	AAGGG**TTCC**TA**GGAA**TAAAC
MRPL10-LRRC46	AGCAG**TTCC**TA**GGAA**GCCGG, TCGGC**TTCC**GTCCATTC**TTCC**GGTGG
MRPL13-MTBP	CGACG**GGAA**TT**TTCCTTCC**CCCCA, CGCAT**TTCC**GG**TTCC**CTTCG, CGGTT**TTCC**GCAGTTT**TTCC**ACCAA, CTGGA**GGAA**ACTCGA**GGAA**GAGGG
MRPL21-IGHMBP2	GTCGT**TTCC**GT**TTCC**GGCCG
MRPL30-MITD1	CAGCA**GGAA**CCAGCTCCTT**TTCC**TCAGG, TGCGC**TTCC**G**GGAA**GTGGT, AGTTC**TTCC**TCTGCTCTGC**TTCC**CTTCGGA**GGAA**AATTT
MRPL32-PSMA2	CCGAT**TTCC**TTTCAT**TTCC**CCGCC, CGGTC**TTCC**AGCAG**GGAA**AATGG, GCTAC**GGAA**GC**TTCC**GCAGA
MRPL37-CYB5RL	GAAAG**GGAA**GTGCC**TTCC**CAAGC, GGATT**TTCC**AGG**TTCC**TCCCA
MRPS7-GGA3	AGGCT**TTCC**CC**TTCC**GCCTC
MRPS12-SARS2	CCAGT**TTCC**CAGTTC**TTCC**TGCGT
MRPS18B-PPP1R10	CTCTC**TTCC**GCCTCC**TTCC**TGCCT, TCCCT**TTCC**TCTGC**TTCC**GCCAC**TTCC**GCCCT, ATGCC**TTCC**CTTTCACGC**TTCC**GTCCT, TGCCT**TTCC**GTCAA**TTCC**TGTCC
MRPS18C-HELQ	CATGG**TTCC**GCGTTTC**TTCC**AC**TTCC**TTTCG**TTCC**AAATCG**TTCC**GAAAGGCCCC**TTCC**GCTGCTC**TTCC**CCTGT
MRPS27-PTCD2	GGAGA**GGAA**ACGTTTCT**GGAA**TCTGA, TCCAT**TTCCTTCC**CTAAA
MRPS30-XR108577.1	AGTTC**TTCCTTCC**ATCTA, TCAGA**TTCC**GCT**TTCC**GATTG
MRPS34-EME2	ACCTC**TTCC**TCGC**TTCC**GGCCG, GCCTC**TTCC**GGTGAC**TTCC**GGCCG
MRRF-RBM18	TGCCT**GGAA**CCCTGGC**TTCC**CGATT
MTRR-FASTKD3	GCGTT**TTCC**TA**GGAA**TGAAA, TACGG**TTCC**CGGA**TTCC**GGCCG
MUT-CENPQ	GACCC**GGAA**GTGGGTG**GGAA**GAAAGC**GGAA**ACGGG
MYO19-PIGW	ACGCG**GGAA**CCAGCCGC**TTCC**GCCTC
NAGS-PYY	ACTCT**TTCC**AGCGCCCTCC**TTCC**AGCCC

NDE1-KIAA0430	TTCAC**TTCC**GC**TTCC**GCACC
NDUFA1-RNF113A	GCAGA**TTCC**GTCGCTTC**TTCC**GGAGC
NDUFA2-IK	AGCCT**TTCC**GC**TTCC**TGTTT**TTCC**CTCCG
NDUFB3-FAM126B	TACTG**GGAA**AATAATCGAC**TTCC**AGCGT
NDUFB9-TATDN1	CCAGC**GGAA**GC**GGAA**GTGGC
NDUFC1-NAA15	TTTCA**TTCC**TT**GGAA**AGAGT
NDUFS1-EEF1B2	AGACC**GGAA**AA**TTCC**TTATA, GCCAC**TTCC**GGC**GGAA**CTGCG
NDUFS3-KBTBD4	AGCCC**GGAA**CCTCCGC**TTCC**GGCTC, CACAC**TTCC**GT**TTCC**GGTCC
NFS1-ROMO1	TCAGG**GGAA**GTAAG**GGAAGGAA**AATCA, GAATA**TTCC**GGAGCC**TTCC**TGTCC
NIT1-PFDN2	AGCTT**TTCC**GGGGACCC**TTCC**CTCTC
NR2F2-NR2F2AS1	AGTTA**TTCC**AGTTTAGGA**GGAA**GATGC, CTCAT**TTCCTTCC**ACAGA
NUDT1-FTSJ2	CCCGG**GGAA**CTGCGACCC**GGAA**TCCTG
OXLD1-CCDC137	GCCAC**TTCC**GCC**TTCC**TGCATGG**TTCC**GCCCC, GGCAC**TTCC**GC**TTCC**TGCGC
PARG-TIMM23B	GCCGC**TTCC**CCCGCCTCC**TTCC**ATGGT, TGACC**TTCC**GGGCGCCGG**TTCC**CGTTA, GCCCC**GGAA**GCT**GGAA**GCGCC, CAGCT**TTCC**GGTGGTG**GGAA**AGTGA
PDHX-APIP	CTTAA**GGAA**GAATCG**TTCC**CATGA
PHB2-EMG1	CAAAT**TTCCTTCC**GGCTG, GGGAC**TTCC**GTATGCGCGA**TTCC**TGTGC
PMPCA-SDCCAG3	GGAGG**GGAA**GCCGTGGGC**GGAA**GC**GGAA**GTGAC
PRDX5-TRMT112	AGGCC**GGAA**CC**GGAA**AAAGG
PRKCQ-PRKCQAS1	GAGTA**GGAA**AT**GGAA**CCAAG, CACCG**GGAA**GAAT**TTCC**CCGCT
PRPF31-TFPT	GTAGT**TTCC**TGT**TTCC**GGCTT
PSMD10-ATG4A	CGACG**GGAA**AAGAAAAG**GGAA**CGA**GGAA**GGCCG, CAGCG**GGAA**GCT**GGAA**GAGTT
PTCD1-CPSF4	AGAGG**GGAAGGAA**GTGCC
PTCD3-POLR1A	CGCGC**GGAA**GCGGTCGCA**GGAA**CGACA, ATTTA**GGAA**AA**TTCC**TCCGA
PTK2B-TRIM35	GTCCC**TTCC**CCT**GGAA**CGCTG, CCCAC**TTCC**GGTGTGCGCG**GGAA**ATCTT, AGGGC**TTCC**GTGTTACT**GGAA**ACCTAC**TTCC**GGCTG, CCAAC**TTCC**TGC**TTCC**GAAGT
PTRH2-VMP1	ACCCA**GGAA**CCCC**GGAA**GAGGT, TTGGA**GGAA**A**GGAA**CAGGC
RMRP-CCDC107	CTCTG**TTCC**TCCCCT**TTCC**GCCTAGG**GGAA**AGTCC, AGACA**TTCC**CCGC**TTCC**CACTC
RNF185-MIR3928	CAAGG**TTCC**TCCGC**TTCC**TGCCC, GCCAT**GGAA**ATTAACCTC**TTCC**GGTTGGGGCC**GGAA**GTCCC
ROM1-EML3	AGAGG**GGAA**GG**GGAA**GCACC, CCGGA**TTCC**CAGGGGACGG**GGAA**GGGAG
RPS6KB1-TUBD1	CGGAC**TTCC**GAGACAG**GGAA**GCTGA
RTN4IP1-QRSL1	AGAGC**GGAA**TAACAG**TTCC**GTATT
SCO1-ADPRM	ACTCC**TTCC**GAC**TTCC**GGA**GGAA**GC**GGAA**CGCTACC**GGAA**ATCGC
SDHAF2-CPSF7	AGGAG**TTCC**C**GGAA**GTGCC
SERAC1-GTF2H5	CAGTG**TTCC**CACACCCCACT**TTCC**CCAGC
SHC1-CKS1B	ACTCG**GGAA**AGTG**GGAA**GCGTG
SIRT3-PSMD13	ACTAG**GGAA**C**TTCC**TCTAC
SKIV2L-RDBP	CGCAC**TTCC**GCCCGGCC**TTCC**ACCGG, TCTAC**TTCC**GCCCG**TTCC**GGGGC
SLC25A11-RNF167	CCGTG**TTCC**CAGCCTCT**GGAA**AAGGGC**TTCC**GGTAG, CGCCT**TTCC**TCTGGT**TTCC**AAATC
SLC25A27-CYP39A1	GTTGG**GGAA**ATTAG**TTCC**ATGTT, GTAGC**TTCCTTCC**TCTGT
SLC25A32-DCAF13	GCGAC**TTCC**GCTT**TTCC**CAGACTAC**TTCC**AGTCA
SQSTM1-MGAT4B	CTCAG**GGAA**GA**GGAA**CAGGC

SSBP1-FLJ40852	TCTGT**TTCC**TTT**TTCC**TCTGG, TTGCG**TTCC**CTGTGCGCC**GGAA**GTGAT, GTGAT**TTCCTTCC**AGTTC
STARD7-LOC285033	ACTGG**TTCC**TTGGGCCCC**GGAA**GCTCG
TAP1-PSMB9	GCCTG**TTCC**TGGGACT**TTCC**GAGAG
TFAP2A-LOC100130275	CTCGC**TTCC**TCTCCCC**TTCC**CCCTC, CCTCT**TTCC**CTCCC**TTCC**TCCTC, TCGAT**TTCC**AGGCATTC**TTCC**CTTAT
TFB2M-CNST	GAGGC**GGAA**GC**GGAA**GTGAG
TIMM8B-SDHD	CTCAC**TTCC**ATCCCC**TTCC**CTGGC, AACAG**GGAA**GA**GGAA**ATGCT, TGAAG**GGAA**A**GGAA**GTTTCACC**TTCC**TTGGT
TMEM186-PMM2	CACGA**GGAA**CTCGGCCC**GGAA**C**TTCC**GGGTT
TOMM70A-LNP1	ACGCT**TTCC**CTTAACCC**GGAA**GTGAT**TTCC**GCCCC
TP53-WRAP53	TCCAT**TTCC**TTTGC**TTCC**TCCGG
TRAK2-STRADB	ACACC**TTCC**C**TTCC**CGGGC, GGGCA**GGAA**ACTACAA**TTCC**CAGCA, GGGGC**TTCC**TGA**GGAA**GCGCG
TRMT5-SLC38A6	GGCCC**TTCC**GGCA**TTCC**GTACT, CGTGG**GGAA**TT**GGAA**TGGTG
TRPV4-MIR4497	CCATT**TTCC**AGGCGA**GGAA**ACTGA, GAGAC**GGAA**GGCACCAGG**TTCC**GCAGG
TUBA1C-LOC100293962	AGCCC**TTCC**TGCCC**TTCC**TCCCC**TTCC**TCCCC**TTCC**TCCCC**TTCC**TCCCC**TTCC**TCCCC**TTCC**TCCCC**TTCC**TCCCC**TTC** **C**TCCTC**TTCC**TGCTC
TXNRD2-COMT	CCGAA**TTCC**CCAG**TTCC**CATCCAGA**TTCC**CCACC
UQCR10-ZMAT5	TTTTT**TTCC**TCA**TTCC**CAGGATA**GGAA**TACCT
USMG5-PDCD11	AATTC**TTCCGGAA**GTGTG, AGACA**GGAA**TC**GGAA**GGGCG, GGTGG**GGAA**TG**GGAA**AGGAG
WARS-WDR25	CTGAA**GGAA**CA**GGAA**GTGAAA**GGAA**GGCGA, GGGCC**GGAA**GTTG**TTCC**GTCCG
WASF1-CDC40	TCATC**TTCC**CTCA**TTCC**CTAGC, CCTCA**GGAA**AGGGG**GGAA**GAGCG
WDR92-PNO1	TGCGT**TTCC**ATTTCGGA**TTCC**ATCCCC**GGAA**ATCTT
WDR93-PEX11A	CGTCC**TTCC**TGC**TTCC**ATCTA, CCTTT**TTCC**CGGGTGCAGG**GGAA**TAGGC, CGGAC**TTCC**GGTTCAAGCC**GGAA**GTTGT
YME1L1-MASTL	ACTAC**TTCC**GCG**GGAA**AGAAC
YRDC-C1orf122	GTCAC**TTCC**TCCC**GGAA**GCGGG

Duplicated GGAA (TTCC) motifs that are located within 630-bp containing both head-head configured transcription start sites (TSSs) of mitochondrial or their function-associated protein encoding genes (shaded) are shown.

Functions of the proteins encoded by some of the genes in Table 2 could be classified into several categories as follows:

1. Subunits of the mitochondrial ATP synthase; *ATP5A1*, *ATP5C1*, *ATP5F1*, *ATP5H*, and *ATP5J*.
2. Components of the 39S subunit of mitochondrial ribosome (mitoribosome); *MRPL10*, *MRPL13*, *MRPL21*, *MRPL30*, *MRPL31*, *MRPL32*, *MRPL37*, and *MRPL43*.
3. Components of the 28S subunit of mitochondrial ribosome (mitoribosome); *DAP3* (*MRPS29*), *MRPS7*, *MRPS12*, *MRPS18B*, *MRPS18C*, *MRPS27*, *MRPS30*, and *MRPS34*.
4. Components of the NADH dehydrogenase (ubiquinone); *NDUFA1*, *NDUFA2*, *NDUFB3*, *NDUFB9*, *NDUFC1*, and *NDUFS3*.
5. Mitochondrial aminoacyl- tRNA synthetases; *EARS2*, *DARS2*, *HARS2*, and *SARS2*.
6. Mitchondrial oxoadipate carriers; *SLC25A11*, *SLC25A27*, and *SLC25A32*.
7. Enzymes or regulators of citrate cycle and respiratory chain; *BCS1L*, *BOLA3*, *CMC2*, *COA5*, *COX10*, *COX15*, *IDH1*, *SCO1*, *SDHD*, and *SDHAF2*.
8. Mitochondrial import proteins; *CHCHD4 (TIMM40)*, *TIMM8B*, *TIMM23B*, and *TOMM70A*.

The number of gene pairs within 500 nucleotide distances that have no apparent duplicated GGAA motifs in the TSS-containing 630-bp regions totalled 93. Although these have not been

examined by further sequence analyses, most of the gene pairs have multiple GGAA-motifs near their TSSs. For example, the GGAA-motifs are located in the bidirectional promoter regions of *COX11-STXBP4*, *COX4I1-EMC8*, *FARS2-LYRM4*, *FASTKD2-MDH1B*, *FOXRED1-SRPR*, *GLUD1-FAM35A*, *IDH3G-SSR4*, *MRPL2-KLC4*, *MRPL14-TMEM63B*, *MRPL27-EME1*, *MRPL46-MRPS11*, *MRPL47-NDUFB5*, *MRPL49-FAU*, *MRPL50-ZNF189*, *NDUFA6-LOC10032273*, *NDUFA7-RPS28*, *NDUFA8-MORN5*, *NDUFAF2-ERCC8*, *NDUFAF3-DALRD3*, *NDUFAF5-ESF1*, *NDUFAF7-CEBPZ*, *NDUFB1-CPSF2*, *NDUFB11-RBM10*, *NDUFS2-ADAMTS4*, *NNT-LOC100652772*, *RARS2-ORC3*, *SDHA-CCDC127*, *SURF1-SURF2*, *TIMM10B-ARFIP2*, and *TIMM17B-PQBP1* gene pairs, but the intervals between two motifs are more than 10 nucleotides. It should be noted that bidirectional promoter regions of HIST-HIST gene pairs do not always carry duplicated GGAA-motifs. These observations imply that bidirectional promoters are not only regulated by the interactions between GGAA motif binding proteins but also by other transcription factors that bind to sequences near the TSSs on these bidirectional promoters.

5. Transcription factors that recognize and bind to bidirectional promoters of the mitochondrial function-associated genes

It has been shown that TATA-box is less frequently identified in bidirectional promoters than unidirectional promoters [30,31]. However, CpG islands and CCAAT box are found two-fold in bidirectional promoters compared with unidirectional promoters [30,31]. In addition, specific binding sites for GABPA, MYC, E2F1, E2F4, NRF1, YY1, NF-Y, and SP1 have been identified within bidirectional promoter regions [32,33]. Notably, higher occupation (approximately 8-fold) of the GABPA to bidirectional promoter regions compared to unidirectional promoters has been shown [32]. Moreover, it has been suggested that hStaf/Znf143 is involved in the regulation of bidirectional promoter activity [34,35]. These observations imply that bidirectional promoters are regulated by transcription factors (TFs) that recognize and bind to various *cis*-elements that play important roles in the regulation of TATA-less promoters. Moreover, analysis of gene ontology indicated that many of the genes driven by bidirectional promoters share similar functions, including RNA-processing, DNA-repair, regulation of cell cycle and metabolism [36,37,38]. Thus, our observation that duplicated GGAA sequences are frequently found in the bidirectional promoters that control mitochondrial function associated gene expression is consistent with the previous analysis of the human genome data bases.

Table 3. The ten adjacent nucleotides around each of the duplicated GGAA motifs located in the bidirectional promoter regions of the human mitochondrial function-associated genes

	-5	-4	-3	-2	-1	GGAA	1	2	3	4	5
A (%)	22	28	27	23	29		26	22	24	23	27
G (%)	37	31	31	29	27		47	30	37	40	33
C (%)	22	27	26	35	31		17	26	20	21	25
T (%)	18	14	14	12	12		10	22	18	15	15
Consensus	V	V	V	V	V	GGAA	R	N	R	R	V

Sequences of fourteen nucleotides containing a GGAA (TTCC) were extracted from Table.1.

The directions of each sequence were aligned as 5'-NNNNNGGAANNNNN-3'. V, R, and N represent (A/G/C), (A/G), and (A/G/C/T), respectively.

The GGAA motif- containing sequences (Table 2) were aligned with the GGAA core placed on the center of the 14 nucleotide sequences, and then the 5'-and 3'- adjacent sequences were examined (Table 3). The consensus sequence was determined as 5'-(A/G/C)(A/G/C)(A/G/C)(A/G/C)(A/G/C)**GGAA**(A/G)N(A/G)(A/G)(A/G/C)-3' or 5'-VVVVV**GGAA**RNRRV-3' in the IUPAC format. Unlike the tentative consensus 14-bp sequence containing GGAA in the human ISG promoters [39], the nucleotide on the 5'-side of GGAA is A, G, or C, suggesting that all ETS family proteins could have access to the 14 nucleotide elements of the mitochondrial function-associated bidirectional promoter region. Table 2 includes 543 sequences carrying GGAA as a core motif. These were analyzed by the JASPAR online free software (http://jasper.genereg.net/cgi-bin/jaspar_db.pl) to search putative binding elements recognized by human TFs within a 90% threshold. The most abundant TF binding site in the 543 sequences is ETS1 (99.8%). The next is SPIB (27.1%), followed by SPI1 (18.4%), ELK1 (14.9%), MZF1-1.4 (11.9%), FEV (8.7%), NFATC (7.0%), and ELK4 (6.1%). The other TF binding sites found in these sequences are GATA2/3, NFIC, FOXA1/C1, YY1, and ZNF354C. Because the binding motif of ETS1 shown by the JASPAR software is thought to represent the binding motif for ETS family proteins, all of them are able to access bidirectional promoters listed in Table 2.In order to examine if the duplicated GGAA motifs in the 151 bidirectional promoters (Table 2) are conserved through evolution, BLAST search analysis of the mouse genomic sequences was executed. As shown in Table 4, conserved GGAA (TTCC) duplications are present close to the TSSs of the 80 mouse gene pairs. Although the duplicated GGAA (TTCC) motifs in the human genes are not conserved, another GGAA duplications are found at other sites near the TSSs of the 47 mouse gene pairs (indicated as "other site near TSS"). Yet, GGAA duplications are not identified near TSSs of remaining 24 mouse genes (indicated as "None"). Eighteen mouse genes (shaded in blue), *Bok*, *Bola3*, *Btg2*, *Cdkn2a*, *Dde11*, *Eci2*, *Gatm*, *Iba57*, *Mrps30*, *Nags*, *Nde1*, *Prkcq*, *Ssbp1*, *Tfap2a*, *Tomm70a*, *Trpv4*, and *Tuba1c* have no partner genes, implying that those in the mouse chromosomes might have been eliminated after evolution from the common ancestor of human and mouse. Alternatively, partner genes, including non-coding RNAs, might have been created or incorporated in the human chromosomes after evolution from the common ancestor. This explanation supports recent findings in genetic research suggesting that conversion of unidirectional into bidirectional promoters has generated novel transcripts with functional relevance [40]. This might partly explain our observation that human-specific non-coding RNAs are transcribed from bidirectional promoter regions.

Table 4. Conserved GGAA duplications in human and mouse bidirectional promoter regions of the mitochondrial function-associated genes

Human Gene Pairs	Mouse Gene Pairs	Human	Mouse	GGAA duplication
ACADVL-DLG4	*Acadvl-Dlg4*			None
ACO2-PHF5A	*Aco2-Phf5a*	CCACCGGAAGCTTCGGGAACTTCCGTCCG	CGGCCGGAAGCTTAGGGAACTTCCGCCCG	
ACOT13-TDP2	*Acot13-Tdp2*			None
ALDH6A1-LIN52	*Aldh6a1-Lin52*	CGGAGTTCCAGGTTAGGGGGGCGTCTTCCCCGTG	CGGAGTTCCCGGTTA-GGGGGCGTCTTCCCTGTG	
ALKBH1-SLIRP	*Alkbh1-Slirp*	CTTTCGGAAACTTTTCCGCTTC	TTTCCGGAAGCTTTTCCGCTTC	
ALKBH4-LRWD1	*Alkbh4-Lrwd1*	CGACCGGAAGGAAGCGGAACCCAG	CGACCGGAAGGAAGTAGAACGCCA	
APEX1-OSGEP	*Apex1-Osgep*			None
APOPT1-BAG5	*Apoptl-Bag5*			None
ATG12-AP3S1	*Atg12-Ap3s1*	CCCCGGGAACAGAGGAACCTGC	CGCCCGGAACGGAGGAACCTCC	
ATM-NPAT	*Atm-Npat*	CCTTCGGAACTGTCGTCACTTCCGTCCT	TCCTCGGAACCGTCGTCATTTCCGTCTG	
ATP5A1-HAUS1	*Atp5a1-Haus1*			None
ATP5C1-KIN	*Atp5c1-Kin*	CAGCCGGAAACGGAACCGGG	CGGTCGGAAGCGGAACTGGC	
ATP5F1-WDR77	*Atp5f1-Wdr77*	TTTCCTTCCGCATCTCCACGGTTCCAACTC	TGTCCTTCCGCATCTCCACGATTCCAAATC	
ATP5H-KCTD2	*Atp5h-Kctd2*	AGATCTTCCCGTGACCCACTTCCGTTAC	AGAACTTCCGGCAACTCTCTTCCGTTCC	
ATP5J-GABPA	*Atp5j-Gabpa*	GCCGATTCCGCGGGAAGGGCC	GCCGATTCCACGGGAAGGGCC	
AUP1-HTRA2	*Aup1-Htra2*			None
BAG1-CHMP5	*Bag1-Chmp5*	TCTTTTTCCGGATTTTCAGCCGGGTCTTCCGGGGA	CCCTTTTCCGTGTTTTCAGCTCCTCCTTCCGGAGA	
BCS1L-ZNF142	*Bcs1l-Zfp142*	CCGCATTCCCTTCCCAGTTCCGCCCC	CGCACTTCCCCTGGTTCCCGCCG	
BOK-BOKAS1	*Bok*			None
BOLA3-BOLA3AS1	*Bola3*			other site near TSS
BRCA1-NBR2	*Brca1-Nbr1*	TTTCATTCCGCAACGCATGCTGGAAATAAT	TTTCGTTCCGCAATGCATGCTGGAATTGGT	
BRE-RBKS	*Bre-Rbks*	AGGAAGGAACTGTCAGGCACCGCGGGGACGCCTTTCCGAGCGCCCGGCTC	TCCCGGGAACCTTCAGGCGCTACCCGGTCGCCTTCCTCAGCTTCCGGCTT	
BTG2-LOC730227	*Btg2*			other site near TSS
C2orf47-TYW5	*9430016H08-Tyw5*			other site near TSS
C7orf10-MPLKIP	*5033411D12-Mplkip*	TCCCCGGAAGCTGCTTCCGCTAC	ACCCCGGAAGTTATTTCCGCCAC	
C10orf2-MRPL43	*Peo1-Mrpl43*	GAGGCTTCCGGTTCCGGGAC	GTGACTTCCGGTTCCGGGGT	
CDK2-PMEL	*Cdk2-Pmel*	GCCAGGGAAACGCGGGAAGCAGG	GCCAGGGAAACGCGGGAAGGAGG	
CDKN2A-CDKN2AAS1	*Cdkn2a*			other site near TSS
CHCHD4-TMEM43	*Chchd4-Tmem43*			other site near TSS
CHPF-TMEM198	*Chpf-Tmem198*	GGCACTTCCGGGGGTCCTTCCCCTTT	GGGCACTTCCGGGGGTCCTTCCCCTTT	
CLRN1-CLRN1AS1	*Clrn1*			other site near TSS
CMC2-CENPN	*Cmc2-Cenpn*			None
COA3-CNTD1	*Coa3-Cntd1*	CTGGATTCCTCGTCCCCTTCCAATGA	TTGGATTCTTCCTTCCCTTTCTAT	
COA5-UNC50	*Coa5-Unc50*	CGGGCTTCCCTTCCCTCAA	GCCACTTCCCTTCCCGCTA	
COQ4-TRUB2	*Coq4-Trub2*	CCATGTTCCACAGCCGGAAGAGGT	CTTGTTTCCTTGGCAGGAAGAGGT	
COQ9-CIAPIN1	*Coq9-Ciapin1*			other site near TSS
COX10-COX10AS1	*Cox10-G20 Rik*	CCGCCGGAAGTGGCGGCCCGGAACTACT	GAGCCGGAAGTGACAAGAGGAAGTCCC	
COX15-CUTC	*Cox15-Cutc*			other site near TSS
DAP3-YY1AP1	*Dap3-Ash1*			None
DARS2-CENPL	*Dars2-Cenpl*	ATCGTTTCCCGACTCGCTCTGCGGCCCCACGCAGGAAGCCTC	ATCATTTCCCGACTTGCCCTGCGGCCG-ACTCGGAAAGCCTC	
DBI-C2orf76	*Dbi-Rik9E18*			other site near TSS
DDX11-DDX11AS1	*Ddx11*			None
DDX19B-AARS	*Ddx19B-Aars*			None
DDX20-FAM212B	*Ddx20-Fam212b*	CCGACTCCCGTCCCTTTCTTCCACTTCCAGGCC	CCGACTTCCGTTCTTCCCCTCCGCTTCCA----	
DDX24-IFI27L1	*Ddx24-Ifi27l1*			other site near TSS
DFFB-CEP104	*Dffb-Cep104*			other site near TSS
DHFR-MSH3	*Dhfr-Msh3*			other site near TSS
DHX29-SKIV2L2	*Dhx29-Skiv2l2*			None
DHX38-TXNL4B	*Dhx38-Txnl4b*			other site near TSS
EARS2-UBFD1	*Ears2-Ubfd1*	TGGCTGGAAGCA-----GTCCC----------CGGAAGTGAC	TGGTTGGAGTGATTTTTTCCCCCCGCCGAGCGGAAGTGAC	
ECI2-LOC100507506	*Eci2*			None
EIF2A-SERP1	*Eif2a-Serp1*			other site near TSS
GATC-TRIAP1	*Gatc-Triap1*			other site near TSS
GATM-LOC145663	*Gatm*			other site near TSS
GLRX5-SNHG10	*Glrx5-Snhg10*	ACACCGGAACCGGAAACTTC CAGGAGGAAAGTCGTCTTCCCTCTT	CAAAAGGAACCGGAAGCTGC CACGAGGAAAGCCGGCTTCCTCTAG	
GTPBP3-ANO8	*Gtpbp3-Ano8*			other site near TSS
HADHA-HADHB	*Hadha-Hadhb*			None
HARS-HARS2	*Hars-Hars2*	TTTCCGGGACAGGAACAAAAGGCCTGGGAAGGAGG	TTTCTGGAACTGAAGTCAAGGGACTGGGAACGAGG	
HSPA1A-HSPA1L	*Hspa1a-Hspa1l*	ACCCTGGAATATTCCCGACC	CTGCTGGAAGATTCCTGGCC	
HSPD1-HSPE1	*Hspd1-Hspe1*	TTTCTGGAAAGTTCTGGAACCGAGCGAGGCCCGGGAACTAGA	CTTCCGGAAGGTTCTAGAACGGACCGTGGCCCAGGAACCAGC	
IBA57-C1orf148	*Iba57*			other site near TSS
IDH1-IDH1AS1	*Idh1-Pikfyve*			None
IMMP1L-ELP4	*Immp1l-Elp4*	CAATAGGAACTCTGGGAACGCAA	CAATCGGAACTCTGGGAACTGGA	
ISCA2-NPC2	*Isca2-Npc2*			other site near TSS
KIAA0391-PPP2R3C	*1110008L16Rik-Ppp2r3c*			other site near TSS
LIPT1-TSGA10	*Lipt1-Tsga10*			other site near TSS
LMNA-MEX3A	*Lmna-Mex3a*	TTTCTTTCCATTATTCCAGATA	TTTCTTTCCATTATTCCAGATA	
LRTOMT-NUMA1	*Lrtomt-Numa1*	ACCCGGGAAAAGGAAGTTTG	ACCCGGGAAAAGGAAGTACA	
MARCH5-CPEB3	*March5-Cpeb3*	ATTTTTTCCCCCTGGAGGAAGGAAACGGG	TTTTTTTCCCCCTGGAGGAAGGAAACGGG	
MLLT11-CDC42SE1	*Mllt11-Cdc42se1*			None
MRP63-SKA3	*Mrp63-Ska3*	GGTGCTTCCCGATCCACTGACGGCCGGGAATGCGG AGTCAGGAACATGGTGCCTGCGGAAGAGAG	GGTGCTTCCCGATCCACTGCCTGCCCGGAATGCGG GGTGAGGAACATGGTGCCTCCTGAGGAAATGAC	
MRPL10-LRRC46	*Mrpl10-Lrrc46*	TCGGCTTCCGTCCATTCTTCCGGTGG	GCAGCGCGCATGGGAACTTCCTGTAG	
MRPL13-MTBP	*Mrpl13-Mtbp*	CGACGGGAATTTTCCTTCCCCCA CGCATTTCCGGTTCCCTTCG	GGATCGAAACTTCCCTTCCCTCCA TGCACTTCCGGTTTCCTGCCG	
MRPL21-IGHMBP2	*Mrpl21-Ighmbp2*	GTCGTTTCCGTTTCCGGCCG	GCTGCTTCCGCTTCCGGTTG	
MRPL30-MITD1	*Mrpl30-Mitd1*	TGCGCTTCCGGGAAGTGGT	TGCGCTTCCAGGAAGCGCA	
MRPL32-PSMA2	*Mrpl32-Psma2*			other site near TSS
MRPL37-CYB5RL	*Mrpl37-Cyb5rl*	GGATTTTCCAGGTTCCTCCCA	GCACCTTCCAGGTTCCACCAA	
MRPS7-GGA3	*Mrps7-Gga3*			other site near TSS
MRPS12-SARS2	*Mrps12-Sars2*	CCAGTTTCCCAGTTCTTCCTGCGT	TCAGTTTCCC----CTTCCAGTAT	
MRPS18B-PPP1R10	*Mrps18b-Ppp1r10*	CCCCTTTCCCTGGCACTCTCTTCCGCCTCCTTCCTGCCT TCCCTTTCCTCTGCTTCCGCCACTTCCGCCCT	CCCCCTTCCCTCGGTTCCCCTCTTCCGTCTCCTTGCTATCA TCCCTTTCCTCTGCCTCCACCACTTCCGCTCC	

MRPS18C-HELQ	*Mrps18c-Helq*	GTTTCTTCCACTTCCTTTCGTTCCAAATC	CGCCCTTCCTCTTCCTTTCTATTTCCCTAAG	
MRPS27-PTCD2	*Mrps27-Ptcd2*			other site near TSS
MRPS30-XR108577.1	*Mrps30*			other site near TSS
MRPS34-EME2	*Mrps34-Eme2*	ACCTCTTCCTCGCTTCCGGCCG GCCTCTTCCGGTGACTTCCGGCCG	CTTCGTTCCTGAAGCTTCCGGACG GCCGCTTCCGGAACGGGAGAGCTCCTGAAG	
MRRF-RBM18	*Mrrf-Rbm18*			other site near TSS
MTRR-FASTKD3	*Mtrr-Fastkd3*			other site near TSS
MUT-CENPQ	*Mut-Cenpq*			other site near TSS
MYO19-PIGW	*Myo19-Pigw*	GCCGCTTCCGCCTCCCGCGCGGCACTTCCGGCCGACGCAGGCAGC TGGCACTTCCGGGCCGGCCGATTCCGCGCG	TCCGCTTCCGCCTCGCGAGCCGCGCTTCCGGCCGTCGCGGCTCAG CCTGTTCTTCCGGGACCGCAGATTCCGCGCG	
NAGS-PYY	*Nags*			None
NDE1-KIAA0430	*Nde1*			None
NDUFA1-RNF113A	*Ndufa1-Rnf113a*			None
NDUFA2-IK	*Ndufa2-Ik*			other site near TSS
NDUFB3-FAM126B	*Ndufb3-Fam126b*	TACTGGGAAAATAATCGACTTCCAGCGT GTATCGGAACGTTAAGCGGCTTCTCCGCTTTCCTGCCG	GTGTGGGAAATAATCTCTACTTCCAGGCT GTACCGGAACGTTTACCATCTCCCCCATTTTCCTGTCG	
NDUFB9-TATDN1	*Ndufb9-Tatdn1*	CCAGCGGAAGCGGAAGTGGC	CCCAGGGAAAAGCGGAAGTGGC	
NDUFC1-NAA15	*Ndufc1-Naa15*	CTCCTTTCCCCACCTCCTCTGGGTTTCGGAGCTTGCCGGGAAACC TG	CGCTCTTCCCCACCCTCGATGCACCGGAGCCTGCCGGGAACGCGG A	
NDUFS1-EEF1B2	*Ndufs1-Eef1b2*	GCCACTTCCGGCGGAACTGCG	GCTACTTCCGGCGGAACTGCA	
NDUFS3-KBTBD4	*Ndufs3-Kbtbd4*	AGCCCGGAACCTCCGCTTCCGGCTC CACACTTCCGTTTCCGGTCC	GATCCGGAACCTCTGCTTCCCGCCC CACATTTCCACTTCCGGTCC	
NFS1-ROMO1	*Nfs1-Romo1*	ATCGCTTCCGGAGCGCCGGGCAGCACTTCCGGGAG	ATCGCTTCCGGAGCACAGCGCAGTTACTTCCGGGAG	
NIT1-PFDN2	*Nit1-Pfdn2*			other site near TSS
NR2F2-NR2F2AS1	*Nr2f2-Gm7656*	AGTTATTCCAGTTTAGGAGGAAGATGC CTCATTTCCTTCCACAGA	AGTTATTCCAGTTTAGAAGGAAGATGC CTCATTTCCTTCCATAGA	
NUDT1-FTSJ2	*Nudt1-Ftsj2*	CCGCCGGAAGTGCCTGGCCTCACTTCCGGTCA	CCGCCGGAAGTCCCGCCCTCGCTTCCAACGC	
OXLD1-CCDC137	*Oxld1-Ccdc137*	GGCACTTCCGCTTCCTGCGC	ATCACTTCCGCTTCCTGCGC	
PARG-TIMM23B	*Parg-Timm23b*	GCCCCGGAAGCTGGAAGCGCCTGACGGCAGCTTTCCGGTGGTGGG AAAGTGA	GGCCCGGAAGTGGGAAGCGCGAGACGGCTGCTTTCCGGTGGTGGG AAAGTGA	
PDHX-APIP	*Pdhx-Apip*			other site near TSS
PHB2-EMG1	*Phb2-Emg1*	GGGACTTCCGTATGCGCGATTCCTGTGC	GGGACTTCCGGATGCGCCATTCCTGTGC	
PMPCA-SDCCAG3	*Pmpca-Sdccag3*	TGGGCGGAAGCGGAAGTGAC	GGGTCGGAAGCGGAAGTGAC	
PRDX5-TRMT112	*Prdx5-Trmt112*			other site near TSS
PRKCQ-PRKCQAS1	*Prkcq*			other site near TSS
PRPF31-TFPT	*Prpf31-Tfpt*			other site near TSS
PSMD10-ATG4A	*Psmd10-Atg4a*			other site near TSS
PTCD1-CPSF4	*Ptcd1-Cpsf4*			other site near TSS
PTCD3-POLR1A	*Ptcd3-Polr1a*			other site near TSS
PTK2B-TRIM35	*Ptk2b-Trim35*			other site near TSS
PTRH2-VMP1	*Ptrh2-Vmp1*	GATCCGGAACTTGTCACCCAGGAACCCCGGAAGAGGT	GATCCGGAACTAGTTTTTCAGGAAACCCGGAAGGAGT	
RMRP-CCDC107	*Rmrp-Ccdc107*	CTCTGTTCCTCCCCTTCCGCCTAGGGAAAGTCC	ACATGTTCCTTATCCTTTCGCCTAGGGAAAGTCCC	
RNF185-MIR3928	*Rnf185-Mir3928*	ACCTCTTCCGGTTGGGGCCGGAAGTCCC	GCCTTTTCCGGCTGGAACCGGAAGTTGT	
ROM1-EML3	*Rom1-Eml3*	AGAGGGGAAGGGGAAGCACC	AGAGGGGAA--GGAAGCACT	
RPS6KB1-TUBD1	*Rps6kb1-Tubd1*	CGGACTTCCGAGACAGGGAAGCTGA	CTGACTTCCGACACAGGGAAGCTGA	
RTN4IP1-QRSL1	*Rtn4ip1-Qrsl1*			other site near TSS
SCO1-ADPRM	*Sco1-Adprm*	CCGACTTCCGGAGGAAGCGGAACGCTACCGGAAATCGC	TCGGCTTCCGGCAGAAGCGGAAGCTCGAGCGGAAATGGA	
SDHAF2-CPSF7	*Sdhaf2-Cpsf7*	AGGAGTTCCCGGAAGTGCC	CGGAATTCCCGGAAGTGGC	
SERAC1-GTF2H5	*Serac1-Gtf2h5*			other site near TSS
SHC1-CKS1B	*Shc1-Csk1b*			None
SIRT3-PSMD13	*Sirt3-Psmd13*			other site near TSS
SKIV2L-RDBP	*Skiv2L-Rdbp*	GTACCGGAAGTTGCCTCTACTTCCGCCCG	GTGTCGGAAGTAATTTCAACTTCCGCCCT	
SLC25A11-RNF167	*Slc25a11-Rnf167*	CCGTGTTCCCAGCCTCTGGAAAGGGCTTCCGGTAG	CTGTCTTCCTATGCCCCTGGAAAAGGGTTTCCGGTAA	
SLC25A27-CYP39A1	*Slc25a27-Cyp39a1*			other site near TSS
SLC25A32-DCAF13	*Slc25a32-Dcaf13*	GCGACTTCCGCTTTTCCCAGACTACTTCCAGTCA	ACTACTTCCGCC-TTCCGCG--GGCTTCCTGTCA	
SQSTM1-MGAT4B	*Sqstm1-Mgat4b*			None
SSBP1-FLJ40852	*Ssbp1*	GCGGAGTTTCTGT--TTCCTTTTTCCTCTGG	GAGATTTCCTGTCTTTCCTTG--CATCTCG	
STARD7-LOC285033	*Stard7-Gm10766*	CTGCGCCCCTCCGGACTGGTTCCTTGGGCCCCGGAAGCTCG	CTGCCTTCCCCCAGCCC--TTCCT----CTCCGGGGCCCCG	
TAP1-PSMB9	*Tap1-Psmb9*			other site near TSS
TFAP2A-LOC100130275	*Tfap2a*			other site near TSS
TFB2M-CNST	*Tfb2m-Cnst*	GAGGCGGAAGCGGAAGTGAG	TAGGCGGAAGCGGAAGCGAG	
TIMM8B-SDHD	*Timm8b-Sdhd*	GACGGGGAGGGTGAAGGGAAAGGAAGTTTC	GATGAGGAAGGAACAGGGGAAGGAGGGTGA	
TMEM186-PMM2	*Tmem186-Pmm2*	CACGAGGAACTCGGCCCGGAACTTCCGGGTT	TACCCGGAACTCTACCCGGAACTTCCGGGTC	
TOMM70A-LNP1	*Tomm70a*			other site near TSS
TP53-WRAP53	*Tp53-Wrap53*			other site near TSS
TRAK2-STRADB	*Trak2-Stradb*	GGGGCGGGCAGGAAACTACAATTCCCAGCA	GGGACGGAAAGCAAACTACAGTTCCCAGAA	
TRMT5-SLC38A6	*Trmt5-Slc38a6*	GGCCCTTCCGGCATTCCGTA-CTTCACCAGGGCCTGGAAGGAGA	GGCCCTTCCGGGCTCTCGTAGCAAATCGCGGGCCCGGAA-GAGA	
TRPV4-MIR4497	*Trpv4*			None
TUBA1C-LOC100293962	*Tuba1c*	TCCTCTTCCTGCTCCTGGCTCCTTCCGACGA	TCCCCTTCCCACTTCCAGCTCCTTCTACGGA	
TXNRD2-COMT	*Txnrd2-Comt*			None
UQCR10-ZMAT5	*Uqcr10-Zmat5*	CGCCCTTCCCAGAGAGCTTTGGGAAT-CTA	C----CTT---CCTAAGGTA------CCGAGAAGGC---- GGAATACGA	
USMG5-PDCD11	*Usmg5-Pdcd11*			other site near TSS
WARS-WDR25	*Wars-Wdr25*	GGGCCGGAAGTTGTTCCGTCCGGACGGCGTTTCCACGGA	AGACCGGAAGTTATTCCAGCCGGCTGGTGTTTCCACGGA	
WASF1-CDC40	*Wasf1-Cdc40*	TCATCTTCCCTCATTCCCTAGC	TCATCTTCCCTCATTCCCGAGC	
WDR92-PNO1	*Wdr92-Pno1*			None
WDR93-PEX11A	*Wdr93-Pex11a*	CGGACTTCCGGTTCAAGCCGGAAGTTGT	TGGAGTTCCGGTTAGTACCGGAAGTGGT	
YME1L1-MASTL	*Yme1l1-Mastl*	GCCCCGGAAGTACTGTTGAGTTAGCGCCTCGCCTTCCGGGGC	GCTCCGGAAGCACTGT--ACCTA-----TGGTCTTCCGGGGCA	
YRDC-C1orf122	*Yrdc-1110065P2Rik*	GTCACTTCCTCCCGGAAGCGGG	GTCACTTCCTCCCGGAAGCGGC	

Upstream regions of mouse genes that correspond to that of the 151 human genes in Table 2 were retrieved from NCBI database. Then, comparison of the duplicated GGAA motifs was carried out by BLAST sequence analysis. Conserved sequences both in the human and the mouse genomes are shown. Although the corresponding sequences are not observed in in the 47 mouse genes, another duplications are found at other site near TSSs of mouse genes. No obvious GGAA duplications are

found in 24 mouse genes. Shadowed mouse gene names indicate that they have no partners.

6. Tanscription factors that may regulate bidirectional promoters of DNA repair factor encoding genes and interferon stimulated genes

Previously, we have identified that *ATM-NPAT*, *APEX1-OSGEP*, and *BRCA1-NBR2* gene pairs are linked with each other by bidirectional promoters [15]. As shown in Table 2, it was revealed that the DNA-repair associated genes, such as *ALKBH1*, *BRE*, *MSH3*, *MTBP*, *CDKN2A*, and *KLLN*, are located upstream of mitochondrial protein encoding genes in a head-head configuration. It has been reported that many cancer or DNA repair associated genes have bidirectional partner genes, and that tandem repeated ETS binding sites are frequently found in the 5'-upstream regions of both genes [36-38,41]. Therefore, expression of many DNA repair factor encoding genes is thought to be regulated by duplicated GGAA motifs in their 5'-upstream or promoter regions.

Surveillance of the human genomic sequence database revealed that several interferon (IFN) stimulated genes (ISGs) have bidirectional partner genes [39]. Similar to the bidirectional promoters involved with DNA repair factor encoding genes, bidirectional ISG promoters contain duplicated GGAA motifs. They are *BAG1-CHMP5*, *BLZF1-NME7*, *EIF3L-ANKRP54*, *CCDC75-HEART5B*, *IFI27L1-DDX24*, *PARP10-PLEC*, *PSMA2-MRPL32*, *RPL22-RNF207*, and *TRADD-FBXL8* [39]. Some of them are listed in Table 1 and Table 2. The *WARS-WDR25* gene pair could be added to this bidirectional ISG group, because *WARS* is also named *IFI53*. The *TOMM70A* gene, whose promoter is linked with that of the *LNP1* (Table 2), might be associated with the antiviral response, because TOM70, a mitochondrial import receptor, has been shown to import antiviral immunity to the mitochondria activating IRF3 [42,43]. It is noteworthy that the bidirectional gene pair *HSPD1-HSPE1*, which encodes the mitochondrial chaperon proteins HSP60 and HSP10, respectively, has been reported to be regulated by interferon (IFN) gamma [44]. The B-cell translocation gene *BTG2* encodes a protein that acts as a proliferation inhibitor [45] and it is listed in Table 2. Moreover, *APOPT1*, *ATG4A*, *BOK*, *KLLN*, *PDCD11*, *TP53*, and *TRIAP1*, as listed in Table 2, are suggested to play roles in the progression of apoptosis or autophagy. Therefore, these genes might be regulated in accordance with the mitochondrial function associated genes in response to immunologically induced signals to stop proliferation or execute cell-death. These findings suggest that the mitochondrial function associated gene promoters carrying duplicated GGAA-motifs could be also regulated by IFN-induced signals. It has been suggested that an antiviral signal to evoke type I IFN gene expression is mediated by a MAVS (mitochondrial antiviral signaling) protein [46]. Although, no bidirectional partner is found upstream of the human *MAVS* gene, we have confirmed that the sequences 5'-ACTTG**GGAA**GCGTGGGGAT**GGAA**TTCTC-3' and 5'-CGGAC**TTCC**CCT**GGAA**GTTGC-3' are present within 300-bp upstream from the most 5'-upstream of the gene. Interestingly, recent study showed that MAVS protein is a potent inhibitor for apoptosis regulating caspase activity [47].

A lot of TFs, especially ETS family proteins, are known to recognize and bind to DNA elements containing GGAA as a core motif [48]. Not only ETS family proteins, but also NF-B/REL [49,50], STAT proteins [51,52], IRF proteins [53], and HSF1/2 [54] can bind to DNA sequences which harbor the GGAA core motif [39], to regulate transcription. Gene expression and binding analysis suggested that STAT1 plays a role in the regulation of bidirectional promoters [55]. In addition, NRF2 (GABP) has been shown to regulate bidirectional transcription of the COX4/NOC4 gene pairs [56]. Although

our survey of the human genome database did not retrieve the COX4/NOC4 gene pair because the distance between them is over 500 nucleotides, the sequence 5'-CGGCT**TTCC**AGCCT**GGAA**GCGCC-3' is located with multiple GGAA-motifs and Sp1 binding sequences [56]. The most frequently found sequence co-localized with ETS binding motifs in human promoters is the Sp1 element with 28.4% occurrence [57]. In accordance with these observations, the duplicated GGAA motif is located at the center of the bidirectional promoter region of *SIRT3-PSMD13* genes surrounded by multiple Sp1 binding sequences [58]. It has been reported that co-operation of the GABP binding site with Sp1/3 and YY1 binding sites plays a role in murine *Gabpa-Atp5j* bidirectional promoter activity [59]. Not only the *Gabpa-Atp5j*, but also the human and murine *Surf1-Surf2* bidirectional promoter, has been suggested to be affected by the co-operation of ETS proteins and YY1 [60]. Moreover, the positive regulatory effect of Sp1 on the *HADHA-HADHB* bidirectional promoter has been reported [61]. The TFs that should be noted as bidirectional transcription regulatory factors are NF-Y [62] and ZNF143 [34], which have been shown to regulate human *MRPS12-SARS2* and *TMEM186-PMM2* bidirectional promoter activity, respectively.

These observations suggest that various TFs, including Sp1, YY1, and other proteins co-operatively work with GGAA-binding factors to regulate the GGAA-motif containing bidirectional promoters. This concept is consistent with the "enhanceosome" that is thought to be involved in the regulation of eukaryotic TATA-less promoters [63].

7. Mitochondria play important roles in the responses to various stresses

Dysfunction or shortening of mammalian telomeres causes p53-mediated suppression of PGC-1, which in turn causes mitochondrial dysfunction to overproduce reactive oxygen species (ROS) [64]. Cellular senescence is thought to be accelerated by telomere shortening and hyper ROS generation [65,66]. In addition, not only telomere-originated signals but also DNA damage on mitochondrial function-associated genes might directly affect various mitochondrial functions. Most of the mitochondrial protein components are translated from mRNAs that have been transcribed from the nuclear genome [4,67]. Recent study suggested that an imbalance between mitochondrial and nuclear proteins exerts a signal to nuclear DNA to induce expression of genes encoding stress-responsive proteins, which in-turn evoke a beneficial condition for longevity of host organisms [67,68]. The mitochondrial imbalance might induce hormesis that is referred to a beneficial outcome from low doses of toxic or other harmful damage, including irradiation, heat shock, or food restriction [69]. Interestingly, a natural compound resveratrol, which has a toxic effect if used in high doses [70], causes the mitochondrial imbalance [68] that might elongate life spans of various organisms [71,72,73]. However, excess damage on chromosomal DNAs will be disadvantageous to mitochondria in which it could lead to dysfunction. If the damage to chromosomes was so severe that mitochondrial protein encoding genes could not produce correct mRNAs, mitochondria might have to exert signals to stop proliferation or to cause cell death. Those events induced by damage on the chromosomal DNAs could lead to apoptosis or autophagy. DNA damage responding signals may affect mitochondrial proteins, such as Bcl-2, Bcl-XL, BAX, and cytochrome c, to control cell death [74,75,76]. Not only apoptosis, but also programmed necrosis (necroptosis) is executed by signals that are exerted from mitochondria [77]. Mitochondria send danger signals to induce not only inner cellular responses but also several extracellular danger signals, including IFN production, inflammasome activation, and neutrophil activation [78]. Moreover, it is noteworthy that active p53,

which is widely known as a tumor suppressing factor, induces transcription of many mitochondrial function associated genes [79]. The other stress may come from neutrients condition or metabolites. The metabolites, including acetyl-CoA, S-adenosylmethionine, and NAD+, are known to affect gene expression [80]. Moreover, it was suggested that the metabolism plays a role in regulating immunity [81]. Thus mitochondria have an important function as stress sensing machinery in a cell. The scenario may partly explain the reason why a lot of mitochondrial function associated genes are regulated by duplicated GGAA motif-containing bidirectional promoters in a similar manner to those genes, including DNA repair-, and apoptosis inducing-factor encoding genes, and ISGs (Fig. 2).

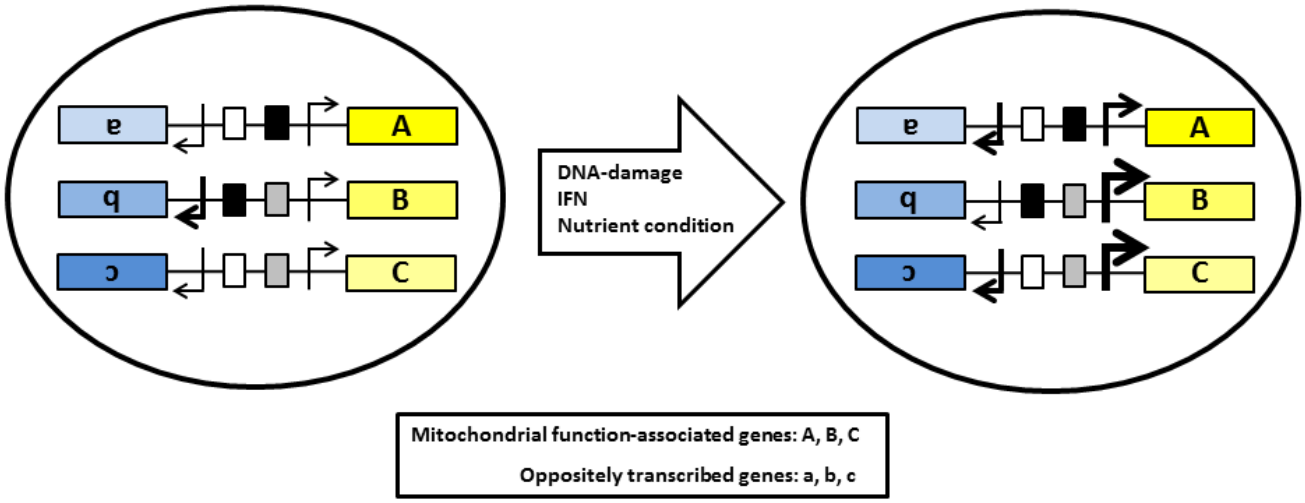

Figure 2. Hypothetical mechanism to regulate expression of genes encoding mitochondrial function-associated proteins. Various stresses, including DNA-damage, IFN-stimulation, and condition of nutrients, may affect transcriptional state in a cell nucleus via various duplicated GGAA motifs. A common signal will alter the activity or quantity of specific transcription factors (TFs) that bind to different GGAA-core motifs. Each promoter may individually respond to the same signal. Although it has not been examined how the direction of transcription from the bidirectional promoter regions are determined or controlled, expression of the bidirectional partner genes (a, b, and c) may co-operate with mitochondrial function-associated genes (A, B, and C) to respond to the induced signals correctly. Arrows indicate transcription start sites. Transcription rates are schematically shown by sizes.

8. Determinants of direction of transcription in mammalian cells

In this article, we have focused on the bidirectional promoters of mitochondrial function-associated genes to find that duplicated GGAA-motifs are very frequently located near TSSs of both promoters. Recently, it was reported that gene-loop formation, which is conducted by an interaction between protein factor Ssu72 and pre-initiation complex, determines the direction of transcription [82]. Gene-loop is suggested to be generated by a juxtaposition of a terminator with its

promoter to make transcription machinery move to one direction [83,84]. The gene-loop formation is thought to be an effective system to recycle transcription machinery repeatedly [85]. Long terminal repeat (LTR) sequences of HIV have been shown to form a gene-loop structure to produce strong promoter activity after integration into the host cellular genome [86]. The gene-loop might be also generated by some interactions between specific sequences that are located at 5' and 3' untranslated regions (UTRs), which are thought to regulate gene expression [87]. Additionally, retrotransposons or transposable elements [88] might be taken into account for understanding the mechanism to generate gene-loops. However, if there were no terminator sequences around a specific gene, or if gene-loop formation was prevented by some sequences/siRNAs, its promoter would allow concomitant bidirectional transcription. Although it is yet to be elucidated, the duplicated GGAA-motifs might function to prevent gene-loop formation.

Given that bidirectional promoters do not naturally form a loop structure, the above observations suggest that some of the mitochondrial function-associated genes have been anchored to the chromosomal DNAs where gene-loop formation is somewhat prevented. Interaction between double stranded DNAs by gene-loop formation might cause a circumstance where circular DNA could be easily released from only two reactions, namely by endonuclease and by DNA ligase, similar to the genomic rearrangement system in immune cells [89]. It should be noted that the gene-loop also plays a part in the double stranded break formation in the meiotic recombination system [90]. Thus, it seems that mitochondrial genes in the chromosomes would remain where they are located presently with bidirectional partner genes. The concept that bidirectional promoter partner genes remain located at the same region of the chromosome seems to be consistent with the observation that Histone protein-encoding genes are linked together by bidirectional promoters (Table 2).

The direction of the RNA pol II at divergent transcription initiation sites in mammalian promoters has been suggested to be determined by well-controlled biological systems [91,92,93]. Recent study showed that upstream antisense RNAs are cleaved and polyadenylated at poly(A) sites (PASs) [94]. In addition, for the sense direction, PAS signals and U1 small nuclear ribonucleoprotein recognition sites are depleted and enriched, respectively [94]. Similar conclusion that the transcriptional direction is affected by the PAS signals was obtained from the analysis of human genome-wide map of promoter-upstream transcripts (PROMPTs) whose transcription initiates from bidirectional promoter activity [95]. Moreover, mutation analysis indicated an element that blocks the reverse transcription of the mouse Ide (insulin-degrading enzyme) gene [96]. These observations suggest that specific cis-acting elements near the RNA pol II binding sites could prevent bidirectional transcription. In other words, absence of these transcriptional-direction regulating sequences is necessary for the bidirectional transcription.

9. Origin of mitochondrial function-associated genes

Proteomic analysis revealed that mitochondria are composed of mosaic of endosymbiotic, non-proteobacterial, and orphan proteins [97]. Mitochondria have specific features such that resemble bacteria, having a double membrane and a circular genome encoding 13 proteins, 22 tRNAs and 2 rRNAs [4]. From the comparison of the small subunit rRNA and the heat-shock protein 60 (HSP90) sequences, it has been suggested that ancestors of mitochondria are α-proteobacteria like cells [98,99]. Among at least 1,100 mitochondrial proteins encoded by the nuclear genome, 400

proteins have a proteobacterial origin, determined by whole genomic sequence analysis of Rickettsia prowazekii, which is thought to be the closest living relative of the ancestral proteobacterial species [98]. Although the protein encoding regions of the Rickettsia genome have been compared with those of eukaryotes, the analysis of the non-protein coding sequences has not yet been performed. The non-coding DNA in the Rikettsia genome has been estimated at 23.7%, which is relatively higher than that of other bacteria (6 to 13%) [100]. During the long process of evolution, non-coding DNAs might have been eliminated from these ancestral bacterial organisms. At present, it is very difficult for us to show directly how protein encoding genes of the α-proteobacteria and other aerobic bacterial organisms have been incorporated into nuclear genomes of eukaryotes. This has been obscured by the long time period of evolution. However, the concept of horizontal gene transfer (HGT), which enables the acquisition of novel traits beyond species, has been postulated to explain transfer of genes from bacteria to eukaryotes by endosymbiosis [101]. Recent study on the analysis of genomic sequence of eukaryotic unicellular red algae supports the idea that HGT facilitated evolution or adaptation to severe environment [102]. The gene transfer from the genome of mitochondrial endosymbiont to the chromosomes of the host could have promoted or modulated the evolution of eukaryotes [103,104]. Contrary to bacterial organisms that have discarded both coding- and non-coding genes, eukaryotes seem to have evolved through receiving exogenous genes and incorporated them at the position where they could not easily be released from the genome. The molecular mechanisms of the incorporating process might have been executed by retroviral integration- or transposon-like systems. The possible scenario for incorporation of the mitochondrial function-associated genes into the eukaryotic genomes is as follows:

1. Protein-coding genes that have originated from α-proteobacteria and other bacterial organisms were inserted into chromosomes of ancestor eukaryotic cells. This could have easily occurred by retroviral LTR or transposon-like gene integration systems.
2. Integration of those genes occasionally occurred at the site where gene looping would not occur by some sequences, including duplicated GGAA (TTCC) elements.
3. Transposable elements or LTR-like sequences were nearly completely lost through evolution of eukaryotic cells except the elements required for recombination in immune systems and meiosis.

The above hypothesis might partly explain the reason why nuclear genes encoding mitochondrial function-associated proteins are regulated by bidirectional promoters.

10. Mitochondrial function and diseases

Recent studies of the human genome, including the ENCODE project, have shown not only protein-encoding genes but also non-coding genes play roles in the regulation of various nuclear events [105]. In the field of medicine, analyses of genomic DNAs are expected to be very useful and powerful diagnostic techniques that would suggest the most suitable treatments for specific diseases including varieties of cancer. In this article, we indicated that a large number of mitochondrial function-associated genes are linked with partner genes by bidirectional promoters, in which frequently GGAA-motif duplications are located. The GGAA-motif containing elements are not only recognized by ETS family proteins but also by NF-B/cREL, IRFs, STAT proteins, and so on. This finding suggests that DNA damage or IFN-induced signals may also affect mitochondria by the alteration of expression of the mitochondrial function-associated genes. It has been revealed that metabolic reactions and signal transduction systems in cancer cells, including not only glycolysis and

oxidative phosphorylation (OXPHOS) but also mTOR/AMPK pathways, are altered from that of normal cells [106,107]. Not only cancer, but also heart/cardio vascular diseases [108,109] and neurodegenerative diseases, including Alzheimer's disease, Parkinson's disease and Huntington's disease [110] are suggested to be caused by dysfunction of mitochondria. The Parkinson's-disease-related kinase, PTEN-induced kinase 1 (PINK1) is known to induce mitochondrial biogenesis and reduce mitochondria-induced apoptosis in neurons [111,112,113]. Recently, it was reported that LKB1 and NUAK1 kinases regulate cortical axon branching through mitochondrial immobilization, suggesting that mitochondrial function affects neural circuits [114]. These observations suggest that mitochondrial dysfunctions could cause various diseases. Thus, novel treatments for these diseases are expected by ameliorating mitochondrial functions. We hope our findings that duplicated GGAA motifs are frequently present in the head-head configured human mitochondrial protein-encoding genes will contribute to the goal.

11. Conclusions

In this article, approximately one-third of genes that have bidirectional partner genes were suggested to associate with mitochondrial functions. We further confirmed that duplicated GGAA motifs are very frequently found in the mitochondrial function-associated bidirectional promoters. At present, biological significance of the duplicated GGAA motifs in the bidirectional promoter regions has not been known yet. However, the motifs are very often found in various DNA-repair/IFN-responding gene promoters, implying that mitochondrial function-associated genes are regulated in concert with DNA repair synthesis or IFN-induced signals. Moreover, mitochondria do not only respond to stresses but also induce signals by modulating amounts of metabolites in accordance with a condition of nutrients. The metabolites, including acetyl-CoA, S-adenosylmethionine, and NAD^+, will in turn affect expression of various genes. Here we propose a putative role of the bidirectional, GGAA motif-containing promoters that respond to stress signals to modulate mitochondrial functions (Fig. 2).

We hope that our findings on the mitochondrial function-associated gene promoters will contribute to studies in both molecular and clinical biology in the future.

Acknowledgments

We are grateful to Eriko Yokoyama, Yoshiko Ogura, Maya Okada, and Megumi Fukui for their outstanding technical assistance. This work was supported in part by JSPS KAKENHI Grant Number 24510270 and a Research Fellowship from the Research Center for RNA Science, RIST, Tokyo University of Science.

Conflict of Interest

The authors declare that there are no conflicts of interest related to this study.

References

1. Carey MF, Peterson CL, Smale ST (2009) Transcription and preinitiation complex assembly in vitro, In: Carey, M.F., Peterson, C.L., Smale, S.T. Eds., *Transcriptional Regulation in Eukaryotes,* 2nd Ed, New York: Cold Spring Harbor Laboratory Press, 439-538.
2. Scarpulla RC, (2008) Transcriptional paradigms in mammalian mitochondrial biogenesis and function. *Physiol Rev* 88：611-638.
3. Campbell CT, Kolesar JE, Kaufman BA, （2012） Mitochondrial transcription factor A regulates mitochondrial transcription initiation, DNA packaging, and genome copy number. *Biochim Biophys Acta* 1819：921-929.
4. Vafai SB, Mootha VK, (2012) Mitochondrial disorders as windows into an ancient organelle. *Nature* 491: 374-383.
5. Pagliarini DJ, Calvo SE, Chang B et al. (2008) A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134: 112-123.
6. Schmidt O, Pfanner N, Meisinger C, (2010) Mitochondrial protein import: from proteomics to functional mechanisms. *Nat Rev Mol Cell Biol* 11: 655-667.
7. Reinders J, Zahedi RP, Pfanner N et al. (2006) Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics. *J. Proteome Res* 5: 1543-1554.
8. Evans MJ, Scarpulla RC (1989) Interaction of nuclear factors with multiple sites in the somatic cytochrome c promoter. Characterization of upstream NRF-1, ATF and intron Sp1 recognition sites. *J Biol Chem* 264: 14361-14368.
9. Scarpulla RC (1997) Nuclear control of respiratory chain expression in mammalian cells. *J Bioenerg Biomembr* 29: 109-119.
10. Whatcott CJ, Meyer-Ficca ML, Meyer RG, et al. (2009) A specific isoform of poly(ADP-ribose) glycohydrolase is targeted to the mitochondrial matrix by a N-terminal mitochondrial targeting sequence. *Exp Cell Res* 315: 3477-3485.
11. Meyer RG, Meyer-Ficca ML, Jacobson EL, Jacobson MK (2003) Human poly(ADP-ribose) glycohydrolase (PARG) gene and the common promoter sequence it shares with inner mitochondrial membrane translocase 23 (TIM23). *Gene* 314: 181-190.
12. Uchiumi F, Sakakibara G, Sato J, et al. (2008) Characterization of the promoter region of the human *PARG* gene and its response to PU.1 during differentiation of HL-60 cells. *Genes Cells* 13: 1229-1248.
13. Uchiumi F, Enokida K, Shiraishi T, et al. (2010) Characterization of the promoter region of the human IGHMBP2 (Sμbp-2) gene and its response to TPA in HL-60 cells. *Gene* 463: 8-17.
14. Koc EC, Burkhart W, Blackburn K, et al. (2001) Large subunit of the mammalian mitochondrial ribosome. *J Biol Chem* 276: 43958-43969.
15. Uchiumi F, Larsen S, Tanuma S (2013) Biological systems that control transcription of DNA repair-and telomere maintenance-associated genes, In: Chen, C. Ed., *DNA repair*, Rijeka, Croatia, InTech Open Access Publisher, 309-325.
16. Matoba S, Kang JG, Patino WD, et al. (2006) p53 regulates mitochondrial respiration. *Science* 312: 1650-1653.
17. Green DR, Kroemer G (2009) Cytoplasmic functions of the tumor suppressor p53. *Nature* 458: 1127-1130.

18. Sahin E, DePinho RA (2012) Axis of aging: telomeres, p53 and mitochondria. *Nat Rev Mol Cell Biol* 13, 397-404.
19. Szczesny B, Tann AW, Longley MJ, et al. (2008) Long patch base excision repair in mammalian mitochondrial genomes. *J Biol Chem* 283: 26349-26356.
20. Martin SA (2011) Mitochondrial DNA repair, In: Storici, F. Ed., *DNA Repair - On the Pathways to Fixing DNA Damage and Errors*, Rijeka, Croatia: InTech-Open Access Publisher, 313-338.
21. Sementchenko VI, Watson DK (2000) Ets target genes: past, present and future. *Oncogene* 19: 6533-6548.
22. Baillat D, Laitem C, Leprivier G, et al. (2009) Ets-1 binds cooperatively to the palindromic Ets-binding sites in the p53 promoter. *Biochem Biophys Res Commun* 378: 213-217.
23. Molloy-Simard V, St-Laurent JF, Vigneault F, et al. (2012) Altered expression of the poly(ADP-ribosyl)ation enzymes in uveal melanoma and regulation of PARG gene expression by the transcription factor ERM. *Invest Ophthamol Vis Sci* 53: 6219-6231.
24. Uchiumi F, Watanabe T, Tanuma S (2010) Characterization of various promoter regions of the human DNA helicase-encoding genes and identification of duplicated ets (GGAA) motifs as an essential transcription regulatory element. *Exp Cell Res* 316: 1523-1534.
25. Uchiumi F, Miyazaki S, Tanuma S (2011) The possible functions of duplicated ets (GGAA) motifs located near transcription start sites of various genes. *Cell Mol Life Sci* 68: 2039-2051.
26. Smeitink JA, Loeffen JL, Triepels RH, et al. (1998) Nuclear genes of human complex I of the mitochondrial electron transport chain: state of the art. *Hum Mol Genet* 7: 1573-1579.
27. Starker LF, Delgado-Verdugo A, Udelsman R, et al. (2010) Expression and somatic mutations of SDHAF2 (SDH5), a novel endocrine tumor suppressor gene in parathyroid tumors of primary hyperparathyroidism. *Endocrine* 38: 397-401.
28. Harris KA, Jones V, Bilbille Y, et al. (2011) YrdC exhibits properties expected of a subunit for a tRNA threonylcarbamoyl transferase. *RNA* 17: 1678-1687.
29. Yang C, Bolotin E, Jiang T, et al. (2007) Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene* 389: 52-65.
30. Trinklein ND, Aldred SF, Hartman SJ, et al. (2004) An aboundance of bidirectional promoters in the human genome. *Genome Res* 14: 62-66.
31. Yang MQ, Elnitski LL (2008) Diversity of core promoter elements comprising human bidirectional promoters. *BMC Genomics* 9 (Suppl. 2): S3.
32. Collins PJ, Kobayashi Y, Nguyen L, et al. (2007) The ets-related transcription factor GABP directs bidirectional transcription. *PLoS Genet* 3: e208.
33. Lin JM, Collins PJ, Trinklein ND, et al. (2007) Transcription factor binding and modified histones in human bidirectional promoters. *Genome Res* 17: 818-827.
34. Anno YN, Myslinski E, Ngondo-Mbongo RP, et al. (2011) Genome-wide evidence for an essential role of the human Staf/ZNF143 transcription factor in bidirectional transcription. *Nucleic Acids Res* 39: 3116-3127.
35. Orekhova AS, Rubtsov PM (2013) Bidirectional promoters in the transcription of mammalian genomes. *Biochem (Moscow)* 78: 335-341.
36. Adachi N, Lieber MR (2002) Bidirctional gene organization: a common architectural feature of the human genome. *Cell* 109: 807-809.

37. Wakano C, Byun JS, Di LJ, et al. (2012) The dual lives of bidirectional promoters. *Biochim Biophys Acta* 1819: 688-693.
38. Xu C, Chen J, Shen B (2012) The preservation of bidirectional promoter architecture in eukaryotes: what is the driving force? *BMC Systems Biol* 6 (Suppl. 1): S21.
39. Uchiumi F, Larsen S, Masumi A, et al. (2013) The putative implications of duplicated GGAA-motifs located in the human interferon regulated genes (ISGs), In: iConcept Ed., *Genomics I- Humans, Animals and Plants*, Hong Kong: iConcept Press, 87-105.
40. Gotea V, Petrykowska HM, Elnitski L (2013) Bidirectional promoters as important drivers for the emergence of species-specific transcripts. *PLoS One* 8: e57323.
41. Yang MQ, Koehly LM, Elinitski LL (2007) Comprehensive annotation of bidirectional promoters identifies co-regulation among breast and ovarian cancer genes. *PLoS Comput Biol* 3, e72.
42. Lin R, Paz S, Hiscott J (2010) Tom70 imports antiviral immunity to the mitochondria. *Cell Res* 20: 971-973.
43. Liu XY, Wei B, Shi HX, et al. (2010) Tom70 mediates activation of interferon regulatory factor 3 on mitochondria. *Cell Res* 20: 994-1011.
44. Kim SW, Kim JB, Kim JH, et al. (2007) Interferon-gamma-induced expressions of heat shock protein 60 and heat shock protein 10 in C6 astroglioma cells: identification of the signal transducers and activators of transcription 3-binding site in bidirectional promoter. *Neuroreport* 18: 385-389.
45. Mauxion F, Chen CY, Séraphin B, et al. (2009) BTG/TOB factors impact deadenylases. *Trends Biochem Sci* 34: 640-647.
46. Koshiba T (2013) Mitochondrial-mediated antiviral immunity. *Biochim Biophys Acta* 1833: 225-232.
47. Okazaki T, Higuchi M, Gotoh Y (2013) Mitochondrial localization of the antiviral signaling adaptor IPS-1 is important for its induction of caspase activation. *Genes Cells* 18: 493-501.
48. Wei GH, Badis G, Berger MF, et al. Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J* 29: 2147-2160.
49. Chen YQ, Sengchanthalangsy LL, Hackett A, et al. (2000) F-kappaB p65 (RelA) homodimer uses distinct mechanisms to recognize DNA targets. *Structure* 8: 419-428.
50. Wei L, Fan M, Xu L, et al. (2008) Bioinformatic analysis reveals cRel as a regulator of a subset of interferon-stimurated genes. *J Interferon Cytokine Res* 28: 541-552.
51. Nguyen VT, Benveniste EN (2000) Involvement of STAT-1 and Ets family members in interferon-γ induction of CD40 transcription in microglia/macrophages. *J Biol Chem* 275: 23674-23684.
52. Aittomäki S, Yang J, Scott EW, et al. (2004) Molecular basis of Stat1 and PU.1 cooperation in cytokine-induced Fcgamma receptor 1 promoter activation. *Int Immunol* 16: 265-274.
53. Meraro D, Gleit-Kielmanowicz M, Hauser H, et al. (2002) IFN-stimulated gene 15 is synergistically activated through interactions between the myelocyte/lymphocyte-specific transcription factors, PU.1, IFN regulatory factor-8/IFN consensus sequence binding protein, and IFN regulatory factor-4: characterization of a new subtype of IFN-stimulated response element. *J Immunol* 168: 6224-6231.
54. Kroeger PE, Morimoto RI (1994) Selection of new HSF1 and HSF2 DNA-binding sites reveals differences in tumor cooperativity. *Mol Cell Biol* 14: 7592-7603.

55. Wang G, Qi K, Zhao Y, et al. (2013) Identification of regulatory regions of bidirectional genes in cervical cancer. *BMC Med Genet* 6 (Suppl. 1): S5.
56. Bachman NJ, Wu W, Schmidt TR, et al. (1999) The 5' region of the COX4 gene contains a novel overlapping gene, NOC4. *Mamm Genome* 10: 506-512.
57. FitzGerald PC, Shlyakhtenko A, Mir AA, et al. Clustering of DNA sequences in human promoters. *Genome Res* 14: 1562-1574.
58. Bellizzi D, Dato S, Cavalcante P, et al. (2005) Characterization of a bidirectional promoter shared between two human genes related to aging: SIRT3 and PSMD13. *Genomics* 89: 143-150.
59. Patton J, Block S, Coombs C, et al. (2005) Identification of functional elements in the murine Gabp alpha/ATP synthase coupling factor 6 bi-directional promoter. *Gene* 369: 35-44.
60. Gaston K, Fried M (1994) YY1 is involved in the regulation of the bi-directional promoter of the Surf-1 and Surf-2 genes. *FEBS Lett* 347: 289-294.
61. Orii KE, Orii KO, Souri M, et al. (1999) Genes for the human mitochondrial trifunctional protein α- and β-subunits are divergently transcribed from a common promoter region. *J Biol Chem* 274: 8077-8084.
62. Zanotto E, Häkkinen A, Teku G, et al. (2009) NF-Y influences directionality of transcription from the bidirectional Mrps12/Sarsm promoter in both mouse and human cells. *Biochim Biophys Acta* 1789: 432-442.
63. Merika M, Thanos D (2011) Enhanceosomes. *Curr Opin Genet Dev* 11: 205-208.
64. Sahin E, Colla S, Liesa M, et al. (2011) Telomere dysfunction induces metabolic and mitochondrial compromise. *Nature* 470: 359-365.
65. O'Sullivan RJ, Karlseder J (2010) Telomeres: protecting chromosomes against genome instability. *Nat Rev Mol Cell Biol* 11: 171-181.
66. Yu BP, Chung HY (2006) Adaptive mechanisms to oxidative stress during aging. *Mech Ageing Dev* 127: 436-443.
67. Wolff S, Dillin A (2013) Beneficial miscommunication. *Nature* 497: 442-443.
68. Houtkooper RH, Mouchiroud L, Ryu D, et al. (2013) Mitochondrial protein imbalance as a conserved longevity mechanism. *Nature* 497: 451-457.
69. Schmacher B (2009) Transcription-blocking DNA damage in aging: a mechanism for hormesis. *Bioessays* 31: 1347-1356.
70. Calabrese EJ, Mattson MP, Calabrase V (2010) Resveratrol commonly displays hormesis: Occurrence and biomedical significance. *Hum Exp Toxicol* 29: 980-1015.
71. Howitz KT, Bitterman KJ, Cohen HY, et al. (2003) Small molecule activators of sirtuins extend Saccharomyces cerevisiae life span. *Nature* 425: 191-196.
72. Wood JG, Rogina B, Lavul S, et al. (2004) Sirtuin activators mimic caloric restriction and delay aging in metazoans. *Nature* 430: 686-689.
73. Stefani M, Markus MA, Lin RC, et al. (2007) The effect of resveratrol on a cell model of human aging. *Ann NY Acad Sci* 1114: 407-418.
74. Kroemer G (1997) Mitochondrial implication in apoptosis. Towards an endosymbiont hypothesis of apoptosis evolution. *Cell Death Differ* 4: 443-546.
75. Tait SW, Green DR (2010) Mitochondria and cell death: outer membrane permeabilization and beyond. *Nat Rev Mol Cell Biol* 11: 621-632.

76. Estaquier J, Vallette F, Vayssiere JL, et al. (2012) The mitochondrial pathways of apoptosis, In: Scatena, R., Bottoni, P., Giardina B. Eds., *Advances in Mitochondrial Medicine*, Dordrecht, Germany: Springer Science+Business Media B.V., 157-183.
77. Vandenabeele P, Galluzzi L, Berghe TV (2010) Molecular mechanisms of necroptosis: an ordered cellular explosion. *Nat Rev Mol Cell Biol* 11: 700-714.
78. Galluzzi L, Kepp O, Kroemer G (2012) Mitochondria: master regulators of danger signaling. *Nat Rev Mol Cell Biol* 13: 780-788.
79. Buchakjian MR, Kornbluth S (2010) The engine driving the ship: metabolic steering of cell proliferation and death. *Nat Rev Mol Cell Biol* 11: 715-727.
80. Gut P (2013) The nexus of chromatin regulation and intermediary metabolism. *Nature* 502: 489-498.
81. Pearce EL, Poffenberger MC, Chang CH, et al. (2013) Fueling Immunity: Insights into metabolism and lymphocyte function. *Science* 342: 1242454.
82. Tan-Wong SM, Zaugg JB, Camblong J, et al. (2012) Gene loops enhance transcriptional directionality. *Science* 338: 671-675.
83. Ansari A, Hampsey M (2005) A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. *Genes Dev* 19: 2969-2978.
84. Hampsey M (2012) A new direction for gene loops. *Science* 338: 624-625.
85. Lykke-Andersen S, Mapendano CK, Jensen TH (2011) An ending is a new beginning. *Cell Cycle* 10: 863-865.
86. Perkins KJ, Lusic M, Mitar I, et al. (2008) Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. *Mol Cell* 29: 56-68.
87. Barrett LW, Fletcher S, Wilton SD (2012) Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cell Mol Life Sci* 69: 3613-3634.
88. McCue AD, Nuthikattu S, Reeder S.H, et al. (2012) Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA. *PLoS Genet* 8: e1002474.
89. Kapitonov VV, Jurka J (2005) RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biol* 3: 181.
90. Acquaviva L, Székvölgyi L, Dichtl B, et al.（2013）The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. *Science* 339：215-218.
91. Core LT, Waterfall JJ, Lis JT（2008）Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322：1845-1848.
92. Seila AC, Calabrese JM, Levine SS, et al.（2008）Divergent transcription from active promoters. *Science* 322：1849-1851.
93. Preker P, Nielsen J, Kammler S, et al.（2008）RNA exosome depletion reveals transcription upstream of active human promoters. *Science* 322：1851-1854.
94. Almada AE, Wu X, Kriz AJ（2013）Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* 499：360-363.
95. Ntini E, Järvelin AI, Bornholdt J et al.（2013）Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. *Nat Struct Mol Biol* 20：923-928.
96. Zhang L, Ding Q, Wang P, et al. (2013) An upstream promoter element blocks the reverse transcription of the mouse insulin-degrading enzyme gene. *Biochem Biophys Res Commun* 430：26-31.

97. Szklarczyk R, Huynen MA（2010） Mosaic origin of the mitochondrial proteome. *Proteomics* 10：4012-4024.
98. Yang D, Oyaizu Y, Oyaizu H, et al. (1985) Mitochondrial origins. *Proc Natl Acad Sci USA* 82: 4443-4447.
99. Viale AM, Arakaki AK（1994） The chaperone connection to the origins of the eukaryotic organelles. *FEBS Lett* 341：146-151.
100. Andersson SG, Zomorodipour A, Andersson JO, et al. (1998) The genome sequence of Rickettsia prowazekii and the origin of mitochondria. *Nature* 396: 133-143.
101. Dunning Hotopp JC (2011) Horizontal gene transfer between bacteria and animals. *Trends Genet* 27: 157-163.
102. Schönknecht G, Chen WH, Ternes CM, et al. (2013) Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science* 339: 1207-1210.
103. Martin WF (2013) Early evolution without a tree of life. *Biol Direct* 6: 36.
104. Rocha EPC (2013) With a little help from prokaryotes. *Science* 339: 1154-1155.
105. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489: 57-74.
106. Scatena R (2012) Mitochondria and cancer, In: Scatena, R., Bottoni, P., Giardina B. Eds., *Advances in Mitochondrial Medicine*, Dordrecht, Germany: Springer Science+Business Media B.V., 287-308.
107. Schultze A, Harris AL (2012) How cancer metabolism is turned for proliferation and vulnerable to disruption. *Nature* 491: 364-373.
108. Camara AK, Bienengraeber M, Stowe DF (2011) Mitochondrial approaches to protect against cardiac ischemia and reperfusion injury. *Front Physiol* 2: 1-34.
109. Griffiths EJ (2012) Mitochondria and Heart Disease, In: Scatena, R., Bottoni, P., Giardina B. Eds., *Advances in Mitochondrial Medicine*, Dordrecht, Germany: Springer Science Business Media B.V., 249-267.
110. Swerdlow LE, Swerdlow RH (2012) Mitochondria in nerodegeneration, In: Scatena, R., Bottoni, P., Giardina B. Eds., *Advances in Mitochondrial Medicine*, Dordrecht, Germany: Springer Science Business Media B.V., 269-286.
111. Petit A, Kawarai T, Paitel E, et al. (2005) Wild-type PINK1 prevents basal and induced neuronal apoptosis, a protective effect abrogated by Parkinson-disease-related mutations. *J Biol Chem* 280: 34025-34032.
112. Dagda RK, Chu CT (2009) Mitochondrial quality control: insights on how Parkinson's disease related genes PINK1, parkin, and Omi/HtrA2 interact to maintain mitochondrial homeostasis. *J Bioenerg Biomembr* 41: 473-479.
113. Hertz NT, Berthet A, Sos ML, et al. (2013) A neo-substrate that amplifies catalytic activity of Parkinson's-disease-related kinase PINK1. *Cell* 154: 737-747.
114. Courchet J, Lewis TL, Lee S, et al. (2013) Terminal axon branching is regulated by the LKB1-NUAK1 kinase pathway via presynaptic mitochondrial capture. *Cell* 153: 1510-1525.