Curr Genet (2016) 62:809–826
DOI 10.1007/s00294-016-0585-3

ORIGINAL ARTICLE

Complete mitochondrial genomes of two gelechioids, *Mesophleps albilinella* and *Dichomeris ustalella* (Lepidoptera: Gelechiidae), with a description of gene rearrangement in Lepidoptera

Jeong Sun Park[1] · Min Jee Kim[1] · Su Yeon Jeong[1] · Sung Soo Kim[2] · Iksoo Kim[1]

Received: 27 January 2016 / Revised: 20 February 2016 / Accepted: 24 February 2016 / Published online: 7 March 2016

Abstract We sequenced the entire mitochondrial genome (mitogenome) of two gelechioids, *Mesophleps albilinella* and *Dichomeris ustalella*, and compared their genome organization and sequence composition to those of available gelechioid mitogenomes for an enhanced understanding of Gelechioidea genomic characteristics. We compared all available lepidopteran mitogenome arrangements, including that of *M. albilinella*, which is unique in Gelechioidea, to comprehend the extensiveness and mechanisms of gene rearrangement in Lepidoptera. The genomes of *M. albilinella* and *D. ustalella* are 15,274 and 15,410 bp in size, respectively, with the typical sets of mitochondrial (mt) genes. The COI gene begins with CGA (arginine) in all sequenced gelechioids, including *M. albilinella* and *D. ustalella*, reinforcing the feature as a synapomorphic trait, at least in the Gelechioidea. Each 353- and 321-bp long A + T-rich region of *M. albilinella* and *D. ustalella* contains one (*D. ustalella*) or two (*M. albilinella*) tRNA-like structures. The *M. albilinella* mitogenome has a unique gene arrangement among the Gelechioidea: ARNESF (the underline signifies an inverted gene) at the ND3 and ND5 junction, as opposed to the ARNSEF that is found in ancestral insects. An extensive search of available lepidopteran mitogenomes, including that of *M. albilinella,* turned up six rearrangements that differ from those of ancestral insects. Most of the rearrangements can be explained by the tandem duplication-random loss model, but inversion, which requires recombination, is also found in two cases, including *M. albilinella*. Excluding the MIQ rearrangement at the A + T-rich region and ND2 junction, which is found in nearly all Ditrysia, most of the remaining rearrangements found in Lepidoptera appear to be independently derived in that they are automorphic at several taxonomic scales, although current mitogenomic data are limited, particularly for congeneric data.

Keywords Mitochondrial genome · Lepidoptera · Gelechioidea · Gene rearrangement

Introduction

Insect mitochondrial genomes (mitogenomes) have been widely used in systematics, phylogeography, diagnostics, and molecular evolution (Cameron 2014). Moreover, mitogenome gene rearrangement, independent of gene sequencing, has been used for comparative and evolutionary genomics and phylogenetic inference in a diverse array of taxonomic groups (Boore et al. 1995, 1998; Boore 1999; Curole and Kocher 1999; Rokas and Holland 2000; Dowton et al. 2009; Cameron et al. 2011; Cameron 2014). In particular, the Hymenoptera and hemipteroids, which revealed exceptional and abundant gene rearrangement, have been extensively studied for their diversity, taxonomic extent, and phylogenetic signals of gene rearrangement (Dowton et al. 2002; Cameron 2014).

Communicated by M. Kupiec.

Electronic supplementary material The online version of this article (doi:10.1007/s00294-016-0585-3) contains supplementary material, which is available to authorized users.

✉ Iksoo Kim
ikkim81@chonnam.ac.kr

[1] Department of Applied Biology, College of Agriculture and Life Sciences, Chonnam National University, 33, Yongbong-ro, Buk-gu, Gwangju 61186, Republic of Korea

[2] Research Institute for East Asian Environment and Biology, Seoul, 24, Arisu-ro 25na-gil, Gangdong-gu, Seoul 05207, Republic of Korea

In contrast, gene rearrangement in Lepidoptera has received little attention. This is because until 2011, all available lepidopteran mitogenome sequences evidenced only one arrangement: the $tRNA^{Met}/tRNA^{Ile}/tRNA^{Gln}$ (MIQ, underline indicates an inverted gene) at the A + T-rich region and ND2 junction (Kim et al. 2011). This Lepidoptera-specific rearrangement differs from the most common type found in other insects: the IQM arrangement (Boore et al. 1998). However, subsequent investigation by Cao et al. (2012) evidenced the presence of different arrangements in the Lepidoptera. Two species of Hepialoidea, which is one of the most ancient lepidopteran lineages, display the most common insect type, presenting IQM instead of the previously known Lepidoptera-specific rearrangement. This finding consequently reduced the extension of the Lepidoptera-specific rearrangement to the Ditrysia, which includes approximately 98 % of all described Lepidoptera (van Nieukerken et al. 2011). Since that time, new arrangements have been reported from individual species of Lepidoptera, although they are not abundant (e.g., Wang et al. 2014). Nevertheless, lepidopteran arrangements have never previously been scrutinized, although more than 270 mitogenome sequences comprising 44 families in 23 superfamilies (as of August 6, 2015) are GenBank-registered (or published).

The Gelechioidea are distributed worldwide, comprising 18,489 species in 1428 genera, and are the second most species-rich group of Lepidoptera (van Nieukerken et al. 2011). Despite the phylogenetic significance of this mega-diverse superfamily for the understanding of the higher phylogeny of Ditrysia (Kaila et al. 2011), prior to this study, only seven mitogenomic sequences, representing 5 of the 19 families, were available, including only a single species of Gelechiidae in the subfamily Pexicopiinae (Park et al. 2014; Timmermans et al. 2014; Zhao et al. 2014). In fact, due in part to the paucity of taxa included, previous mitogenome-based lepidopteran phylogeny suffered in resolving the relationships of some early-derived groups, including the Gelechioidea (Timmermans et al. 2014). Therefore, additional mitogenomic sequences from a diverse group of Gelechioidea are essential for the inference of interfamilial relationships within Gelechioidea and superfamilial relationships within the Lepidoptera in the future.

In the present study, we sequenced the entire mitogenomes of two Gelechiidae: *Mesophleps albilinella,* which belongs to the Anacampsinae, and *Dichomeris ustalella,* which belongs to the Dichomeridinae (Park 1990, 1991; Parsons 1995). *M. albilinella* is found in Korea and China, whereas *D. ustalella* is distributed extensively in southeastern Siberia, the Caucasus, Transcaucasia, Korea, Japan, China, Denmark, Belgium, France, and Italy (Park 1990, 1991; Parsons 1995; Li and Sattler 2012). The genome organization and sequence composition of the two mitogenome sequences were compared to those of available gelechioid mitogenomes. We also report that *M. albilinella* has a unique gene arrangement never previously found in Gelechioidea. The mechanism responsible for this rearrangement appears to involve inversion, which is a rare mechanism (Cameron 2014). Furthermore, the mitogenome sequences of all available lepidopterans were collected from public databases, and their gene arrangements were analyzed to determine the extent of gene rearrangement in Lepidoptera, to infer the major mechanism responsible for genome rearrangements, and to determine the evolutionary independence (or sharing) of any given rearrangement.

Materials and methods

Genomic DNA extraction, PCR, and sequencing

Adult *Mesophleps albilinella* and *Dichomeris ustalella* specimens were collected from Geojedo Island, Gyeongsangnam-do Province in Korea on September 25 and August 25, 2012, respectively. Total DNA was extracted from two hind legs using a Wizard™ Genomic DNA Purification Kit according to the manufacturer's instructions (Promega, USA). The complete mitogenomes were amplified into three overlapping long fragments (LF1–LF3), using genomic DNA as a template, and 26 subsequent overlapping short fragments (SF1–SF 26), using the LFs as templates. The primers for both the LFs and SFs were adapted from Kim et al. (2012), and detailed sequences are presented in Table 1.

LF PCR was performed using LA Taq™ (Takara Biomedical, Japan) under the following conditions: initial denaturation for 2 min at 96 °C, followed by 30 cycles of 10 s at 98 °C and 15 min at 50 °C, and a subsequent 10-min final extension at 72 °C. For SF PCR, AccuPower PreMix (Bioneer, Korea) was used under the following conditions: initial denaturation for 5 min at 94 °C, followed by 35 cycles of 1 min at 94 °C, 1 min at 48–51 °C, and 1 min at 72 °C, with a subsequent final 7-min extension at 72 °C. SF1–SF25 were directly sequenced, whereas SF26, which encompasses the whole A + T-rich region, was sequenced after cloning. Cloning was carried out using a pGEM-T Easy vector (Promega, USA) and HIT-competent cells (Real Biotech Corporation, Taiwan). The resultant plasmid DNA was isolated using a Plasmid Mini Extraction Kit (Bioneer, Korea). DNA sequencing was conducted using an ABI PRISM® BigDye® Terminator v3.1 Cycle Sequencing Kit with an ABI 3100 Genetic Analyzer (PE Applied Biosystems, USA). All products were sequenced from both directions.

Table 1 List of primers used to amplify and sequence two mitochondrial genomes of Gelechiidae

Fragment name	Primer name	Gene	Direction[a]	Sequence (5–3′)	Nucleotide position[b]
Long fragments					
LF1	Lep-COI-F1	COI	F	TTCTACAAATCATAAAGATATTGG	1498–1521
LF1	Lep-ND4-R1	ND4	R	ATTGGTCATGGTTTATGTTCTTC	8592–8614
LF2	Lep-ND5-F1	ND5	F	CTAAAAGGAATTTGAGCTCT	7511–7530
LF2	Lep-lrRNA-R1	lrRNA	R	CTGTACAAAGGTAGCATAATCATT	13281–13304
LF3	Lep-lrRNA-F1	lrRNA	F	TGTAAGATTTTAATGATCGAACAGAT	12860–12885
LF3	Lep-COI-R1	COI	R	CTTCAGGATGACCAAAAAATC	2181–2201
Short fragments					
SF1	LF03-S05-F1	$tRNA^{Met}$	F	AAGCTTTTGGGYTCATACC	20–38
SF1	LF03-S05-R2	ND2	R	CAWCCTAAATTATTAATWGAWGA	803–825
SF2	LF03-S06-F1	ND2	F	TCWTCHWTATTAATAAAAATAGG	566–588
SF2	LF03-S06-R2	$tRNA^{Tyr}$	R	GCGATAAATTGTAAATTTAT	1438–1457
SF3	LF03-S07-F1	$tRNA^{Trp}$	F	AATCTTCAAAATTATTTATAAAG	1310–1332
SF3	Lep-COI-R1	COI	R	CTTCAGGATGACCAAAAAATC	2181–2201
SF4	LF01-S01-F1	COI	F	GGTATTTCATCAATTTTAGG	1934–1953
SF4	LF01-S01-R2	COI	R	GTCGAGGTATTCCTGCTA	2772–2789
SF5	LF01-S02-F2	COI	F	ACWGTAGGAGGATTAACAGG	2519–2538
SF5	LF01-S02-R2	COII	R	GTTCAAATTAATTCAATTATTTG	3256–3278
SF6	LF01-S03-F2	COII	F	TAGAAATGGCAACWTGATC	3077–3095
SF6	LF01-S03-R1	$tRNA^{Lys}$	R	CTTGCTTTCAGTCATCTAAT	3759–3778
SF7	LF01-S04-F1	COII	F	CAGGTCGWTTAAATCAAAC	3602–3620
SF7	LF01-S04-R2	ATP6	R	GTTCCTTGDGGAATTATATG	4444–4463
SF8	LF01-S05-F1	ATP6	F	TTATTTTCAATTTTTGATCC	4072–4091
SF8	LF01-S05-R1	COIII	R	CTCGTCATCATTGATATAT	4896–4914
SF9	LF01-S06-F2	COIII	F	GTWGATTATAGHCCWTGACC	4767–4786
SF9	LF01-S06-R2	$tRNA^{Gly}$	R	GATTGGAAGTCAAATATACT	5542–5561
SF10	LF01-S07-F1	COIII	F	AGCATATGAATATWTWGAAGC	5267–5287
SF10	LF01-S07-R1	$tRNA^{Asn}$	R	CAATTTTATCATTAACAGTGA	6112–6132
SF11	LF01-S08-F1	ND3	F	TAGAAATTGCATTAATTTTHCC	5798–5819
SF11	LF01-S08-R2	ND5	R	CCTTATATAATTTATTTACC	6758–6777
SF12	LF01-S09-F1	ND5	F	AWAHTTCTCTTCAACCYAWATC	6522–6543
SF12	LF01-S09-R2	ND5	R	GCTTTATCWACTTTAAGWCA	7277–7296
SF13	LF01-S10-F1	ND5	F	TCYTTWGAATAAAAYCCAG	7025–7043
SF13	LF01-S10-R1	ND5	R	GATGGDTTAGGDTTAGTTTCTT	7746–7767
SF14	LF01-S11-F1	ND5	F	AAAAAATATAATTTCAWCTHCC	7605–7626
SF14	LF01-S11-R1	ND4	R	CATTGATTWCCTTTAAATAT	8226–8245
SF15	LF01-S12-F1	$tRNA^{His}$	F	ATATTTTTGAYHCCACAAATC	8141–8161
SF15	LF01-S12-R2	ND4	R	CAGGTTCAATAATTTTAGC	8868–8886
SF16	LF02-S01-F1	ND4	F	TTATAATACCHCCAATWAC	8656–8674
SF16	LF02-S01-R1	ND4	R	GGTTTAATTTTATTAAGAATTTG	9327–9349
SF17	LF02-S02-F1	ND4	F	ATATTAAAGTAGGAATTAAWC	9151–9171
SF17	LF02-S02-R2	$tRNA^{Pro}$	R	TAATTTTGGAGATTATWGAT	9904–9923
SF18	LF02-S03-F1	ND4L	F	CCTAAAGCHCCYTCACAAAC	9597–9616
SF18	LF02-S03-R1	ND6	R	GTAATTTTTACWACTGCAATTA	10425–10446
SF19	LF02-S04-F2	ND6	F	TNTCAAGAATTGCHTCWAATG	10155–10175
SF19	LF02-S04-R1	Cytb	R	GATATTTGTCCYCAAGGTA	10902–10920
SF20	LF02-S05-F2	Cytb	F	TATHTHCATATTGGACGAGG	10787–10806
SF20	LF02-S05-R1	Cytb	R	CCTTGDATTTTTTTATTAAADGT	11423–11445

Table 1 continued

Fragment name	Primer name	Gene	Direction[a]	Sequence (5–3′)	Nucleotide position[b]
SF21	LF02-S06-F2	Cytb	F	ACHCCHRTTCATATTCAACC	11294–11313
SF21	LF02-S06-R2	ND1	R	GCTGAAACTAATCGAACTC	12078–12096
SF22	LF02-S07-F1	ND1	F	AACGAGGTAAWGTHCCHCG	11848–11868
SF22	LF02-S07-R2	lrRNA	R	CTGAGTTCAAACCGGTGTRA	12897–12916
SF23	LF02-S08-F2	$tRNA^{Leu}$	F	GAHTTCTAAAAYCATTAC	12655–12672
SF23	LF02-S08-R1	lrRNA	R	GACTGTACAAAGGTAGCATAAT	13345–13366
SF24	LF03-S02-F1	lrRNA	F	ATTATGCTACCTTTGTACAGTC	13345–13366
SF24	LF03-S02-R1	$tRNA^{Val}$	R	GTATTTCATTTACATTGAAAAGA	14152–14174
SF25	LF03-S03-F3	lrRNA	F	CTCTGATACACAAGATAC	13936–13953
SF25	LF03-S03-R3	srRNA	R	CCAGCAGTTGCGGTTAAAC	14785–14803
SF26	LF03-S04-F1	srRNA	F	AATAGGTGATCTAATCCTAG	14631–14650
SF26	LF03-S04-R1	$tRNA^{Ile}$	R	CTATCAGAATAATCCTTTWA	82–101

[a] F and R, forward and reverse direction of trsnacription

[b] Nucleotide positions are with respective to *Mesophleps albilinella* mitochondrial genome

Genome annotation

Gene identification, boundary delimitation, and secondary structure predictions for *M. albilinella* and *D. ustalella* tRNAs were made using tRNAscan-SE 1.21 with the search mode set as default, Mito/Chloroplast as the search source, invertebrate mitogenomes as the genetic code for tRNA isotype prediction, and a Cove score cutoff of 1 (Lowe and Eddy 1997). By this method, 21 tRNAs were found in both species. $tRNA^{Ser}$(AGN), which has a truncated DHU arm, was found in a hand-drawn secondary structure by the alignment of predicted regions of other lepidopteran mitochondrial (mt) $tRNA^{Ser}$(AGN), with particular consideration given to the anticodons. MAFFT ver. 6 (Katoh et al. 2002) was used for this process, with the gap opening penalty set to 1.53 and the offset value (≈gap extension penalty) set to 0.5. The individual mt protein-coding gene (PCG) was identified and its boundary delimitated using the blastn program in BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The start and stop codons of the PCGs were further confirmed by the alignment of *M. albilinella* and *D. ustalella* PCGs with other lepidopteran mt PCGs, including those of other gelechioids. Two rRNAs and the A + T-rich region were identified and delimitated using the nucleotide blast program in BLAST and further confirmed by alignment with other lepidopteran mt rRNA genes and A + T-rich region sequences. The sequence data were deposited into the GenBank database under the accession numbers KU366707 for *M. albilinella* and KU366706 for *D. ustalella*.

Comparative mitochondrial gene analyses

For the genomic comparison, seven gelechioid mitogenomes were downloaded from either GenBank or AMiGA (Feijao et al. 2006) (Table 2). The nucleotide composition of each gene, the whole genome, and the codon positions of the PCGs were calculated using MEGA 6 (Tamura et al. 2013). Translation of nucleotide sequences and calculation of the codon frequency of the PCGs were performed by MEGA 6 based on the invertebrate mt DNA genetic code (Tamura et al. 2013). Gene overlap and intergenic-space sequences were hand-counted. The taxonomic scope of the comparison was limited to the available gelechioids, since the genomic features of this group have never been extensively analyzed.

Results

General perspectives on the *M. albilinella* and *D. ustalella* mitogenomes

The mitogenomes of *M. albilinella* and *D. ustalella* were 15,274 and 15,410 bp in size, respectively, and contained typical sets of mt genes (13 PCGs, 22 tRNAs, and 2 rRNAs) and one major non-coding region, known as the A + T-rich region in insects (Table 2). The extra tRNAs that have been infrequently detected in other Lepidoptera were not found in these species (e.g., *Coreana raphaelis* and *Ctenoptilum vasava* in Papilionoidea; Kim et al. 2006; Hao et al. 2012). Twelve PCGs of *M. albilinella* and *D. ustalella* started with the typical start codons ATN, but COI began with CGA (arginine) in both *M. albilinella* and *D. ustalella*, and all sequenced gelechioids also started with CGA (Fig. 1). The mt PCGs of both *M. albilinella* and *D. ustalella* ended with TAA in nine genes, but ended with an incomplete stop codon consisting of a single thymine

Table 2 Genomic summary of Gelechioidea

Gene	Anticodon	Start codon	Stop codon	Mesophleps albilinella	Dichomeris ustalella	*Pectinophora gossypiella*	*Atrijuglans hetaohei*	*Ethmia eupostica*
$tRNA^{Met}$	CAT			1–68 (68)	1–68 (68)	1–69 (69)	1–67 (67)	1–69 (69)
$tRNA^{Ile}$	GAT			69–135 (67)	73–139 (67)	70–135 (66)	68–132 (65)	70–135 (66)
$tRNA^{Gln}$	TTG			133–201 (69)	159–227 (69)	133–201 (69)	130–198 (69)	133–201 (69)
ND2		$ATA^a/ATT^b/ATC^c$	TAA	$256–1269^b$ (1014)	$279–1292^b$ (1014)	$257–1270^b$ (1014)	$263–1267^a$ (1005)	$264–1265^b$ (1002)
$tRNA^{Trp}$	TCA			1271–1337 (67)	1291–1359 (69)	1269–1335 (67)	1266–1334 (69)	1264–1331 (68)
$tRNA^{Cys}$	GCA			1330–1395 (66)	1352–1419 (68)	1328–1392 (65)	1327–1391 (65)	1324–1392 (69)
$tRNA^{Tyr}$	GTA			1405–1470 (66)	1446–1510 (65)	1396–1463 (68)	1395–1461 (67)	1402–1468 (67)
COI		CGA	T	1473–3003 (1531)	1514–3044 (1531)	1466–2996 (1531)	1463–2993 (1531)	1471–2998 (1528)
$tRNA^{Leu}$ (UUR)	TAA			3004–3070 (67)	3045–3112 (68)	2997–3063 (67)	2994–3061 (68)	2999–3065 (67)
COII		ATG	T	3071–3755 (685)	3113–3794 (682)	3064–3745 (682)	3062–3743 (682)	3066–3747 (682)
$tRNA^{Lys}$	CTT			3756–3826 (71)	3795–3865 (71)	3746–3816 (71)	3744–3814 (71)	3748–3818 (71)
$tRNA^{Asp}$	GTC			3829–3896 (68)	3865–3933 (69)	3821–3888 (68)	3820–3885 (66)	3875–3942 (68)
ATP8		$ATA^a/ATT^b/ATC^C$	TAA	$3897–4061^b$ (165)	$3934–4098^b$ (165)	$3889–4050^a$ (162)	$3886–4044^b$ (159)	$3943–4110^c$ (168)
ATP6		ATG	TAA	4055–4732 (678)	4092–4769 (678)	4044–4721 (678)	4038–4715 (678)	4104–4781 (678)
COIII		ATA^a/ATG^b	T^1/TAA^2	$4734–5526^{a/1}$ (793)	$4776–5564^{b/2}$ (789)	$4721–5509^{b/2}$ (789)	$4715–5503^{b/2}$ (789)	$4802–5590^{b/2}$ (789)
$tRNA^{Gly}$	TCC			5527–5592 (66)	5580–5645 (66)	5513–5579 (67)	5506–5571 (66)	5594–5660 (67)
ND3		$ATA^a/ATT^b/ATC^c$	TAA^1/TAG^2	$5590–5946^{a/1}$ (357)	$5643–5999^{a/1}$ (357)	$5580–5933^{b/1}$ (354)	$5572–5925^{b/1}$ (354)	$5664–6014^{b/1}$ (351)
$tRNA^{Ala}$	TGC			5949–6014 (66)	6002–6067 (66)	5940–6004 (65)	5935–6002 (68)	6022–6087 (66)
$tRNA^{Arg}$	TCG			6014–6077 (64)	6088–6157 (70)	6004–6069 (66)	6002–6067 (66)	6124–6191 (68)
$tRNA^{Asn}$	GTT			6077–6142 (66)	6159–6226 (68)	6071–6136 (66)	6067–6132 (66)	6201–6266 (66)
$tRNA^{Ser}$ (AGN)	GCT			6218–6283 (66)	6226–6291 (66)	6136–6202 (67)	6132–6194 (63)	6291–6356 (66)
$tRNA^{Glu}$	TTC			6150–6217 (68)	6293–6357 (65)	6205–6273 (69)	6259–6325 (67)	6357–6422 (66)
$tRNA^{Phe}$	GAA			6311–6376 (66)	6356–6423 (68)	6274–6339 (66)	6373–6440 (68)	6428–6494 (67)
ND5		ATT	$T^1/TA^2/TAA^3$	$6377–8108^1$ (1732)	$6424–8164^1$ (1741)	$6340–8071^1$ (1732)	$6441–8180^3$ (1740)	$6494–8229^2$ (1736)
$tRNA^{His}$	GTG			8109–8175 (67)	8165–8232 (68)	8072–8138 (67)	8181–8245 (65)	8230–8296 (67)
ND4		ATG	$T^1/TA^2/TAA^3$	$8180–9523^3$ (1344)	$8233–9571^1$ (1339)	$8139–9477^1$ (1339)	$8246–9584^1$ (1339)	$8296–9635^2$ (1340)
ND4L		ATG	TAA	9526–9816 (291)	9574–9867 (294)	9478–9768 (291)	9585–9857 (273)	9638–9928 (291)
$tRNA^{Thr}$	TGT			9819–9882 (64)	9874–9938 (65)	9771–9836 (66)	9862–9926 (65)	9931–9994 (64)
$tRNA^{Pro}$	TGG			9883–9948 (66)	9939–10,003 (65)	9837–9903 (67)	9927–9991 (65)	9995–10,060 (66)
ND6		ATA^a/ATT^b	T^1/TAA^2	$9951–10,484^{b/2}$ (534)	$10,006–10,536^{a/2}$ (531)	$9906–10,436^{a/2}$ (531)	$10,003–10,527^{a/2}$ (525)	$10,066–10,593^{a/2}$ (528)
CytB		$ATA^a/ATG^b/ATC^c$	TAA	$10,493–11,641^b$ (1149)	$10,541–116,89^b$ (1149)	$10,452–11,600^b$ (1149)	$10,540–11,691^b$ (1152)	$10,596–11,750^a$ (1155)
$tRNA^{Ser}$ (UCN)	TGA			11,652–11,718 (67)	11,692–11,756 (65)	11,599–11,666 (68)	11,690–11,757 (68)	11,750–11,816 (67)
ND1		ATA^a/ATG^b	TAA	$11,735–12,670^a$ (936)	$11,778–12,716^b$ (939)	$11,684–12,622^b$ (939)	$11,798–12,733^b$ (936)	$11,836–12,771^b$ (936)
$tRNA^{Leu}$ (CUN)	TAG			12,674–12,745 (72)	12,717–12,784 (68)	12,624–12,696 (73)	12,735–12,805 (71)	12,776–12,844 (69)

Table 2 continued

Gene	Anticodon	Start codon	Stop codon	Mesophleps albilinella	Dichomeris ustalella	*Pectinophora gossypiella*	*Atrijuglans hetaohei*	*Ethmia eupostica*
lrRNA				12,746–14,077 (1332)	12,785–14,225 (1441)	12,697–14,052 (1356)	12,806–14,172 (1367)	12,845–14,188 (1344)
$tRNA^{Val}$	TAC			14,078–14,144 (67)	14,226–14,294 (69)	14,053–14,116 (64)	14,173–14,242 (70)	14,189–14,253 (65)
srRNA				14,145–14,921 (777)	14,295–15,089 (795)	14,117–14,893 (777)	14,243–15,016 (774)	14,254–15,036 (783)
A + T–rich region				14,922–15,274 (353)	15,090–15,410 (321)	14,894–15,202 (309)	15,017–15,379 (363)	15,037–15,347 (311)

Gene	Anticodon	Start codon	Stop codon	*Perimede* sp.	*Endrosis sarcitrella*	*Promalactis suzukiella*	*Oegoconia novimundi*
$tRNA^{Met}$	CAT			1–67 (67)	1–68 (68)	1–68 (68)	1–68 (68)
$tRNA^{Ile}$	GAT			69–134 (66)	69–134 (66)	74–137 (64)	69–132 (64)
$tRNA^{Gln}$	TTG			132–200 (69)	132–200 (69)	167–235[b] (69)	134–202[a] (69)
ND2		ATA[a]/ATT[b]/ATC[c]	TAA	229–1269[a] (1041)	247–1275[c] (1029)	291–1307 (1017)	237–1259 (1023)
$tRNA^{Trp}$	TCA			1268–1334 (67)	1274–1341 (68)	1306–1373 (68)	1259–1326 (68)
$tRNA^{Cys}$	GCA			1335–1405 (71)	1334–1397 (64)	1366–1430 (65)	1319–1386 (68)
$tRNA^{Tyr}$	GTA			1418–1484 (67)	1419–1486 (68)	1432–1496 (65)	1386–1449 (64)
COI		CGA	T	1493–3023 (1531)	1490–3020 (1531)	1502–3032 (1531)	1457–2987 (1531)
$tRNA^{Leu}$ (UUR)	TAA			3024–3090 (67)	3021–3087 (67)	3033–3100 (68)	2988–3053 (66)
COII		ATG	T	3091–3772 (682)	3088–3769 (682)	3102–3782 (681)	3055–3736 (682)
$tRNA^{Lys}$	CTT			3773–3842 (70)	3770–3840 (71)	3789–3869 (81)	3737–3807 (71)
$tRNA^{Asp}$	GTC			3842–3908 (67)	3850–3915 (66)	3883–3950 (68)	3821–3887 (67)
ATP8		ATA[a]/ATT[b]/ATC[C]	TAA	3909–4070[c] (162)	3916–4080[b] (165)	3951–4112[c] (162)	3888–4061[c] (174)
ATP6		ATG	TAA	4064–4747 (684)	4074–4751 (678)	4106–4783 (678)	4055–4732 (678)
COIII		ATA[a]/ATG[b]	T^1/TAA^2	4747–5535[b/2] (789)	4756–5544[b/2] (789)	4783–5571[b/2] (789)	4738–5526[b/2] (789)
$tRNA^{Gly}$	TCC			5538–5603 (66)	5547–5614 (68)	5574–5641 (68)	5529–5597 (69)
ND3		ATA[a]/ATT[b]/ATC[c]	TAA^1/TAG^2	5607–5957[b/2] (351)	5618–5968[b/1] (351)	5642–5995[b/1] (354)	5601–5951[c/1] (351)
$tRNA^{Ala}$	TGC			5956–6019 (64)	5989–6052 (64)	6059–6124 (66)	5983–6048 (66)
$tRNA^{Arg}$	TCG			6019–6080 (62)	6052–6115 (64)	6127–6194 (68)	6101–6162 (62)
$tRNA^{Asn}$	GTT			6083–6147 (65)	6116–6181 (66)	6195–6261 (67)	6188–6253 (66)
$tRNA^{Ser}$ (AGN)	GCT			6148–6208 (61)	6216–6281 (66)	6271–6340 (70)	6268–6333 (66)
$tRNA^{Glu}$	TTC			6210–6276 (67)	6283–6349 (67)	6339–6403 (65)	6389–6452 (64)
$tRNA^{Phe}$	GAA			6295–6361 (67)	6374–6440 (67)	6406–6472 (67)	6466–6532 (67)
ND5		ATT	T^1/TA^2/TAA^3	6361–8078[2] (1718)	6451–8169[3] (1719)	6473–8210[1] (1738)	6532–8249[2] (1718)
$tRNA^{His}$	GTG			8097–8162 (66)	8185–8249 (65)	8211–8278 (68)	8268–8335 (68)
ND4		ATG	T^1/TA^2/TAA^3	8162–9498[2] (1337)	8249–9588[2] (1340)	8279–9617[1] (1339)	8335–9674[2] (1340)
ND4L		ATG	TAA	9499–9789 (291)	9590–9877 (288)	9630–9917 (288)	9675–9965 (291)
$tRNA^{Thr}$	TGT			9797–9860 (64)	9883–9946 (64)	9923–9987 (65)	9968–10033 (66)
$tRNA^{Pro}$	TGG			9861–9924 (64)	9947–10,011 (65)	9988–10,052 (65)	10,034–10,099 (66)

Table 2 continued

Gene	Anticodon	Start codon	Stop codon	*Perimede* sp.	*Endrosis sarcitrella*	*Promalactis suzukiella*	*Oegoconia novimundi*
ND6		ATA[a]/ATT[b]	T[1]/TAA[2]	9948–10,449[a/1] (502)	10,014–10,544[b/2] (531)	10,055–10,585[b/2] (531)	10,115–10,636[c/2] (522)
CytB		ATA[a]/ATG[b]/ATC[c]	TAA	10,457–11,608[a] (1152)	10,551–11,699[b] (1149)	10,635–11,783[b] (1149)	10,642–11,805[b] (1164)
$tRNA^{Ser}$ (UCN)	TGA			11,607–11,670 (64)	11,715–11,781 (67)	11,829–11,896 (68)	11,811–11,875 (65)
ND1		ATA[a]/ATG[b]	TAA	11,691–12,623[b] (933)	11799–12,737[b] (939)	11,914–12,858[b] (945)	11,897–12,835[a] (939)
$tRNA^{Leu}$ (CUN)	TAG			12,625–12,691 (67)	12,739–12,809 (71)	12,860–12,926 (67)	12,836–12,904 (69)
lrRNA				12,692–13,990 (1299)	12,808–14,178 (1371)	12,927–14,293 (1367)	12,905–14,254 (1350)
$tRNA^{Val}$	TAC			13,991–14,054 (64)	14,179–14,245 (67)	14,294–14,363 (70)	14,255–14,319 (65)
srRNA				14,055–14,831 (777)	14,246–15,027 (782)	14,364–15,138 (775)	14,320–15,093 (774)
A + T-rich region				14,832–15,131 (300)	15,028–15,317 (290)	15,139–15,507 (369)	15,094–15,408 (315)

Superscripts indicate identical start codon and stop codon among gelechioid species, respectively. Values in parentheses indicate gene size (bp)

in four genes (Table 2). Such an incomplete termination codon can become a complete stop codon (TAA) by post-translational modifications that occur during the mRNA maturation process (Ojala et al. 1981).

Most mt genes of *M. albilinella* and *D. ustalella* were well within the size range found in other gelechioids, but a few were slightly larger (Table 2). For example, the size of COIII in *M. albilinella* was 793 bp, but ranged from 681 to 789 bp in other gelechioids. In addition, the size of $tRNA^{Arg}$ in *D. ustalella* was 70 bp, but ranged from 62 to 68 bp in other gelechioids.

Nucleotide composition

The nucleotide composition of the *M. albilinella* and *D. ustalella* mitogenomes was also biased toward A/T nucleotides, at 80.5 and 81.1 %, respectively, similar to other gelechioids, where it ranged from 77.6 % (*Oegoconia novimundi*) to 81.5 % (*Promalactis suzukiella*) (Table 3). The A/T content varied profoundly between RNAs (85.4 % in srRNA, 84.1 % in lrRNA, 82.3 % in tRNAs in *M. albilinella*; and 86.5 % in srRNA, 86.2 % in lrRNA, and 82.2 % in tRNAs in *D. ustalella*) and PCGs (79.0 % in *M. albilinella* and 79.3 % in *D. ustalella*), and this trend was always observed in the sequenced gelechioids (Table 3). The biased usage of A/T nucleotides was also reflected in the form of codon usage (Table 4). Among the 64 available codons, *M. albilinella* and *D. ustalella* utilized TTA (Leucine), ATT (Isoleucine), TTT (Phenylalanine), and ATA (Methionine) most frequently, accounting for 40.70 and 40.46 %, respectively. On the other hand, *Oegoconia novimundi,* belonging to the Autostichidae, had the lowest frequency of the four codons at 35.99 % (Table 4), and this species had the lowest A/T content in the whole genome (77.6 %) as well (Table 3). Currently, *O. novimundi* is the only species available from this family. Thus, it would be interesting to have more data on this family.

The nucleotide composition of 13 concatenated PCGs in the *M. albilinella* mitogenome was as follows: A, 33.0 %; T, 46.0 %; C, 10.3 %; and G, 10.7 %. In the *D. ustalella* mitogenome the composition was A, 33.4 %; T, 45.9 %; C, 10.1 %; and G, 10.7 % (Table 5). The base composition at each codon position of the PCGs in *M. albilinella* and *D. ustalella* showed that the third codon position (93.2 and 94.5 %, respectively) harbored a higher A/T content than the first (73.6 and 73.1 %) and second (70.2 and 70.1 %) codon positions, revealing slightly higher content in the first codon position. A similar pattern was also detected in other sequenced gelechioids, with an average of 73.04 % in the first position, 69.88 % in the second position, and 92.69 % in the third position (Table 5).

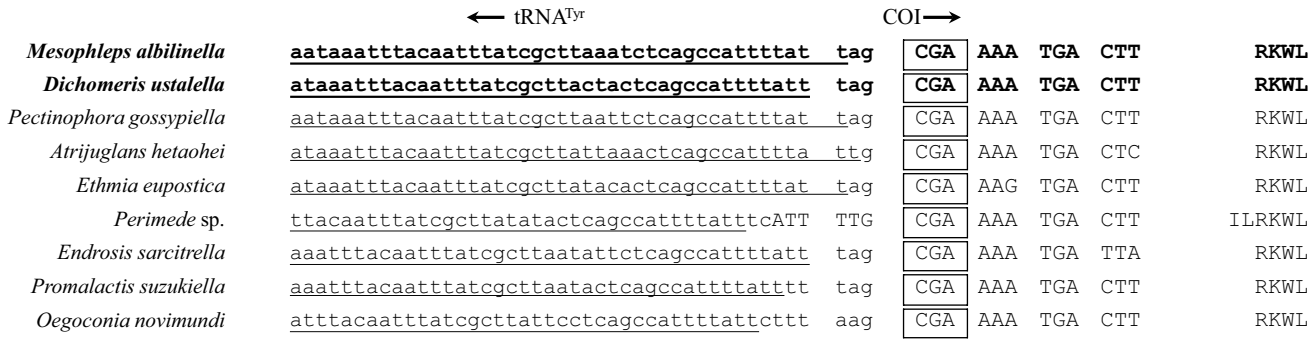

Fig. 1 Alignment of the initiation context of the COI genes of Gelechioidea, including those of *Mesophleps albilinella* and *Dichomeris ustalella*. The amino acid sequences of the first four to six codons are shown on the *right-hand* side of the figure. Underlined nucleotides indicate the adjacent partial sequence of $tRNA^{Tyr}$. *Arrows* indicate the transcriptional direction. *Boxed* nucleotides indicate currently proposed translation initiators

Non-coding spacer sequences

The *M. albilinella* and *D. ustalella* mt genes are interleaved with a total of 151 and 180 bp intergenic spacer sequences, spread over 18 and 16 regions ranging in size from 1 to 54 bp and 1 to 51 bp, respectively (Sup. Table 1). Two relatively longer spacer sequences are noteworthy enough to mention here. One is found between $tRNA^{Gln}$ and ND2 at 54 bp in *M. albilinella* and at 51 bp in *D. ustalella*. Another is found between $tRNA^{Ser}$(UCN) and ND1 at 16 bp in *M. albilinella* and at 21 bp in *D. ustalella* (termed Spacers 1 and 2, respectively). Spacer 1 is consistently found in all other gelechioids, at a size ranging from 28 to 66 bp (Sup. Table 1). Spacer 2 is composed of 81.3 and 90.5 % A/T nucleotides in *M. albilinella* and *D. ustalella*, respectively, and also is consistently found in all other gelechioids, at sizes ranging from 17 to 40 bp (Sup. Table 1). With the exception of these two longer spacer sequences, *M. albilinella* has relatively longer sequences at 26 bp. *D. ustalella* has three sequences longer than 10 bp, but no peculiar aspects were found except that some of these are composed mainly of TA repeats (data not shown). The two gelechioid species had overlapping sequences ranging from 1 to 8 bp spread over 8 (*M. albilinella*) and 10 (*D. ustalella*) locations, for a total of 29 bp (*M. albilinella*) and 27 bp (*D. ustalella*), respectively (Sup. Table 1).

The A + T-rich region

The lengths of the A + T-rich regions in *M. albilinella* and *D. ustalella* were 353 and 321 bp, respectively (Table 2), and were composed of 94.6 and 94.4 % of A/T nucleotides (data not shown). The 369-bp *P. suzukiella* was the longest and *M. albilinella* was the third-longest among the gelechioids. The shortest region was found in *Pectinophora gossypiella,* at 300 bp (Timmermans et al. 2014).

The A + T-rich regions of all gelechioids commonly possess the motif ATAGA close to a 5′-end of the srRNA, with a varying length of poly-T stretch (16 to 19 bp) (Fig. 2a). This motif and the poly-T stretch have been suggested as the site of replication origin of minority strands of mtDNA in the lepidopteran *Bombyx mori* (Saito et al. 2005). Along with the motif, the A + T-rich regions of most gelechioids, including *M. albilinella* and *D. ustalella,* commonly possess the ATTTA sequence, the function of which is unknown; a variable length of TA-repeat (excluding *M. albilinella*); and a complete or interrupted poly-A stretch immediately upstream of the $tRNA^{Met}$ (excluding *P. suzukiella*) (Fig. 2a). Two tRNA-like structures in *M. albilinella* ($tRNA^{Met}$-like and $tRNA^{Asn}$-like structures) and one in *D. ustalella* ($tRNA^{Leu}$-like structure) were found in the A + T-rich region (Fig. 2b).

RNAs

Two rRNA genes (srRNA and lrRNA) were identified at 792 and 1347 bp in *M. albilinella* and at 795 and 1441 bp in *D. ustalella* (Table 3). In other gelechioids, the srRNA ranged from 774 (*Atrijuglans hetaohei* and *O. novimundi*) to 783 bp (*E. eupostica*) and the lrRNA ranged from 1299 (*Perimede* sp.) to 1441 bp (*D. ustalella*); thus, the present *D. ustalella* lrRNA is the largest of any known gelechioid lrRNA (Table 3).

Most (21 of 22) tRNAs were folded into a cloverleaf secondary structure, except for a $tRNA^{Ser}$(AGN) that lacked the DHU stem in both *M. albilinella* and *D. ustalella* (Sup. Figures 1 and 2), as has been shown in many other metazoans (Garey and Wolstenholme 1989). The length of the 22 tRNAs ranged from 64 bp ($tRNA^{Arg}$ and $tRNA^{Thr}$) to

Table 3 Characteristics of mitochondrial genomes of Gelechioidea

Taxon	Size (bp)	A/T content (%)	PCG		srRNA		lrRNA		tRNA		A + T-rich region		GenBank accession no.	References
			No. codons[a]	AT (%)	Size (bp)	AT (%)	Size (bp)	AT (%)	Size (bp)	AT (%)	Size (bp)	AT (%)		
Gelechioidea														
Gelechiidae														
Anacampsinae														
Mesophleps albilinella	**15,274**	**80.5**	**3687**	**79.0**	**792**	**85.4**	**1347**	**84.1**	**1476**	**82.3**	**353**	**94.6**	**KU366707**	**This is study**
Dichomeridinae														
Dichomeris ustalella	**15,410**	**81.1**	**3727**	**79.3**	**795**	**86.5**	**1441**	**86.2**	**1485**	**82.2**	**321**	**94.4**	**KU366706**	**This is study**
Pexicopiinae														
Pectinophora gossypiella	15,202	80.7	3720	79.3	777	84.4	1356	84.4	1481	81.6	309	94.8	KM225795	Zhao et al. (2014)
Stathmopodidae														
Atrijuglans hetaohei	15,379	81.3	**3710**	79.7	774	85.3	1367	85.2	1475	81.6	363	95.6	KT581634	Unpublished
Dpressariidae														
Ethmiinae														
Ethmia eupostica	15,347	79.6	3717[b]	77.6	783[c]	85.7	1344[c]	83.3	1478	81.2	311[c]	94.9	KJ508047	Timmermans et al. (2014)
Cosmopterigidae														
Chrysopeleiinae														
Perimede sp.	15,131	80.4	3714[b]	79.0	777[c]	85.3	1299[c]	84.1	1452	81.5	300[c]	94.7	KJ508041	Timmermans et al. (2014)
Oecophoridae														
Oecophorinae														
Endrosis sarcitrella	15,317	79.6	3719[b]	77.8	782[c]	85.4	1369[c]	83.6	1468	81.2	290[c]	93.1	KJ508037	Timmermans et al. (2014)
Promalactis suzukiella	15,507	81.5	3723	79.7	775	85.8	1367	85.4	1491	81.7	369	97.0	KM875542	Park et al. (2014)
Autostichidae														
Symmocinae														
Oegoconia novimundi	15,408	77.6	3723[b]	75.5	774[c]	83.7	1346[c]	81.6	1464	80.6	315[c]	93.3	KJ508036	Timmermans et al. (2014)

The two species sequenced in this study were bold-faced

[a] Termination codons were excluded in total codon count

[b] COI sequences include 23–24 bp of non-decided nucleotides

[c] Corresponding genes and region were reannotated in this study

Table 4 Content of four most frequently used codons in mitochondrial genomes of Gelechioidea

Species	Codon				Total
	TTA (L)	ATT (I)	TTT (F)	ATA (M)	
Gelechioidea					
Gelechiidae					
Mesophleps albilinella	**12.42**	**11.62**	**9.87**	**6.79**	**40.70**
Dichomeris ustalella	**11.73**	**11.83**	**10.06**	**6.84**	**40.46**
Pectinophora gossypiella	12.80	11.53	9.14	6.72	40.19
Stathmopodidae					
Atrijuglans hetaohei	12.48	11.86	9.70	7.28	41.32
Dpressariidae					
Ethmia eupostica	11.81	11.16	9.00	6.77	38.74
Cosmopterigidae					
Perimede sp.	12.12	11.71	9.93	7.04	40.80
Oecophoridae					
Endrosis sarcitrella	12.21	10.97	9.41	6.23	38.82
Promalactis suzukiella	12.81	11.85	9.72	6.88	41.26
Autostichidae					
Oegoconia novimundi	9.85	10.77	8.86	6.51	35.99

Within parenthesis indicate corresponding amino acid: *L* leucine, *I* isoleucine, *F* phenylalanine, *M* methionine

The two species sequenced in this study were bold-faced

72 bp [$tRNA^{Leu}$(CUN)] in *M. albilinella*, and from 65 bp [($tRNA^{Tyr}$, $tRNA^{Glu}$, $tRNA^{Thr}$, $tRNA^{Pro}$, and $tRNA^{Ser}$(UCN)] to 71 bp ($tRNA^{Lys}$) in *D. ustalella*. The anticodons for each tRNA isotype were identical in all gelechioids, including *M. albilinella* and *D. ustalella* (Table 2).

Rearrangement in *M. albilinella*

The orientation and gene order of the *M. albilinella* mitogenome differed from any other gelechioid arrangement, including that of *D. ustalella* (Fig. 3). At the ND3 and ND5 junction, $tRNA^{Ser}$(AGN)/$tRNA^{Glu}$ (SE) in the $tRNA^{Ala}$/$tRNA^{Arg}$/$tRNA^{Asn}$/$tRNA^{Ser}$(AGN)/$tRNA^{Glu}$/$tRNA^{Phe}$ (ARNSEF; underline indicates an inverted gene) cluster region has been rearranged to ES, both with inversion, resulting in ARNESF.

Rearrangement in Lepidoptera

Six different mitogenome rearrangements (excluding one arrangement with duplicated tRNA) differing from the ancestral arrangement have been reported among the Lepidoptera, including that of *M. albilinella* (Fig. 4). The typical lepidopteran arrangement, which is found in the majority of the Ditrysia, differs from the ancestral one found in a variety of insect orders in the "three tRNA region" at the A + T-rich region and ND2 junction. MIQ is the typical lepidopteran arrangement, whereas IQM is the ancestral

Table 5 Codon position-based nucleotide composition of concatenated 13 PCGs of Gelechioidea

Species	Overall				1st codon position				2nd codon position				3rd codon position			
	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G
Gelechioidea																
Gelechiidae																
Mesophleps albilinella	**33.0**	**46.0**	**10.3**	**10.7**	**36.1**	**37.5**	**10.3**	**16.1**	**21.8**	**48.4**	**16.6**	**13.3**	**41.0**	**52.2**	**4.1**	**2.7**
Dichomeris ustalella	**33.4**	**45.9**	**10.1**	**10.7**	**36.5**	**36.6**	**10.4**	**16.5**	**21.8**	**48.3**	**16.4**	**13.4**	**41.8**	**52.7**	**3.4**	**2.2**
Pectinophora gossypiella	33.8	45.6	10.1	10.6	36.6	37.0	10.0	16.3	22.0	48.0	16.7	13.3	42.7	51.6	3.5	2.2
Stathmopodidae																
Atrijuglans hetaohei	33.8	45.9	9.8	10.5	37.1	37.0	9.9	16.1	21.8	48.3	16.6	13.3	42.4	52.4	3.1	2.1
Dpressariidae																
Ethmia eupostica	32.9	44.7	11.0	11.2	36.1	36.2	10.9	16.6	21.6	48.0	16.9	13.3	41.0	49.9	5.3	3.6
Cosmopterigidae																
Perimede sp.	33.2	45.8	10.1	10.7	36.4	37.4	10.0	16.0	21.5	48.6	16.3	13.4	41.8	51.4	3.9	2.8
Oecophoridae																
Endrosis sarcitrella	32.9	45.0	10.8	11.1	35.7	36.7	10.7	16.6	21.5	48.3	16.6	13.3	41.4	49.8	5.1	3.4
Promalactis suzukiella	33.9	45.8	9.7	10.6	36.4	37.1	9.9	16.7	21.6	48.2	16.9	13.4	43.6	52.1	2.4	1.9
Autostichidae																
Oegoconia novimundi	31.6	43.9	12.4	11.9	36.0	35.0	11.9	17.0	20.8	48.4	16.7	13.9	38.0	48.4	8.5	4.9

Stop codon was excluded in the count

The two species sequenced in this study were bold-faced

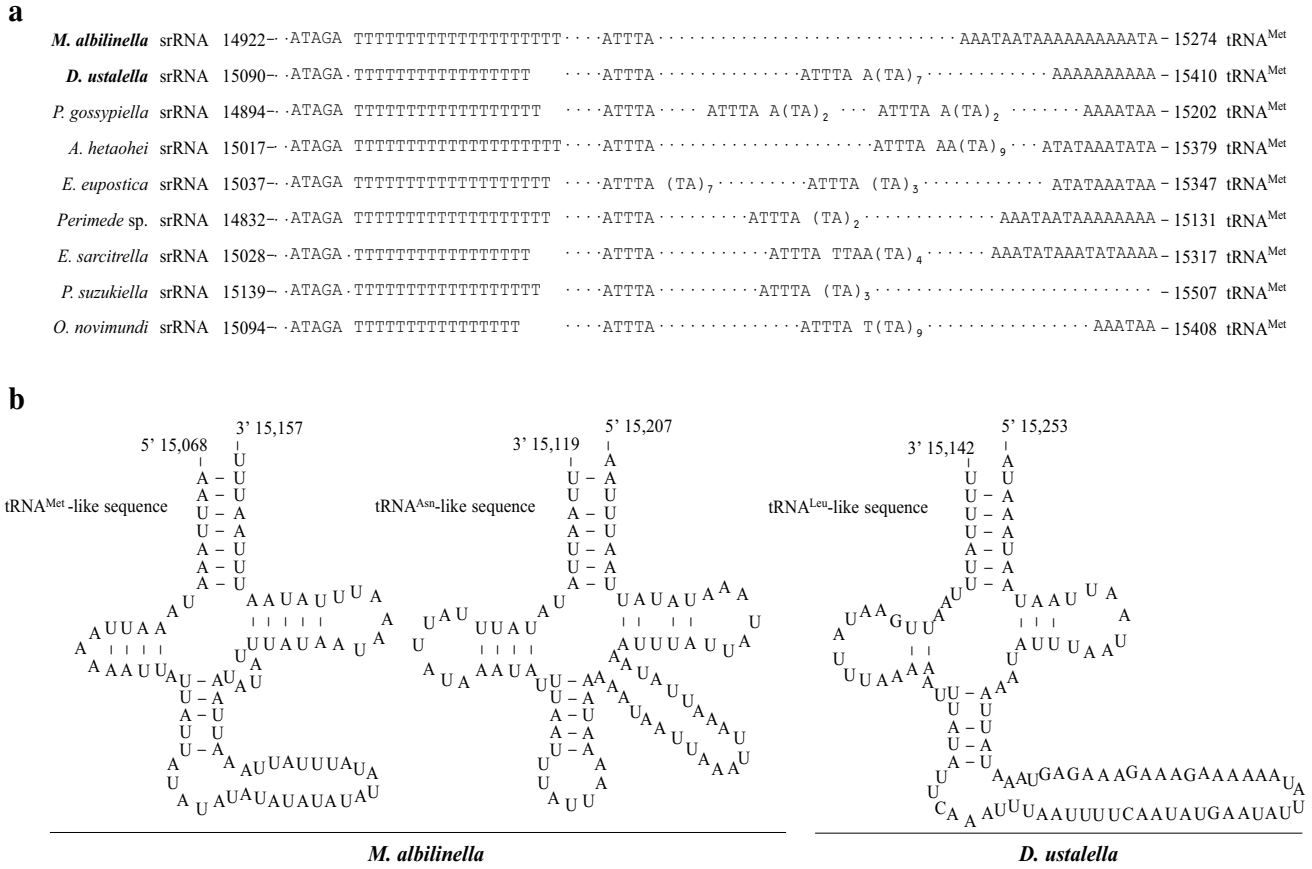

Fig. 2 Structural elements found in the A + T-rich region of Gelechioidea. **a** Schematic illustration of the A + T-rich region. The presented nucleotides indicate the conserved sequences, such as the ATAGA motif, poly-T stretch, ATTTA sequence, microsatellite-like TA repeat sequence, and poly-A stretch. *Dots* between sequences indicate omitted sequences **b** secondary structure of the tRNA-like sequence found in Gelechioidea, including *Mesophleps albilinella* and *Dichomeris ustalella*. Subscript indicates the repeat number. The nucleotide position is indicated at the beginning and end sites of the tRNA-like sequence

insect arrangement. This arrangement also is found in the Hepialoidea and Nepticuloidea, which are ancient, non-ditrysian lepidopteran groups (Cao et al. 2012; Timmermans et al. 2014). Apart from "MIQ", IMQ rearrangement in the three tRNA region is found uniquely in *Euripus nyctelius* (Nymphalidae in Papilionoidea; Xuan et al. 2015). Another rearranged region in the Lepidoptera is the "ARNSEF cluster region", which accounts for four of the six lepidopteran rearrangements, including that of *M. albilinella* (Fig. 4). The *Astrotischeria* sp., which belongs to one of the ancient, non-ditrysian lepidopteran groups (Tischeriidae in Tischerioidea) was reported to have an RNSAEF rearrangement in this cluster region (Timmermans et al. 2014). Within the Ditrysia, *Erynnis montanus* (Hesperiidae in Papilionoidea) was reported to have an SN rearrangement, resulting in an ARSNEF (Wang et al. 2014). The remaining two rearrangements, including that of *M. albilinella,* involve inversion. The ARESNF rearrangement found in *Lacosoma valva* (Mimallonidae, Mimallonoidea) has inverted genes, along with translocated ones, as compared to ancestral ARNSEF (Timmermans et al. 2014).

Discussion

Genomic characteristics

The CGA start codon for COI has been regarded as a synapomorphy in the Lepidoptera (Kim et al. 2009), but several exceptions also exist, presenting typical ATN codons (e.g., *Ctenoptilum vasava* and *Lobocla bifasciatus* in Hesperiidae; Hao et al. 2012; Kim et al. 2014). Thus, the start codon for COI may not yet be fixed in the Lepidoptera, or a secondary change may be the source of the ATN start codon that is found infrequently in Lepidoptera (Kim et al. 2014). Nevertheless, the conservancy of the start codon as CGA in Gelechioidea may indicate that this feature is a synapomorphic character, at least in the Gelechioidea, if not in all

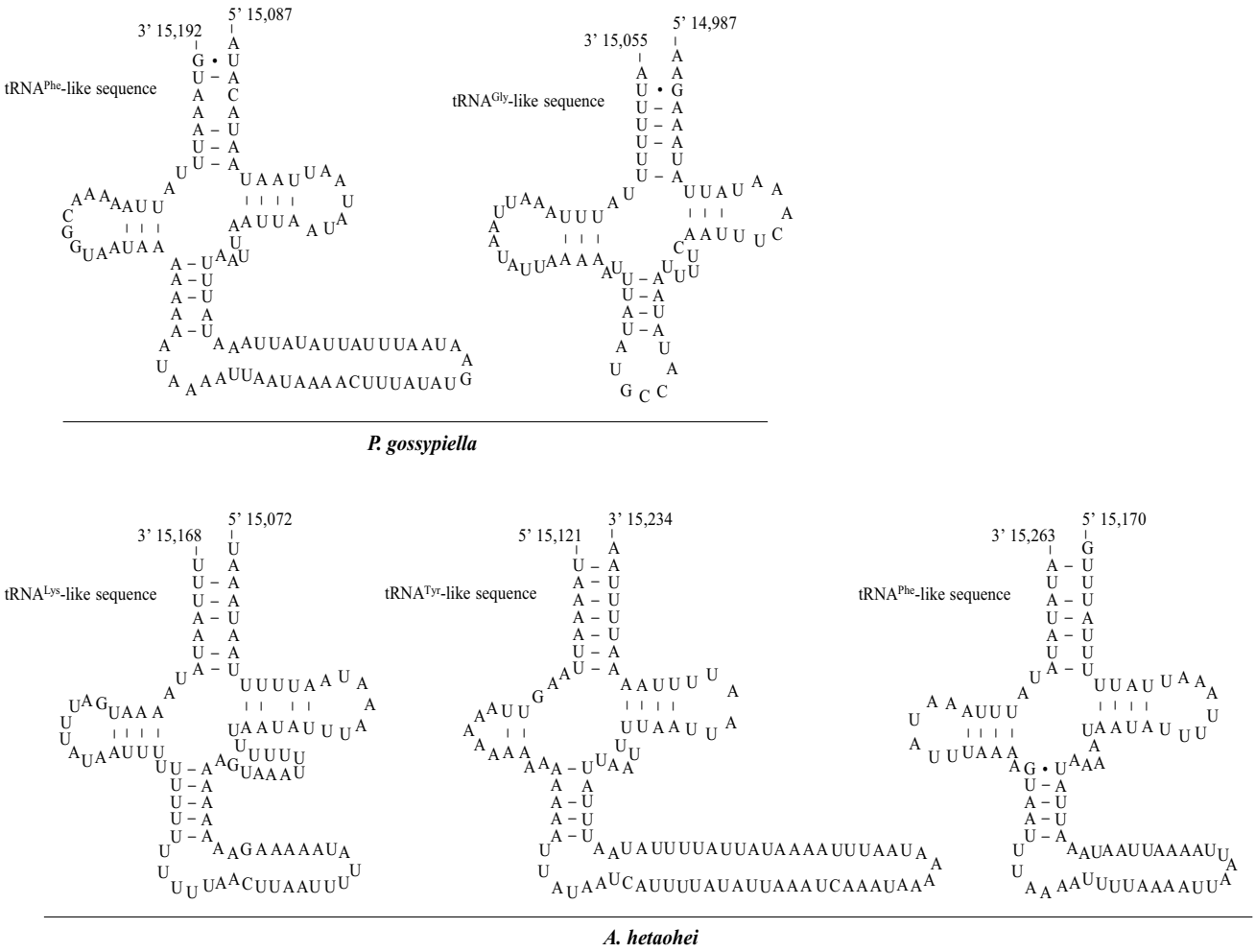

Fig. 2 continued

Lepidoptera. However, additional transcriptional data are required to clarify this issue, although recent expressed sequence tag data from a species of Crambidae in the Pyraloidea showed the start codon for COI as CGA (Margam et al. 2011).

With respect to Spacer 1 found between tRNA^{Gln} and ND2 (Sup. Table 1), a previous study indicated that this spacer originated in the course of a gene rearrangement, leading to tRNA^{Met}/tRNA^{Ile}/tRNA^{Gln} (MIQ, underline indicates an inverted gene) in the ditrysian Lepidoptera from the ancestral IQM (Kim et al. 2014). When the ancestral IQM block duplicated, a partial ND2 may have also been duplicated, resulting in IQMIQM-partial ND2. The subsequent deletion process may have accompanied the deletion of the first copy of IQ, second copy of M, and a portion of duplicated ND2, resulting in the MIQ arrangement plus a leftover portion of ND2, such as the 54 bp in *M. albilinella* and 51 bp in *D. ustalella*. If this assumption is plausible, there should be some trace of duplication, such as high sequence homology between the leftover portion of ND2 and functional ND2. In fact, sequence alignment of Spacer 1 in gelechioids including *M. albilinella* and *D. ustalella* shows substantially high sequence homology to a portion of neighboring ND2, ranging from 58 % (*Ethmia eupostica*) to 82 % (*Perimede* sp.), and this identity is obviously higher than can be attributed to chance (Fig. 5). Furthermore, species of Bombycidae, Papilionidae, Pieridae, Lycaenidae, Hesperiidae, Nymphalidae, and Saturniidae also have Spacer1 with substantially high sequence homology to a portion of neighboring ND2 (Kim et al. 2010, 2012, 2014). Spacer 2, located between tRNA^{Ser}(UCN) and ND1, is known to have a conserved motif sequence, TTAGTAT (Fig. 6), and this sequence has been suggested as the possible recognition site for the transcription termination peptide mtTERM, since it is located just past the final PCG (CytB gene) in the major strand of the mitogenome (Taanman 1999; Cameron and Whiting 2008) and found consistently in all other gelechioids (Fig. 6).

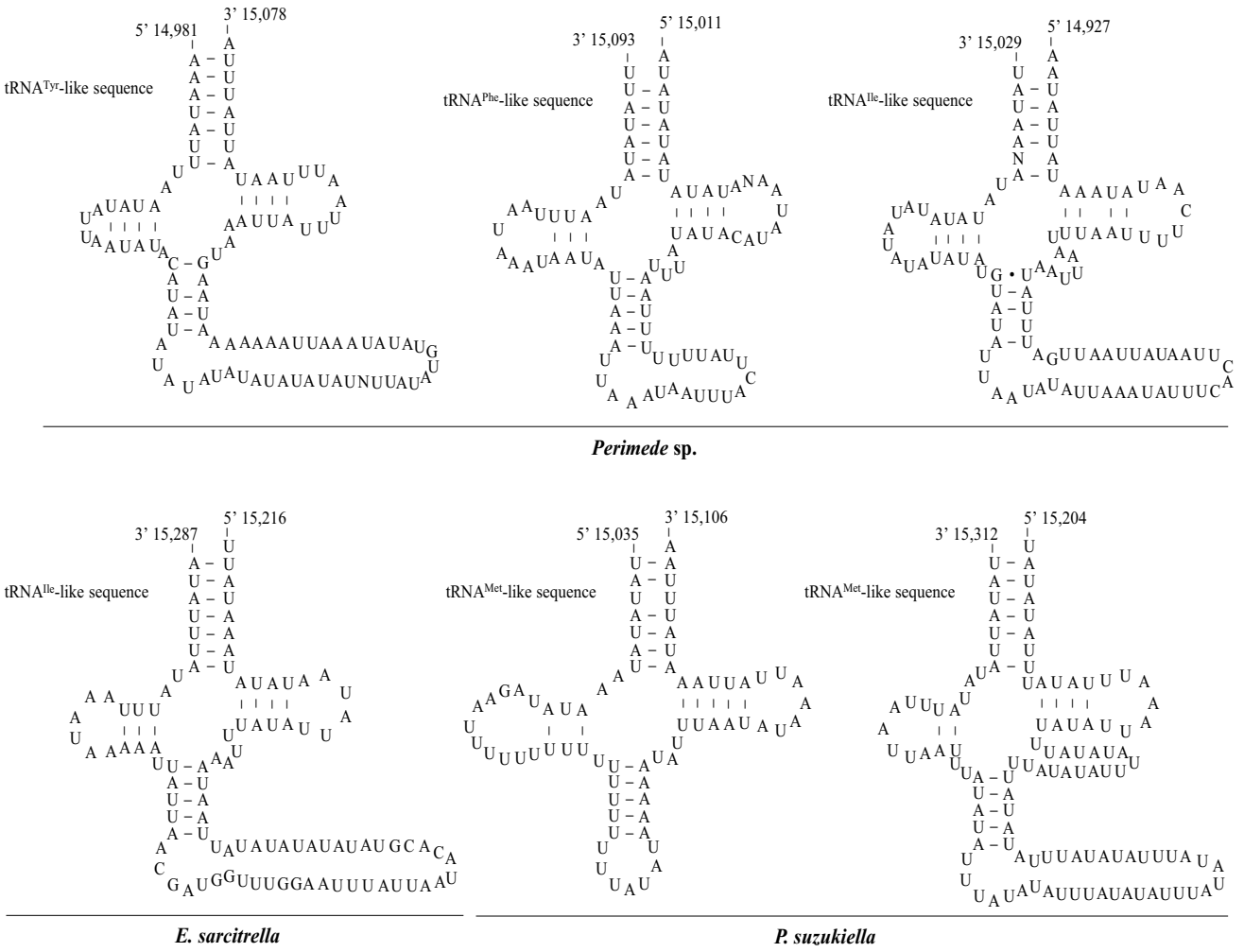

Fig. 2 continued

tRNA-like structures in the A + T-rich region have frequently been reported in Lepidoptera (Kim et al. 2009, 2011, 2014; Wan et al. 2013), and we found two tRNA-like structures in *M. albilinella* and one in *D. ustalella* (Fig. 2b). These sequences are folded into cloverleaf structures that harbor the proper anticodons and well-matched stem regions, but either the anticodon loop or variable loop contain extra nucleotides. This results in a substantial difference in length and sequence divergence between the regular tRNAs and the tRNA-like pseudo-genes for a given isotype. Our careful analysis of all available gelechioids showed that five of the seven species possessed at least one tRNA-like structure (Fig. 2b). This finding emphasizes the prevalence of tRNA-like pseudogenes in the lepidopteran A + T-rich region. These have been suggested to be spurious tRNAs of random secondary structures, owing to the reduced sequence complexity (>90 % A + T) in this non-coding region (Cameron et al. 2007). Alternatively, the tRNA-like pseudogene in the A + T-rich region has been explained in terms of the failure of the tRNA removal that functions as a primer in replication (Brown et al. 1986; Cantatore et al. 1987).

Rearrangement in Lepidoptera, including *M. albilinella*

The inverted transposition found in the *M. albilinella* mitogenome (Fig. 3) can be described through a combination of the tandem duplication-random loss (TDRL) model and recombination (Cameron 2014). The translocation of SE probably occurred by duplication of the SE block, resulting in an SESE arrangement. A subsequent random loss of S in the first copy and E in the second copy may have resulted in the change from SE to ES. Up to this point, this was the rearrangement generated by a typical TDRL model (Mortiz et al. 1987), and the most widely accepted mechanism for mitogenome rearrangements in insects (Cameron 2014; Dowton and Austin 1999). TDRL explains most of the observed mitogenomic rearrangements in insects, but

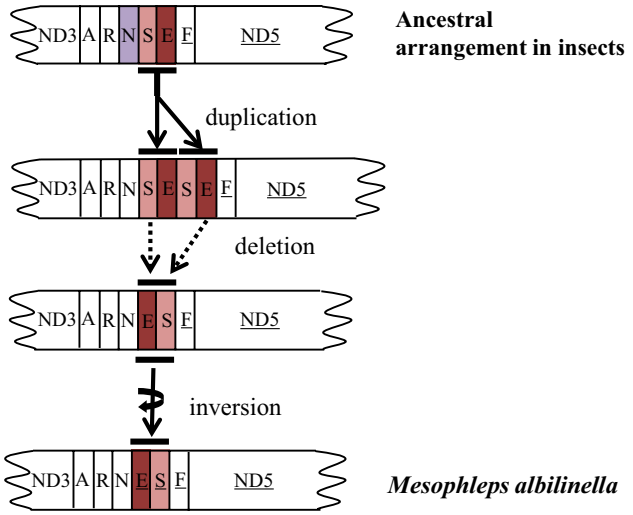

Fig. 3 Schematic illustration of the mitochondrial gene rearrangement in *Mesophleps albilinella*. Gene sizes are not drawn to scale. Gene names that are not *underlined* indicate a forward transcriptional direction, whereas *underlines* indicate a reverse transcriptional direction. tRNAs are denoted by *one-letter symbols* in accordance with the IUPAC-IUB single-letter amino acid codes. Genes and arrangements that are identical to the Ditrysia in Lepidoptera are omitted

this model cannot explain inversions without recombination (Dowton and Campbell 2001). Thus, the local inversion of ES in *M. albilinella* may have been caused by a recombination (Dowton and Campbell 2001). The sequential processes that occurred may be double-strand breakage of the mitogenome at either of the two tRNAs (either between N and ES or between ES and F); incorporation of a short, inverted segment of the two tRNAs; and re-association of the breakage, resulting in the inverted "ES" in the ARNESF cluster from the original ARNESF at the ND3 and ND5 junction (Fig. 3). However, the precipitating event between TDRL and inversion remains uncertain.

Among the six different mitogenome rearrangements found in Lepidoptera, the MIQ rearrangement found in Ditrysia, the IMQ rearrangement in *E. nyctelius* (Nymphalidae in Papilionoidea; Xuan et al. 2015), the RNSAEF rearrangement in *Astrotischeria* sp. (Timmermans et al. 2014), and the RNSAEF rearrangement in *E. montanus* (Hesperiidae in Papilionoidea; Wang et al. 2014) can all be explained in terms of TDRL (Mortiz et al. 1987). However, the rearrangements of *Lacosoma valva* (Mimallonidae in Mimallonoidea; Timmermans et al. 2014) and *M. albilinella* require inversion and translocation to achieve their current order.

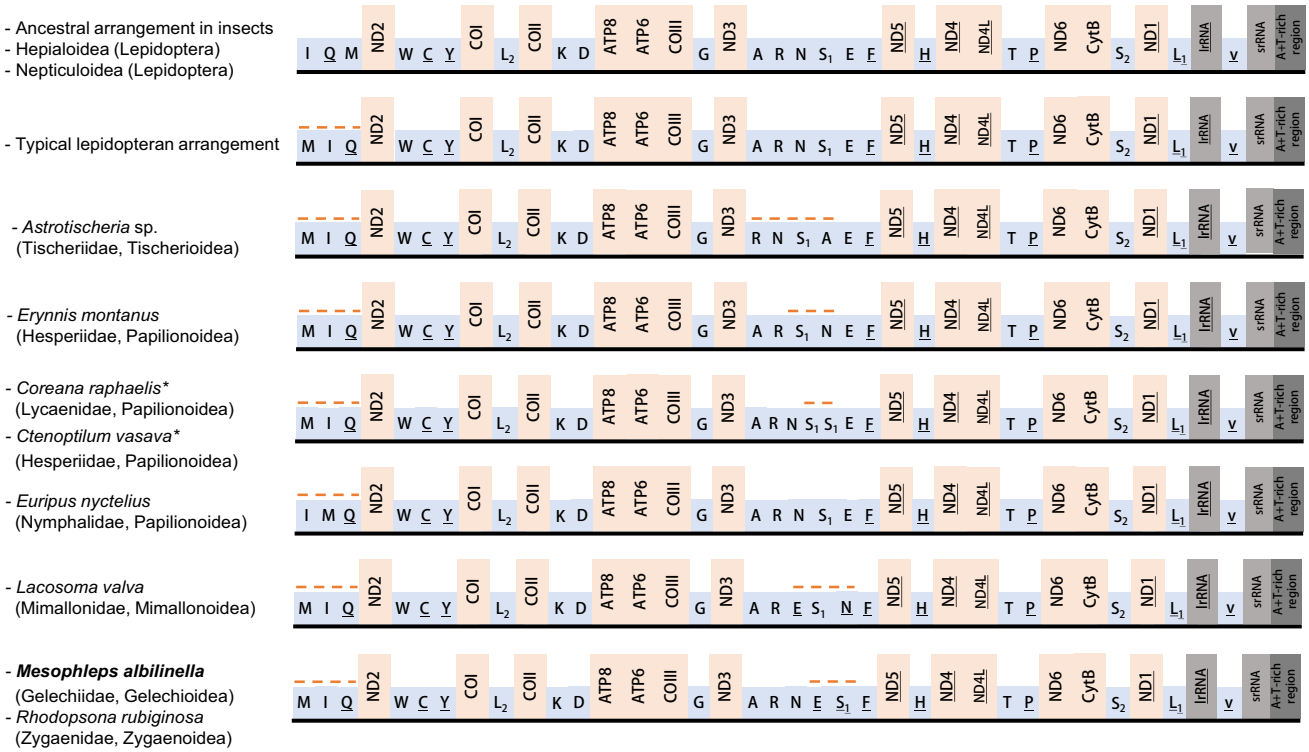

Fig. 4 Schematic illustration of the available mitochondrial gene arrangements in Lepidoptera. Gene sizes are not drawn to scale. Gene names that are not *underlined* indicate a forward transcriptional direction, whereas *underlines* indicate a reverse transcriptional direction. tRNAs are denoted by *one-letter symbols* in accordance with the IUPAC-IUB single-letter amino acid codes. *Dotted lines above* the gene names indicate rearranged genes relative to the ancestral arrangement in insects. *Note that *Coreana raphaelis* (Lycaenidae in Papilionoidea) and *Ctenoptilum vasava* (Hesperiidae in Papilionoidea) have duplicated $tRNA^{Ser}(AGN)$ (S_1) instead of gene rearrangement

***Mesophleps albilinella* (76%)**

```
ND2     5′-1061 ATTTTTATTTTTTA-AGTTTT--ATTTTTATTATTATAAGATTAATTATAATATTTTTT −3′ 1116
Spacer  5′-202  TTTTTTATAAAAAAGAATTTTTAATTCTTCTTATTATA---TT--TTATAATATTTTTT-3′ 255
                *******     * * ****  *** ** ********   **  **************
```

***Dichomeris ustalella* (76%)**

```
ND2     5′-1071 TTAATTTTAAATAAAATTTTTATTTTAACTTTTATTTTTATTTTTATAAGATTAATTATA −3′ 1130
Spacer  5′-228  ACAATTATAAA-GAAATTTTTATTTT--CTTAAA-----ATTTTTTTAA-ATTATTTTTA −3′ 278
                  **** ****  *************  ***  *     ****** *** **** ** **
```

Pectinophora gossypiella (68%)

```
ND2     5′-688  TTTTAATAAAAACT----TTTTAATTATGATTATAATTTTCAATACT-ATTATTG −3′ 737
Spacer  5′-202  TTTTAATAAAAAAAGGAATTTAAATTTCCCCCAAAATTTAAATTATTTATTTTTT −3′ 256
                ************      *** ****      * *****  * ** * *** **
```

Atrijuglans hetaohei (65%)

```
ND2     5′-1028 TGAATTG-TTATTAATTTTTTAATTATAAATAAAATTTTTTTATTAACTTTTATTTTTATTTTTATA −3′ 1093
Spacer  5′-199  TAAATTAATTAAAGAAATTATAATTTTTAACAGAAATTTTTATTT--CTTATATTTC-ATTTTTAAA −3′ 262
                * ****  ***   *  ** ***** * ** * ** *****  **  *** *****  ******* *
```

Ethmia eupostica (58%)

```
ND2     5′-468 TTTTTATTTATTATTTTATTAAAAATATTTTTTTTTAAAAACTTTGAATTTAATAACTTAATTTCA −3′ 533
Spacer  3′-263 TTTAAAAAAAATAAAAAAT-AAGAAAATTTTCTTTGAAGAATTATAAATTCAAT---TTAATAAAA −5′ 202
               ***  *   * **    ** ** ** ***** *** ** ** * * **** ***   *****   *
```

Perimede sp. (82%)

```
ND2     5′-1241 TAATTTTAAGAACTTTTATTTTTTTTTA −3′ 1268
Spacer  5′-202  AAATTTTAAGAATTTTTATTTCTTTTAG −3′ 228
                ********** ******** ****
```

Endrosis sarcitrella (72%)

```
ND2     5′-1166 AATTAAAATGATTTAAAATTAATATTAAAAATAATAATTTTTTAAT-TATTAAT −3′ 1218
Spacer  3′-246  ATTTAATTTATCTTAAAA--------AGAAATAATAATTCTTTTTTATACAAAA −5′ 281
                * ****  *   ******        * ********** ***  * **  **
```

Promalactis suzukiella (69%)

```
ND2     5′-1197 TTTTGAAATAAATAATTTAATCTCAATTATAATTAACTCTACTTTATTAATAAAAATAGGTTCTGT −3′ 1236
Spacer  5′-203  TTTTCATAAAAAGAATTTA----TAATTCT--TTAA---TAATTT--TAATTTAAATTTATTTTAT −3′ 236
                **** * * *** ******     **** *  ****   ** ***  ****  ****   ** * *
```

Oegoconia novimundi (73%)

```
ND2     5′-1197 ATTAATATTTTTAGATTCATTTCTTTATCTGGAATTTTAC −3′ 1236
Spacer  5′-203  TTTAAAAG----AATTTCAATTCTT--TCTGTATTTTTAC −3′ 236
                 **** *     *  **** *****  **** * ******
```

Fig. 5 Alignment of the intergenic spacer sequence (Spacer 1) located between $tRNA^{Gln}$ and ND2 and neighboring partial ND2 of Gelechioidea. *Asterisks* indicate consensus sequences in the alignment. *Bars* (–) were introduced to maximize sequence alignment. Sequence homology between the spacer and the ND2 is shown in the parentheses next to the species name. The nucleotide position is indicated at the beginning and end sites of the sequence

In *L. valva,* the rearrangement appears to involve two independent TDRLs, resulting in ESN, and also two independent inversions resulting in the current ESN (one for the E inversion and the other for the N inversion). Consequently, only two of the six available rearrangements in Lepidoptera involve inversion. These results are consistent with a recent summary on genomic rearrangement in insects indicating that short-range rearrangement by the TDRL model is most common and that inversion is found infrequently (Cameron 2014). In fact, inversion is rarely found in insect groups (e.g., Dermaptera, Hymenoptera, Thysanoptera, Hemiptera, and Phthiraptera) (Shao et al. 2001; Shao and Barker 2003; Thao et al. 2004; Dowton et al. 2009; Wan et al. 2012; Cameron 2014).

Gene rearrangement, which is utilized for phylogenetic markers, is the second major use of mitogenomic data in the evolutionary perspective on insects (Dowton et al. 2003). In particular, the taxonomic extent and synapomorphic status of given rearrangements have received considerable attention, particularly in the Hymenoptera and

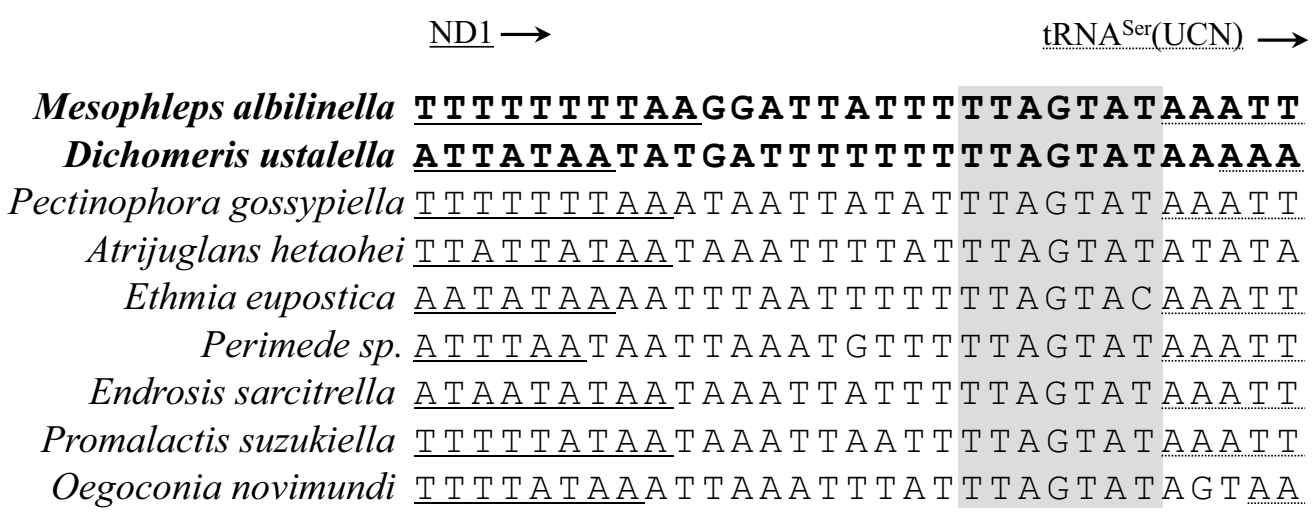

Fig. 6 Alignment of the internal spacer region (Spacer 2) located between ND1 and tRNA^{Ser}(UCN) of the Gelechioidea. The *gray-shaded* nucleotides indicate the conserved heptanucleotide region (TTAGTAT). *Underlined* nucleotides indicate the adjacent partial sequences of ND1 and tRNA^{Ser}(UCN), respectively. *Arrows* indicate the transcriptional direction

hemipteroids, which show extremely high rates of gene rearrangement (Cameron 2014). In order to understand the taxonomic extent of gene rearrangement in Lepidoptera, all available complete mitogenome sequences registered in GenBank were obtained (274 mitogenomes from 44 families in 23 superfamilies as of August 6, 2015; Sup. Table 2). The MIQ rearrangement seems to be synapomorphic in the Ditrysia (Cameron 2014), whereas the rearrangement to IMQ from MIQ occurring uniquely in *E. nyctelius* (Nymphalidae in Papilionoidea; Xuan et al. 2015) can be explained in terms of the secondary loss of the synapomorphic gene arrangement in this particular species. With the exception of the MIQ rearrangement, the rearrangements in Lepidoptera seem to be automorphic at several taxonomic scales, although such an inference is premature, since only limited data are currently available for many taxonomic groups. However, the ARSNEF rearrangement found in *E. montanus* (Pyrginae, Hesperiidae in Papilionoidea; Wang et al. 2014) is automorphic at the subfamily, family, and superfamily levels in that this rearrangement is unique among 4 species of the same subfamily, 14 species of the same family, and 15 species of the same superfamily. Likewise, the IMQ rearrangement in *E. nyctelius* (Apaturinae, Nymphalidae in Papilionoidea; Xuan et al. 2015) is automorphic at the subfamily, family, and superfamily levels. However, mitogenome sequences for other congenerics of these two species are not currently available. Thus, the status of the genus-level synapomorphy of the rearrangements remained unanswered in the current study. In addition, the sharing of the ARNESF rearrangement, which was found in a single gelechioid species, *M. albilinella* (Gelechiidae in Gelechioidea), and another superfamilial member, *Rhodopsona rubiginosa* (Zygaenidae in Zygaenoidea; Tang et al. 2014), indicates the evolutionary independence of the gene rearrangement. Therefore, the current available data suggest that most gene rearrangement in Lepidoptera is evolutionarily independent, excluding the MIQ rearrangement. This result is consistent with that of a previous investigation of mitogenome arrangement in Hymenoptera, which found that the vast majority of mt gene rearrangements are independently derived (Dowton et al. 2009). Nevertheless, more sequence data from diverse species are obviously required for more robust inference.

In summary, the six different mitogenome rearrangements found in Lepidoptera were explained mainly using TDRL, but the gene rearrangements in *M. albilinella*, and *L. valva* involve inversion, indicating that gene inversion does occur in Lepidoptera, although it is rare. Except for the MIQ rearrangement, the remaining rearrangement supports the evolutionary independence in Lepidoptera, indicating the limited utility of gene rearrangement as a phylogenetic marker. Nevertheless, future research focused on congenerics could clarify evolutionary independence at the generic level.

Acknowledgments This study was financially supported by Chonnam National University, 2013.

Compliance with ethical standards

The authors declare that they have no conflict of interest. For this type of study formal consent is not required. This article does not contain any studies with human participants performed by any of the authors. Informed consent was obtained from all individual participants included in the study.

References

Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Res 27:1767–1780

Boore JL, Collins TM, Stanton D, Daehler LL, Brown WM (1995) Deducing the pattern of arthropod phylogeny from mitochondrial DNA rearrangements. Nature 376:163–165

Boore JL, Lavrov DV, Brown WM (1998) Gene translocation links insects and crustaceans. Nature 392:667–668

Brown GG, Gadaleta G, Pepe G, Saccone C, Sbisà E (1986) Structural conservation and variation in the D loop-containing region of vertebrate mitochondrial DNA. J Mol Biol 192:503–511

Cameron SL (2014) Insect mitochondrial genomics: implications for evolution and phylogeny. Annu Rev Entomol 59:95–117

Cameron SL, Whiting MF (2008) The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. Gene 408:112–123

Cameron SL, Lambkin CL, Barker SC, Whiting MF (2007) A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision. Syst Entomol 32:40–59

Cameron SL, Yoshizawa K, Mizukoshi A, Whiting MF, Johnson KP (2011) Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). BMC Genom 12:394

Cantatore P, Gadaleta MN, Roberti M, Saccone C, Wilson AC (1987) Duplication and remodeling of tRNA genes during the evolutionary rearrangement of mitochondrial genomes. Nature 329:853–855

Cao YQ, Ma C, Chen JY, Yang DR (2012) The complete mitochondrial genomes of two ghost moths, *Thitarodes renzhiensis* and *Thitarodes yunnanensis*: the ancestral gene arrangement in Lepidoptera. BMC Genom 13:276

Curole JP, Kocher TD (1999) Mitogenomics: digging deeper with complete mitochondrial genomes. Trends Ecol Evol 14:394–398

Dowton M, Austin AD (1999) Evolutionary dynamics of a mitochondrial rearrangement 'hotspot' in the Hymenoptera. Mol Biol Evol 16:298–309

Dowton M, Campbell NJH (2001) Intramitochondrial recombination: is it why some mitochondrial genes sleep around? Trends Ecol Evol 16:269–271

Dowton M, Castro LR, Austin AD (2002) Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome 'morphology'. Invertebr Syst 16:345–356

Dowton M, Castro LR, Campbell SL, Bargon SD, Austin AD (2003) Frequent mitochondrial gene rearrangements at the Hymenoptera *nad3-nad5* junction. J Mol Evol 56:517–526

Dowton M, Cameron SL, Dowavic JI, Austin AD, Whiting MF (2009) Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral. Mol Biol Evol 26:1607–1617

Feijao PC, Neiva LS, Azeredo-Espin AML, Lessinger AC (2006) AMiGA: the arthropodan mitochondrial genomes accessible database. Bioinformatics 22:902–903

Garey JR, Wolstenholme DR (1989) Platyhelminth mitochondrial DNA: evidence for early evolutionary origin of a $tRNA^{ser}AGN$ that contains a dihydrouridine arm replacement loop, and of serine-specifying AGA and AGG codons. J Mol Evol 28:374–387

Hao J, Sun Q, Zhao H, Sun X, Gai Y, Yang Q (2012) The complete mitochondrial genome of *Ctenoptilum vasava* (Lepidoptera: Hesperiidae: Pyrginae) and its phylogenetic implication. Comp Funct Genomics 2012:1–13

Kaila L, Mutanen M, Nyman T (2011) Phylogeny of the mega-diverse Gelechioidea (Lepidoptera): adaptations and determinants of success. Mol Phylogenet Evol 61:801–809

Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066

Kim I, Lee EM, Seol KY, Yun EY, Lee YB, Hwang JS, Jin BR (2006) The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). Insect Mol Biol 15:217–225

Kim MI, Baek JY, Kim MJ, Jeong HC, Kim KG, Bae CH, Han YS, Jin BR, Kim I (2009) Complete nucleotide sequence and organization of the mitogenome of the redspotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects. Mol Cells 28:347–363

Kim MJ, Wan X, Kim KG, Hwang JS, Kim I (2010) Complete nucleotide sequence and organization of the mitogenome of endangered Eumenis autonoe (Lepidoptera: Nymphalidae). Afr J Biotechnol 9:735–754

Kim MJ, Kang AR, Jeong HC, Kim KG, Kim I (2011) Reconstructing intraordinal relationships in Lepidoptera using mitochondrial genome data with the description of two newly sequenced lycaenids, *Spindasis takanonis* and *Protantigius superans* (Lepidoptera: Lycaenidae). Mol Phylogenet Evol 61:436–445

Kim JS, Park JS, Kim MJ, Kang PD, Kim SG, Jin BR, Han YS, Kim I (2012) Complete nucleotide sequence and organization of the mitochondrial genome of eri-silkworm, *Samia cynthia ricini* (Lepidoptera: Saturniidae). J Asia Pac Entomol 15:162–173

Kim MJ, Wang AR, Park JS, Kim I (2014) Complete mitochondrial genomes of five skippers (Lepidoptera: Hesperiidae) and phylogenetic reconstruction of Lepidoptera. Gene 549:97–112

Li HH, Sattler K (2012) A taxonomic revision of the genus *Mesophleps* Hübner, 1825 (Lepidoptera: Gelechiidae). Zootaxa 3373:1–82

Lowe TM, Eddy SR (1997) TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

Margam VM, Coates BS, Hellmich RL, Agunbiade T, Seufferheld MJ, Sun W et al (2011) Mitochondrial genome sequence and expression profiling for the legume pod borer *Maruca vitrata* (Lepidoptera: Crambidae). PLoS One 6:e16444

Mortiz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu Rev Ecol Syst 18:269–292

Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. Nature 290:470–474

Park KT (1990) Three new species of genera *Brachyacma* Meyrick and *Arisloielia* Hiibner (Lepidoptera I Gelechiidae). J Appl Entomol 29:136–143

Park KT (1991) Gelechiidae (Lepidoptera) from North Korea with description of two new species. Annls Hist-Nat Mus Natn Hung 83:117–123

Park JS, Kim SS, Kim KY, Kim I (2014) Complete mitochondrial genome of Suzuki's Promolactis moth *Promalactis suzukiella* (Lepidoptera: Oecophoridae). Mitochondrial DNA Early Online:1–2

Parsons MS (1995) A review of the scarce and threatened ethmiine, stathmopodine and gelechiid moths of Great Britain (No. 16). Joint Nature Conservation Committee, Peterborough

Rokas A, Holland PWH (2000) Rare genomic changes as a tool for phylogenetics. Trends Ecol Evol 15:454–459

Saito S, Tamuea K, Aotsuka T (2005) Replication origin of mitochondrial DNA in insects. Genetics 171:433–448

Shao R, Barker SC (2003) The highly rearranged mitochondrial genome of the plague thrips, *Thrips imaginis* (Insecta: Thysanoptera): convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes. Mol Biol Evol 20:362–370

Shao R, Campbell NJH, Barker SC (2001) Numerous gene rearrangements in the mitochondrial genome of the wallaby

louse, *Heterodoxus macropus* (Phthiraptera). Mol Biol Evol 18:858–865

Taanman JW (1999) The mitochondrial genome: structure, transcription, translation and replication. Biochim Biophys Acta 1410:103–123

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729

Tang M, Tan M, Meng G, Yang S, Su X, Liu S, Song W, Li Y, Wu Q, Zhang A, Zhou X (2014) Multiplex sequencing of pooled mitochondrial genomes-a crucial step toward biodiversity analysis using mito-metagenomics. Nucleic Acids Res 42:e166

Thao ML, Baumann L, Baumann P (2004) Organization of the mitochondrial genomes of whiteflies, aphids and pysillids (Hemiptera: Sternorrhyncha). BMC Evol Biol 4:25

Timmermans MJTN, Lees DC, Simonsen TJ (2014) Towards a mitogenomic phylogeny of Lepidoptera. Mol Phylogenet Evol 79:169–178

van Nieukerken EJ, Kaila L, Kitching IJ, Kristensen NP, Lees DC, Minet J et al (2011) Order Lepidoptera Linnaeus, 1758. In: Zhang ZQ (ed) Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness. Zootaxa 3148:212–221

Wan X, Kim MI, Kim MJ, Kim I (2012) Complete mitochondrial genome of the free-living earwig, *Challia fletcheri* (Dermaptera: Pygidicranidae) and phylogeny of Polyneoptera. PLoS One 7:e42056

Wan X, Kim MJ, Kim I (2013) Description of new mitochondrial genomes (*Spodoptera litura*, Noctuoidea and *Cnaphalocrocis medinalis*, Pyraloidea) and phylogenetic reconstruction of Lepidoptera with the comment on optimization schemes. Mol Biol Rep 40:6333–6349

Wang AR, Jeong HC, Han YS, Kim I (2014) The complete mitochondrial genome of the mountainous duskywing, *Erynnis montanus* (Lepidoptera: Hesperiidae): a new gene arrangement in Lepidoptera. Mitochondrial DNA 25:93–94

Xuan S, Song F, Cao L, Wang J, Li H, Cao T (2015) The complete mitochondrial genome of the butterfly *Euripus nyctelius* (Lepidoptera: Nymphalidae). Mitochondrial DNA Early Online:1–2

Zhao J, Sun Y, Xiao L, Tan Y, Dai H, Bai L (2014) Complete mitochondrial genome of the pink bollworm *Pectinophora gossypiella* (Lepidoptera: Gelechiidae). Mitochondrial DNA Early online:1–2