An intact mitochondrial *cox1* gene and a pseudogene with different genomic configurations are present in apple cultivars 'Golden Delicious' and 'Delicious' : Evolutionary aspects

Akihito Wakatsuki*, Kazuyoshi Kitazaki*, Sumie Kato*, Yuji Kishima**, Tetsuo Mikami**

Faculty of Agriculture, Hokkaido University, N-9, W-9, Sapporo 060-8589, Japan

*These authors contributed equally to this work
**Corresponding authors

Keywords:
Apple, Cytoplasmic diversity, Evolution, Mitochondrial genome

ABSTRACT

We have characterized the mitochondrial *cox1* gene copies in two apple cultivars 'Golden Delicious' and 'Delicious'. Both the cultivars contained an intact copy and a truncated copy of *cox1*. The intact 'Golden Delicious' and 'Delicious' *cox1* genes, designated *G-cox1* and *D-cox1*, respectively, were both found to be actually transcribed to give an RNA of approximately 1.7kb. The two intact *cox1* and two truncated copies (*G-φcox1* and *D-φcox1*) shared a common 1115-bp segment flanked by four combinations of two different 5'- and 3'- sequences. PCR assay demonstrated that the configurations bearing *G-cox1* and *G-φcox1* existed in substoichiometric amounts within the mitochondrial genome of 'Delicious' whereas substoichiometric molecules carrying *D-φcox1* were present in the 'Golden Delicious' mitochondrial genome. Although ancestor/descendant relationships cannot be inferred between the *G-cox1* and *D-cox1* arrangements, the results led us to hypothesize that (1) the 1115-bp segment containing part of the progenitor *cox1* was duplicated, thereby generating a pseudo-*cox1* copy, and (2) this was followed by homologous recombination across a portion of the 1115-bp repeats which gave rise to the descendant *cox1* and pseudo-*cox1* arrangements.

1. Introduction

The cultivated apple (*Malus* × *domestica*) is a complex interspecific hybrid (Korban and Skirvin 1984; Way et al. 1990). Its primary wild ancestor has been considered to be *Malus sieversii* whose range is centered at the border between western China and the former Soviet Union (Hokanson et al. 1997). Other *Malus* species which have contributed to the genetic makeup of apple include the wild Caucasian apple (*M. orientalis*), as well as crabapples from Europe (*M. sylvestris*), Siberia (*M. baccata*), Manchuria (*M. mandshurica*), and China (*M. prunifolia*) (Janick et al. 1996; Hokanson et al. 1997). This theory of the primary ancestor has recently been challenged, however, by Coart et al.

(2006), who found the extensive sharing of plastome types between *M. x domestica* and *M. sylvestris*. They also reported that the main plastome types shared by both species were nearly absent from *M. sieversii*.

As with other important fruit crops, it is difficult to determine exactly when the apple was first domesticated, but the Greeks and Romans were growing apples at least 2500 years ago (Hancock 2004). Ten thousand or more named apple cultivars are documented, yet only a few dozen are grown on a commercial scale world-wide (Janick et al. 1996). We should also note that the world production of apples is mostly based on two cultivars: 'Delicious' and its red sports, and 'Golden Delicious', with their seedlings making up a high proportion of the new cultivars (Janick et al. 1996; Hokanson et al. 1997). The two cultivars originated in the USA as chance seedlings in the late 19th century (Way et al.1990). 'Delicious' may have reportedly an old US cultivar 'Yellow Bellflower' in its parentage, while 'Golden Delicious' seems to be derived from a cross between 'Grimes Golden' and 'Golden Reinette' (Way et al.1990). Moreover, Savolainen et al. (1995) described that the 'Red Delicious' family including 'Delicious' might have a cytoplasmic origin in a landrace 'Yellow Transparent' grown in Russia and the Baltic states. We previously used mitochondrial DNA (mtDNA) polymorphisms to characterize the cytoplasmic diversity within a range of apple cultivars and landraces (Ishikawa et al. 1992; Kato et al. 1993). The distribution of mtDNA polymorphism patterns allowed the classification of the apple genotypes into four distinct cytoplasmic groups. The four groups were described by representative cultivar or landrace within each group: 'Golden Delicious'-type, 'Delicious'-type, 'McIntosh'-type, and 'Dolgo Crab'-type.

As part of our effort to understand the molecular basis of mitochondrial genome variation giving rise to these diverse cytoplasm types, we investigated the genomic regions that distinguished the 'Golden Delicious' -type and 'Delicious'-type cytoplasms from each other. We present herein an analysis of the mtDNA rearrangement involving the *cox1* (cytochrome c oxidase subunit 1) locus. The implications of our data are discussed with respect to the evolutionary mechanisms for the generation of

mitochondrial genome diversity in apples.

2. Materials and methods

2.1 Plant material and nucleic acid preparation

Eight apple cultivars representing each of the 'Golden Delicious'-and 'Delicious'-type cytoplasms were chosen for our analysis: the former cytotype cultivars were 'Golden Delicious', 'Tsugaru' ('Golden Delicious' x 'Jonathan'), 'Jonagold' ('Golden Delicious' x 'Jonathan') and 'Fuji' ('Ralls Janet' x 'Delicious'); and the latter cytotype cultivars were 'Delicious', 'Hopa Crab', 'Red Astrachan' and 'Cellini' (Kato et al. 1993; Harada et al. 1993). Total genomic DNA was prepared from young fresh leaves by following a CTAB-based DNA extraction method (Doyle and Doyle 1990). Total RNA was isolated from leaves according to the procedure described by Chomczynski and Sacchi (1987).

2.2 Cloning and sequence analysis

Standard techniques were used in the preparation and analysis of recombinant plasmid clones in *Escherichia coli* (Hanahan 1983; Sambrook et al. 1989). The apple *cox1* locus was isolated from a *Hin*dⅢ library or a *Pst* I library of total DNA by colony hybridization using the sugarbeet *cox1* sequence as a probe (Senda et al.1991). Single-stranded DNA sequencing was conducted using the dideoxynucleotide chain-termination method (Sanger et al. 1977).

2.3 Hybridization

DNA fragments were separated by electrophoresis on 1.0% agarose gels and then blotted onto a Hybond N+ membrane (GE Healthcare Biosciences, Amersham Place, England). Total RNA was denatured, fractionated on 1.0% agarose gels containing 0.66 M formaldehyde, and

transferred to a Hybond N+ membrane according to the protocol recommended by the membrane supplier. Probes were labeled with [^{32}P] dCTP and hybridizations were performed under the conditions described by Sambrook et al. (1989).

2.4 PCR amplification

The PCR reaction mixture consisted of 100-150ng of template DNA, 5pmol of each forward and reverse primer, and GoTaq (Promega, Madison, WI) as manufacturer's instruction manual. Amplification was performed with initial denaturation at 94℃ for 5 min followed by 35 cycles of incubation at 94℃ for 30s, 54-60℃ for 30s, and 72℃ for 2-3.5 min with final extension at 72℃ for 5 min. Annealing temperature (based on Tm of the primers) and/or the extension time (according to the length of the amplicons) varied between PCRs. To obtain cDNA sequences of *cox1* transcripts, DNase-treated total RNA was subjected to reverse transcriptase-PCR (RT-PCR) using the Superscript Ⅲ First-Strand Synthesis System (Invitrogen, Carlsbad, CA). The amplified product was cloned into the Bluescript KS+ plasmid. Primers used are given in Table 1.

3. Results and discussion

3.1 Sequence and expression of the cox1 gene in 'Golden Delicious'

Kato et al. (1993) previously described that a heterologous *cox1* probe from sugarbeet hybridized to 10.0-kb and 4.2-kb *Hin*d Ⅲ fragments in 'Golden Delicious', which were not present in 'Delicious', which instead showed hybridization to 9.1-kb and 5.0-kb *Hin*d Ⅲ fragments. Hybridization of *Pst* I -digested DNAs to the same probe detected 13.0-kb and 6.5-kb fragments in 'Golden Delicious', but not in 'Delicious', where 17.0-kb and 2.5-kb fragments were identified (data not shown). First, the 6.5-kb *Pst* I fragment from 'Golden Delicious' was cloned and used for fine

mapping and sequencing. A 1563-bp open reading frame was identified as *cox1* on the basis of DNA sequence homology (Senda et al. 1991; Kadowaki et al. 1995; Quiñones et al. 1995), and was designated *G-cox1* (Fig.1, see Supplementary Fig.S1).

RNA blot analysis was carried out using a hybridization probe derived from the *G-cox1* coding region. The coding sequence showed hybridization to a transcript of approximately 1.7kb (Fig.2). To determine whether the *G-cox1* mRNA undergoes editing, reverse transcripts were PCR-amplified using primers P5 and P6. Comparison of the cDNA sequences to the genomic sequences revealed the presence of 14 C-to-U editing sites; all editing events modify the identity of the encoded amino acids (Table.2). It should also be noted that the start codon of the *G-cox1* transcript is created by editing.

3.2 cox1 pseudogene in 'Golden Delicious'

We constructed a restriction map of the region surrounding the 6.5-kb *Pst* I fragment. The restriction sites of *Bam*HI, *Eco*R I, *Hin*dIII, and *Sac*I, were mapped on the basis of the hybridization experiments with the *G-cox1* probe. Single and double digests enabled us to locate the restriction sites relative to each other (Fig.1). The map indicated that the 6.5-kb *Pst*I fragment overlapped with the 10.0-kb *Hin*dIII fragment, but not with the 4.2-kb *Hin*dIII fragment. This prompted us to clone the 4.2-kb *Hin*dIII fragment from 'Golden Delicious'. A 2.4-kb *Hin*dIII-*Pst*I subfragment that exhibited homology to the *G-cox1* was then subjected to sequence analysis. The analysis revealed the presence of a truncated *cox1* sequence (Fig.1, see Supplementary Fig.S1). This *cox1* pseudo-copy (named *G-φcox1*) and intact *G-cox1* sequences were virtually identical (97.0% sequence identity at aligned positions) from nucleotide -854 to +261 (viz. 1115-bp repeat, Fig.1, see Supplementary Fig.S1). After the *cox1* - similar region, the *G-φcox1* reading frame was extended at least another 191 codons by sequences of uncertain origin: the total length of this chimeric reading frame is unknown.

199
200 *3.3 Recombination events across the cox1-related sequences*
201
202 If the 1115-bp reiterated sequence functions as a recombination repeat,
203 we would expect to find it present in four genomic environments. For
204 example, the 6.5-kb and 13.0-kb *Pst*I fragments could, through a
205 reciprocal recombination event, give rise to the 17.0-kb and 2.5-kb *Pst* I
206 fragments. Likewise, a recombination event between 10.0-kb and 4.2-kb
207 *Hin*dIII fragments via the 1115-bp repeat could yield the 5.0-kb and
208 9.1-kb *Hin*dIII fragments (Fig.1). As mentioned above, these four
209 recombinant fragments were found in 'Delicious'.
210 The 2.5-kb *Pst*I fragment from 'Delicious' was cloned and sequenced.
211 The results indicated the presence of a truncated *cox1* copy. This *cox1*
212 pseudo-copy, named *D-φcox1,* shared a nearly identical sequence (99.2%
213 identity) with *G-φcox1*, starting 854bp upstream of the *D-φcox1* initiation
214 codon and extending at least up to the extremity of the sequenced region
215 (Fig.1, see Supplementary Fig.S1).
216 The sequence data of the upstream flank of *G-φcox1* and the
217 downstream flank of *G-cox1* were used to synthesize oligonucleotides that
218 were subsequently employed for PCR amplification of the reciprocal
219 products of the recombination in question. The primer pair P1 and P2
220 amplified an expected DNA fragment (2902bp) from 'Delicious'.
221 Sequencing analysis of the fragment allowed us to identify an intact *cox1*
222 copy (termed *D-cox1*), which shared 100% sequence identity with *G-cox1*
223 (Fig.1, see Supplementary Fig.S1). The sequence homology to *G-cox1*
224 continued for at least 320bp past the stop codon. In the 5' flanking region,
225 the homology with *G-cox1* stopped at position -854 of *cox1*, whereas the
226 homology between *D-cox1* and *G-φcox1* stretched for at least an additional
227 110-bp-long sequence, as far as the limit of the available *D-cox1* sequence
228 (Fig.1, see Supplementary Fig.S1). Moreover, *D-cox1* was judged to
229 express a transcript of 1.7kb, based upon RNA blot analysis (Fig. 2).
230
231 *3.4 Low-copy-number molecules carrying cox1 and φcox1*

We next wished to determine whether or not the products of the 1115-bp repeat-mediated recombination are maintained at a substoichiometric level in the 'Golden Delicious'- and 'Delicious'-type mitochondrial genomes. To address the issue, PCR experiments were performed using four primers (primers P1, P2, P3 and P4) which can be annealed to the unique sequences lying outside the 1115-bp repeat (Table1, Fig.1). When the two primers P2 and P3 were used with DNAs from four 'Delicious'-type cultivars ('Delicious', 'Hopa Crab', 'Red Astrachan' and 'Cellini'), a 2.8-kb region specific to the *G-cox1* arrangement was amplified (Fig.3, Supplementary Fig.S2). In addition, PCR experiments using primers P1 and P4 allowed the detection of a 1.8-kb signal specific to the reciprocal recombination product (*G-φcox1* arrangement) (Fig.3, Supplementary Fig.S2).

Amplification products corresponding to the 1.7-kb amplicon (*D-φcox1* arrangement) were also obtained when two primers (P3 and P4) were used with DNAs from four 'Golden Delicious'-type cultivars ('Golden Delicious', 'Tsugaru' 'Jonagold' and 'Fuji') (Fig.3, Supplementary Fig.S2). The nature of the PCR products was checked by nucleotide sequencing (data not shown). In higher plants, substoichiometric DNA molecules carrying recombined sequences from the so-called master genome have been repeatedly observed in their mitochondrial genomes (Small et al. 1989; Bonen and Brown 1993; Gutierres et al. 1997; Bellaoui et al. 1998; Senda et al. 1998; Woloszynska et al. 2001). Thus it seems reasonable to conclude that the recombinant configurations bearing *G-cox1* and *G-φcox1* existed in substoichiometric amounts within the mitochondrial genomes of 'Delicious'-type cultivars whereas substoichiometric molecules carrying the *D-φcox1* arrangement were present in the mitochondrial genomes of 'Golden Delicious'-type cultivars.

Our results also indicated that the reciprocal products (*D-cox1* arrangement) of the recombination events producing the *D-φcox1* substoichiometric molecules were never detected in the 'Golden Delicious' –type mitochondrial genome. The primers (P1 and P2) used here were

shown to be efficient in the PCR experiments (see Fig.3). It remains to be seen whether the *D-cox1* molecules lacked an active replication origin and were subsequently lost or whether an asymmetric recombination mechanism resulted in the formation of a single recombination product (Mackenzie 2007).

3.5 Possible scenario for the evolution of apple cox1

The 1115-bp repeat copies can be distinguished from one another by 30 nucleotide substitutions and 5 insertions / deletions (Table 3). A sequence comparison of the repeat copies gives a hint for the evolution of the apple *cox1*.

Ancestor / descendant relationships cannot be inferred between the *G-cox1* and *D-cox1* arrangements yet. However, provided that the ancestral form is the *G-cox1* arrangement, one could hypothesize that the 1115-bp segment containing the 5' coding and its flanking sequences of *G-cox1* was duplicated in an unknown progenitor genome, thereby generating *G-φcox1.* This was followed by the occurrence of the above-mentioned mutations in the common region. In the region from -384 to +261, as shown in Table 3 (also see Supplementary Fig.S1), the *G-cox1* and *G-φcox1* sequences exhibited complete identity to the *D-cox1* and *D-φcox1* copies, respectively. On the other hand, in the region from - 854 to -752 *G-cox1* and *G-φcox1* shared sequence identity with *D-φcox1* and *D-cox1*, respectively, and *G-cox1* and G-*φcox1* showed perfect sequence identity to each other from -789 to -329. *D-cox1* is therefore considered to have been generated by a recombination between the *G-cox1* and *G-φcox1* arrangements between positions -789 and -329. This recombination must also have created *D-φcox1.* Finally, we can add that six *D-cox1*-unique nucleotide substitutions (boxed in Table 3) may have occurred at some time after the recombination event in question.

Acknowledgements

Leaf samples were kindly supplied by Drs. Takeo Harada (Faculty of

Agriculture and Life Science, Hirosaki University, Japan) and Megumi Igarashi (Aomori Prefectural Industrial Technology Research Center, Japan). We are grateful to Dr Mineo Senda (Faculty of Agriculture and Life Science, Hirosaki University, Japan) for providing sugarbeet *cox1* gene. This work was supported in part by Grants-in-Aids for Scientific Researches from Ministry of Education, Science, and Culture, Japan.

References

Bellaoui, M., Martin-Canadell, A., Pelletier, G., Budar, F., 1998. Low-copy- number molecules are produced by recombination, actively maintained and can be amplified in the mitochondrial genome of Brassicaceae: relationship to reversion of the male sterile phenotype in some cybrids. Mol. Gen. Genet. 257, 177-185.

Bonen, L., Brown, G. G., 1993. Genetic plasticity and its consequences: perspectives on gene organization and expression in plant mitochondria. Can. J. Bot. 71, 645-660.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidium thiocyanate- phenol-chloroform extraction. Anal. Biochem. 162, 156-159.

Coart, E., Van Glabeke, S., De Loose, M., Larsen, A. S., Roldán-Ruiz, I., 2006. Chloroplast diversity in the genus *Malus*: new insights into the relationship between the European wild apple (*Malus sylvestris* (L.) Mill.) and the domesticated apple (*Malus domestica* Borkh.) Mol. Ecol. 15, 2171-2182.

Doyle, J. J., Doyle, J. L., 1990. Isolation of plant DNA from fresh tissue. Focus 12, 13-15.

Gutierres, S., Lelandais, C., De Paepe, R., Vedel, F., Chétrit, P., 1997. A

331 mitochondrial sub-stoichiometric *orf87- nad3- nad1* exon A
332 co-transcription unit present in Solanaceae was amplified in the genuo
333 *Nicotiana*. Curr. Genet. 31, 55-62.
334
335 Hanahan, D., 1983. Studies on transformation of *Escherichia coli* with
336 plasmids. J. Mol. Biol. 166, 557-580.
337
338 Hancock, J. F., 2004. Plant evolution and the origin of crop species. CABI
339 Publishing, Wallingford, UK.
340
341 Harada, T., Matsukawa, K., Sato, T., Ishikawa, R., Niizeki, M., Saito, K.,
342 1993. DNA-RAPDs detect genetic variation and paternity in *Malus*.
343 Euphytica 65, 87-91.
344
345 Hokanson, S. C., McFerson, J. R., Forsline, P. L., Lamboy, W. F., Ludy, J.
346 J., Djangaliev, A. D., Aldwinckle, H. S., 1997. Collecting and managing
347 wild *Malus* germplasm in its center of diversity. HortScience 32, 173-176.
348
349 Ishikawa, S., Kato, S., Imakawa, S., Mikami, T., Shimamoto, Y., 1992.
350 Organelle DNA polymorphism in apple cultivars and rootstocks. Theor.
351 Appl. Genet. 83, 963-967.
352
353 Janick, J., Cummins, J. N., Brown, S. K., Hemmat, M., 1996. Apples. In:
354 Janick, J., Moore, J. M. (Eds.), Fruit breeding. Vol. II Tree and tropical
355 fruits. Wiley, New York, pp.1-77.
356
357 Kadowaki, K., Ozawa, K., Kazama, S., Kudo, N., Akihama, T., 1995.
358 Creation of an initiation codon by RNA editing in the *cox1* transcript from
359 tomato mitochondria. Curr. Genet. 28, 415-422.
360
361 Kato, S., Ishikawa, S., Imakawa, S., Komori, S., Mikami, T., Shimamoto,
362 Y., 1993. Cytoplasmic relatedness of apple landraces and cultovars : a
363 molecular analysis. Euphytica 66, 99-102.

Korban, S. S., Skirvin, R. M., 1984. Nomenclature of the cultivated apple. HortScience 19, 177-180.

Mackenzie, S. A., 2007. The unique biology of mitochondrial genome instability in plants. In: Logan, D. C. (Ed.), Plant mitochondria. Blackwell Publishing, Oxford, pp.36-49.

Quiñones, V., Zanlungo, S., Holuigue, L., Litvak, S., Jordana, X., 1995. The *cox1* initiation codon is crated by RNA editing in potato mitochondria. Plant Physiol. 108, 1327-1328.

Sambrook, J., Fritsch, E. F., Maniatis, T., 1989. Molecular cloning : a laboratory manual. 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sanger, F., Nicklen, S., Coulson, A. R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Savolainen, V., Corbaz, R., Moncousin, C., Spichiger, R., Manen, J. F., 1995. Chloroplast DNA variation and parentage analysis in 55 apples. Theor. Appl. Genet. 90, 1138-1141.

Senda, M., Harada, T., Mikami, T., Sugiura, M., Kinoshita, T., 1991. Genomic organization and sequence analysis of the cytochrome oxidase subunit II gene from normal and male-sterile mitochondria in sugar beet. Curr. Genet. 19, 175-181.

Senda, M., Onodera, Y., Mikami, T., 1998. Recombination events across the *atpA*-associated repeated sequences in the mitochondrial genomes of bees. Theor. Appl. Genet. 96, 964-968.

Small, I., Suffolk, R., Leaver, C. J., 1989. Evolution of plant mitochondrial

397 genomes via substoichiometric intermediates. Cell 58, 69-76.
398
399 Way, R. D., Aldwinckle, H. S., Lamb, R. C., Rejman, A., Sansavini, S.,
400 Shen, T., Watkins, R., Westwood, M. N., Yoshida, Y., 1990. Apples (*Malus*).
401 In: Moore, J. N., Ballington, J. R. (Eds.), Genetic resources of temperate
402 fruit and nut crops. Intl. Soc. Hort. Sci., Wageningen, Netherlands,
403 pp.3-62.
404
405 Woloszynska, M., Kieleczawa, J., Ornatowska, M., Wozniak, M., Janska,
406 H., 2001. The origin and maintenance of the small repeat in the bean
407 mitochondrial genome. Mol. Genet. Genomics 265, 865-872.
408
409 Figure captions
410
411 Fig. 1. Restriction maps of the intact *cox1* genes (*G-cox1* and *D-cox1*)
412 and the pseudocopies *(G-φcox1* and *D-φcox1*) from 'Golden Delicious' and
413 'Delicious'. The polarity of transcription is indicated by horizontal arrow.
414 Restriction sites are B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; Ps, *Pst*I; and Sc,
415 *Sac*I. The locations of primers P1-P4 are also given. The grey box shows
416 the 1115-bp repeat sequence shared by the two *cox1* and two *φcox1*
417 arrangements.
418
419 Fig. 2. Northern-blot analysis of *G-cox1* (lane G) and *D-cox1* (lane D).
420 Total RNA was hybridized with a 0.9-kb *Bam*HI fragment containing
421 *G-cox1* (see Fig.1). Sizes of transcripts are indicated in kb.
422
423 Fig. 3. PCR detection of substoichiometric *G-cox1* and *G-φcox1* in
424 'Delicious' and of substoichiometric *D-φcox1* in 'Golden Delicions'. The
425 experiments were performed using primers P3/P2 for the detection of
426 *G-cox1* arrangement, P1/P4 for *G-φcox1* arrangement, P1/P2 for *D-cox1*
427 arrangement, and P3/P4 for *D-φcox1* arrangement (see Table 1 and Fig.1).
428 Total DNAs from 'Golden Delicious (lane G)' and 'Delicious (lane D)' were
429 used as template. Amplified substoichiometric DNA fragments were

sequenced to confirm their identity: a 2.8-kb PCR product from 'Delicious' and a 1.7-kb product from 'Golden Delicious' were isolated from agarose gels and used to prepare sequencing template employing primer pairs P3*/P2* and P3*/P4*, respectively; a 1.8-kb amplicon from 'Delicious' was excised and used to prepare sequencing template employing the primer pair P1/P4.

Supplementary Figure caption

Fig. S1 Nucleotide sequence of *G-cox1* (1), *G-φcox1* (2), *D-cox1* (3) and *D-φcox1* (4). Dots denote positions of base identity with respect to the reference sequence, *G-cox1* (1). Dashes correspond to positions of gaps inserted to optimize alignment. Positions of the primers used are underlined. Nucleotide sequence data have been deposited in the DDBJ given the accession numbers AB588943-AB588946.

Fig. S2. PCR detection of substoichiometric *G-cox1* and *G-φcox1* in 'Delicious'-type cultivars ['Hopa Crab' (lane5), 'Red Astrachan' (6) 'Cellini' (7) and 'Delicious' (8)] and of substoichiometric *D-φcox1* in 'Golden Delicions'-type cultivars ['Golden Delicious' (lane1), 'Tsugaru' (2), 'Jonagold' (3) and 'Fuji' (4)]. The experiments were performed using primers P3/P2 for the detection of *G-cox1* arrangement, P1/P4 for *G-φcox1* arrangement, P1/P2 for *D-cox1* arrangement, and P3/P4 for *D-φcox1* arrangement (see Table 1 and Fig.1).

Table 1

List of primers used along with the purpose (see Supplementary Fig.S1 for the positions of primers)

For amplification of *cox1* and *Φcox1*	
P1	5'-GAACTTAGTGCAAAAGTAAGTGGCGGAGA-3'
P2	5'-CATAAACGACTGGAAAGCGGAATCCTG-3'
P3	5'-CTGGACCAACCAGATTTCCGACAAG-3'
P4	5'-GGAGAGAGATATGAGAATGCTTGAGAGTG-3'
P2*	5'-AGAAAGCTGGGTGGAGAGAGATATGAGAATGCTTGAGAGTG-3'
P3*	5'-AAAAAGCAGGCTCTGGACCAACCAGATTTCCGACAAG-3'
P4*	5'-AGAAAGCTGGGTCATAAACGACTGGAAAGCGGAATCCTG-3'
For RT-PCR	
P5	5'-CGCTCCGGCCCCTCTCTGACAA-3'
P6	5'-TGGGGCATTGCTTGTTCTGTTAGGTTC-3'

P2,P3* and P4* have 12 additional nucleotides at the 5'-termini of P2, P3 and P4, respectively.

Table 2

Amino-acid changes predicted by RNA editing in apple *G-cox1* transcripts.
The edited nucleotides are underlined

Amino acid position*	Codon	Amino acid change	Amino acid position	Codon	Amino acid change
1	A<u>C</u>G	T→M	246	C<u>C</u>C	P→L
78	T<u>C</u>T	S→F	251	T<u>C</u>C	S→F
82	T<u>C</u>T	S→F	279	A<u>C</u>A	T→I
148	T<u>C</u>T	S→F	393	<u>C</u>AC	H→Y
181	T<u>C</u>A	S→L	465	<u>C</u>GT	R→C
220	T<u>C</u>T	S→F	466	<u>C</u>GT	R→C
236	<u>C</u>GG	R→W	475	T<u>C</u>A	S→L

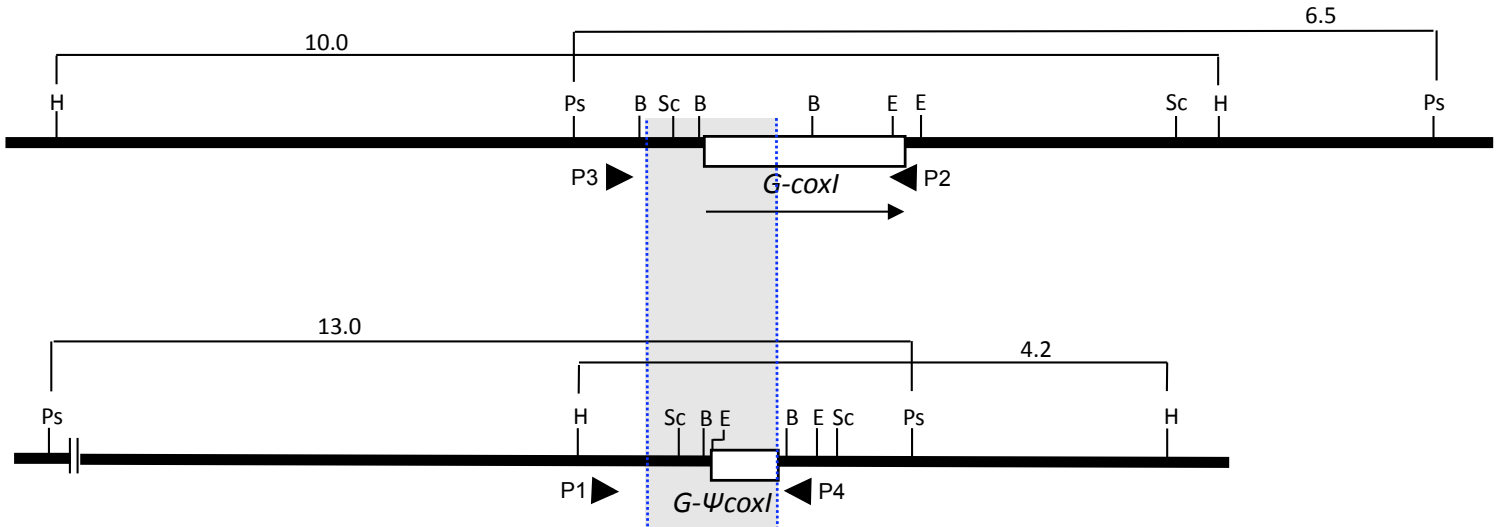
∗Deduced from the nucleotide sequence of *G-cox1* (see Supplementary Fig.S1).

Table 3

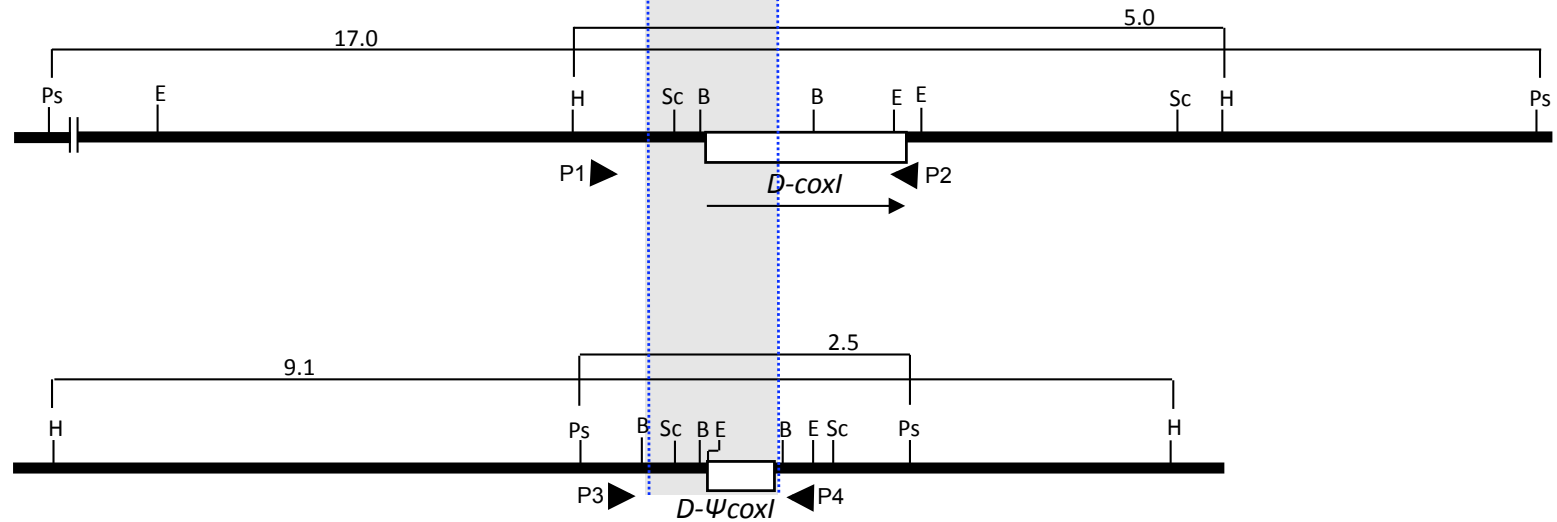
Summary of the nucleotide differences detected within the 1115-bp repeat copies.
Numbering begins with the first nucleotide of the presumed start codon (+1) of *G-cox1*. Dashes correspond to positions of gaps inserted to optimize alignment. *D-cox1*-unique nucleotide substitutions are boxed

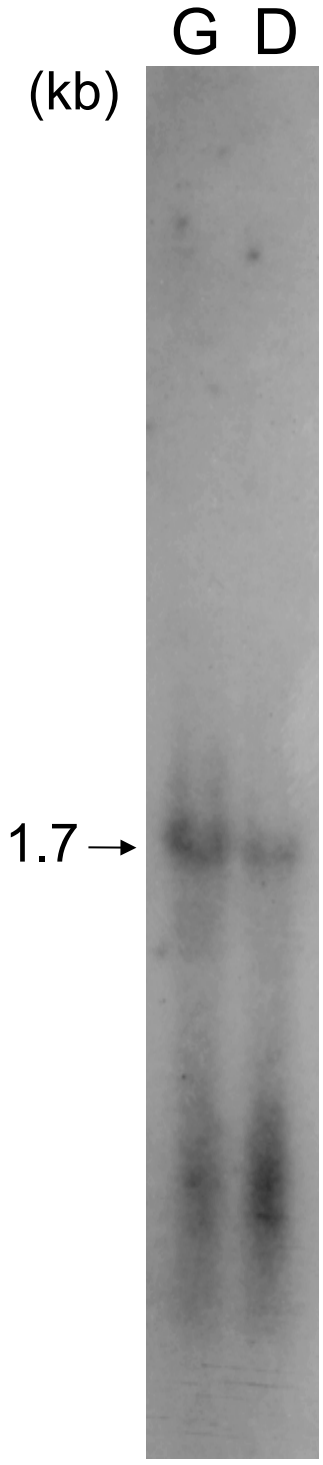
Locus	Polymorphic sites																																			
	-839	-838	-837	-836	-827	-823	-820	-791	-790	-751	-746	-742	-723	-400	-385	-328	-327	-303	-289	-274	-268	-267	-238	-205	-202	-199	-178	-175	-106	+34	+38	+233	+247	+252	+253	
G-cox1	C	A	C	A	T	G	A	G	G	G	A	A	C	A	T	T	T	G	A	C	T	G	A	G	C	G	G	G	G	G	T	C	G	G	–	
G-φcox1	–	–	–	–	G	C	C	C	T	G	A	A	C	A	T	A	A	A	G	A	G	T	C	C	A	C	A	T	A	T	C	A	T	T	C	
D-cox1	–	–	–	–	G	C	C	C	T	T	C	G	T	G	C	T	T	G	A	C	T	G	A	G	C	G	G	G	G	G	T	C	G	G	–	
D-φcox1	C	A	C	A	T	G	A	G	G	G	A	A	C	A	T	A	A	A	G	A	G	T	C	C	A	C	A	T	A	T	C	A	T	T	C	

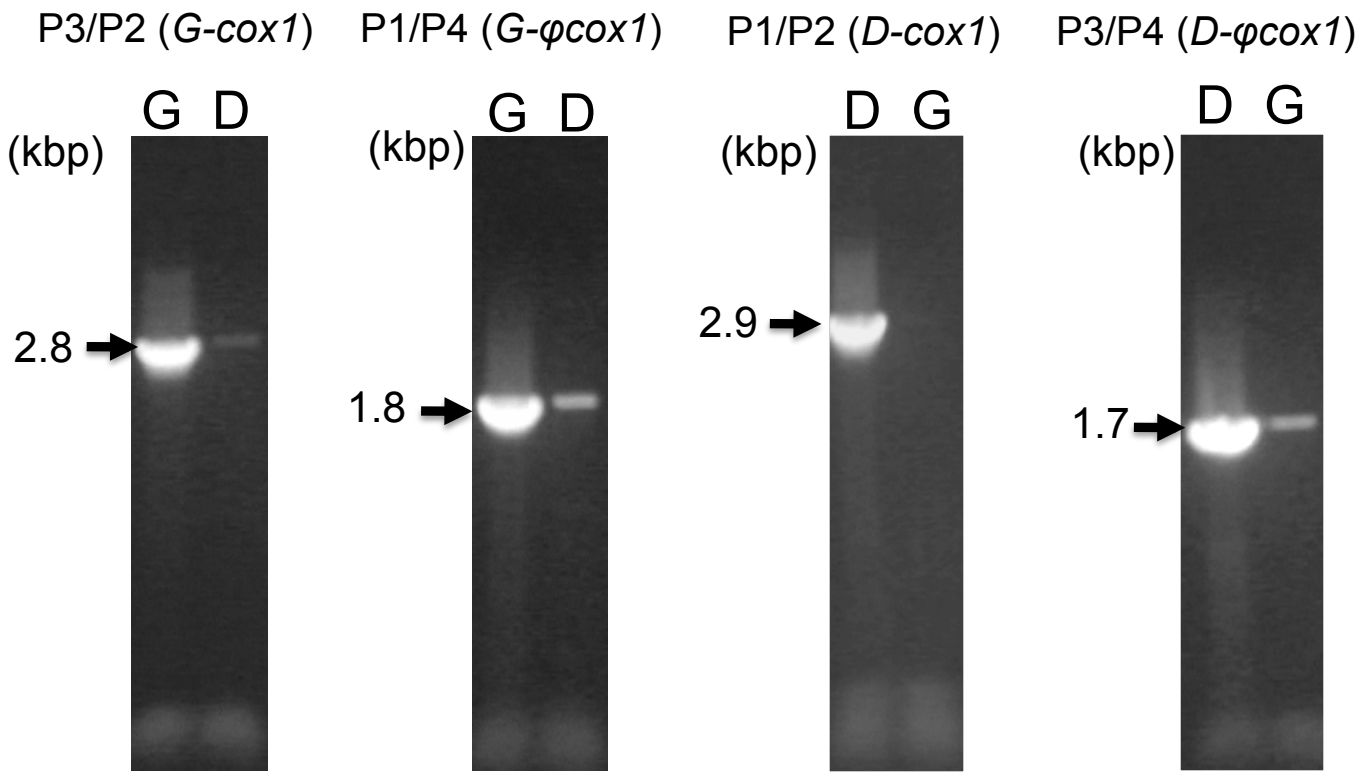
Golden Delicious



Delicious

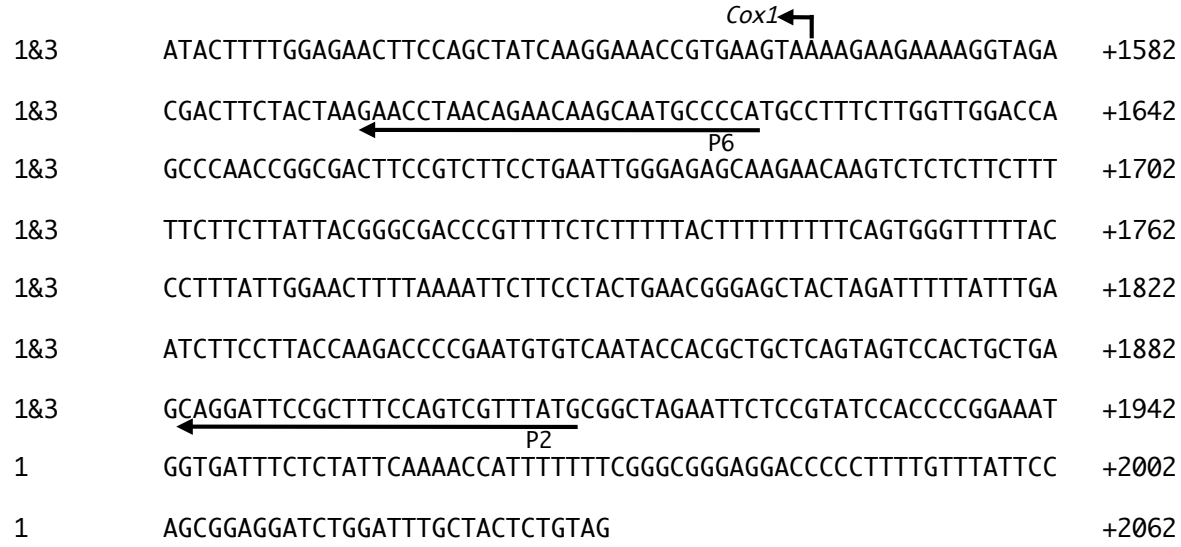



G D
(kb)
1.7

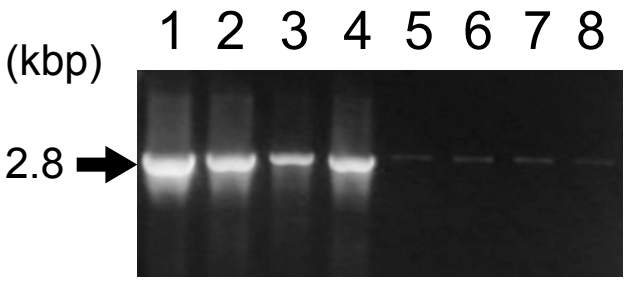

P3/P2 (G-cox1)
P1/P4 (G-φcox1)
P1/P2 (D-cox1)
P3/P4 (D-φcox1)
G D
D G
(kbp)
2.8
1.8
2.9
1.7

```
1           CTGCAGTTAAATAGGCCTTTCGTGCGATAAGGAAGAAAGTCCCGTTAAGCAAGTACTAAT   -1179

1           CATAAGGTGCGAAATGTGATGTATATCAAGACGGAGAAGGAGACGCAAGCGGAAAGATTT   -1119

1           AGATTTACATATACAGCTTTATTTAATAAAGAATTGGTAAGAACCCCATATCTTAAGTAA   -1059
2                                                          TTG ..G.AG.TGC.TG

1           GTAGTTTTATGCCTCAWAATAGAAAGGTTGCCACAGTCAAATAGAATGCTGTAGGTATAG   -999
2           TACT...C..TTTATGTCGAGATTTA.....TGTTCAGTGTACCTG.TGAAG.CC.TA.C

1           GTGGATGAAAAAAATAATGAAATCAAAATAGCGTATGAAATACGCCCAAAAGTCCCATGT   -939
2&3         ..AAGGA.CTT.GTGC.AA.GTAAGTGGCG.A.ATCC.C.A.G.TAACCGGA.G..TG.G
                 P1  P3
1           TTTTCTGGACCAACCAGATTTCCGACAAGTCTTTCTGTTAGAGCAAGAAGCGGAACTACA   -879
2&3         GC..AG.ACAAGGAGC.TCCGGA.T.TCT.AAGAAA.GA.AGAA..A...TTATCA.GA.

1&4         AGTGAATCTTTTCCGAAATGGATCCTATCAAATATTTCACATTTTCTATGATCATGTTAT   -819
2&3         GA.ACC.AGG.....TC.CATT.T.............----.................

1&4         TCTTAGGTAGACCACTCGTTAGGCTTGGGTTCGCTTACTTTTATTTGCTTTTTTGGAATT   -759
2&3         ............................CT..............................

1,2&4       GGCTTTTGGTAAACGCACCTTTTGCTGACATCGATCTTATGCATTTTTTTTCGGACTCTT   -699
3           .......T....C...G.................T.......................

1,2,3&4     TGGCCGTGAACTCTGGCTCAATCGCCGATTCCAATTCAGTTCATAATATACCTATCAATG   -639

1,2,3&4     ATCCCCCTTTCAATCCTATTGATCAAGAACCTAACAATAATGTAGTACCTCTCGACGTTG   -579

1,2,3&4     TCGATCAATTTCGTATACTCAACCAAGAGACGGAATTGGAACTCTTTGCTCGGATACGAC   -519

1,2,3&4     TTCTAGAGAACCGGTTGATTGACGGCTTGCCACCGCAAATGAACTTTGGCGAGTACGAAG   -459

1,2&4       CCTTAGTCAGAGGGTTTCTCGATGACACCCTGACCATTGGACACTACCGTAATGTCCTAA   -399
3           ................G...........................................

1,2&4       GCAATGAGCTCTTTGATCTTCGCCTGTTAGAGTTTAAGGCGAATCTCTTAGAGCAACTCG   -339
3           ............C...............................................

1&3         TCAATTTGTTTTTGAGCGAAACAACTGAGCGGCTCGCTCAAATATTGGCAGACTCGCCCT   -279
2&4         ..........AA......................A.............G..........

1&3         TTCGCGAGGGTGCCATAAGGACCGAAGCTCTTTCATTTATAAACGACTTTATGAGTCGTT   -219
2&4         ....A.....GT............................C..................

1&3         TAAATCTGAATGAGCCCCAGTCCCCTTTTGATAAAGGATTGGTGGAAGCTATGCTGCGGT   -159
2&4         ............C..A..C...................A..T................

1&3         ACTGGATCCAAGATGCCCAGCAAAACGGGAATCTCTCCGCCGTGTATAGAGAGTTCCTCC   -99
2&4         ................................................A.......

1,2,3&4     GGTATTTCCTTGGAAACTAAACCGGCCTTCCCATTTCAATAAAAAAGCCCCGCTCCGGCC   -39
                                                    Cox1          P5
1,2,3&4     CCTCTCTGACAAGGAAAGAAAGAAATAATCTCAATTTTACGACACGATGGCTGTTCTCCA   +22

1&3         CTAACCACAAGGATATAGGGACTCTATATTTCATCTTCGGTGCTATTGCTGGAGTGATGG   +82
2&4         ..........T...C.............................................
```

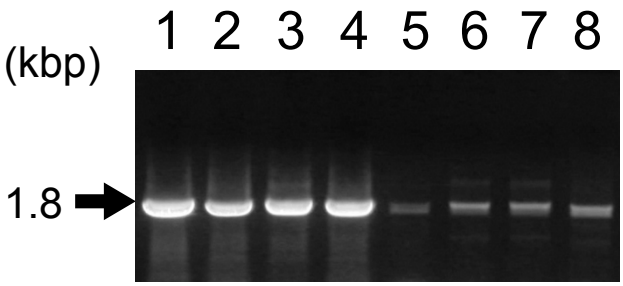
```
1,2,3&4   GCACATGCTTCTCAGTACTGATTCGTATGGAATTAGCACGACCCGGCGATCAAATTCTTG   +142

1,2,3$4   GTGGTAATCATCAACTTTATAATGTTTTAATAACGGCTCACGCTTTTTTAATGATCTTTT   +202

1&3       TTATGGTTATGCCGGCGATGATAGGCGGATCTGGTAATTGGTCTGTTCCGATTCTGATAG   +262
2&4       ............................A.............T....TCA.TCTGATT

1&3       GTGCACCTGACATGGCATTTCCACGATTAAATAATATTTCATTCTGGTTGTTGCCACCTA   +322
2&4       T..GTTT.CGT..CA.GGA..TTGCT.CGCGCTT...GG.GAAAAT.C.CACCGG.TTGC

1&3       GTCTCTTGCTCCTATTAAGCTCAGCCTTAGTAGAAGTGGGTAGCGGCACTGGGTGGACGG   +382
2&4       T.T.TC.TTCT..CC.GG.TAA..TAG..TCCCTTT.ATT..ATTTTTT.AA....GTAT

1&3       TCTATCCGCCCTTAAGTGGTATTACCAGCCATTCTGGAGGAGCTGTTGATTTAGCAATTT   +442
2&4       .ACT.TATTGTGGCTTCTCGGGGTTTCCAGGGATCCTCACT..CTG.TCGG.G.TCCC.A

1&3       CTAGTCTTCATCTATCTGGTGTTTCATCCATTTTAGGTTCTATCAATTTTATAACAACTA   +502
2&4       .CGAC.CCGTC.CCCTGCTA.AG.TC.A.C.CAATATGC.CCA.C.AGACCC.AACTGGG

1&3       TCTCCAACATGCGTGGACCTGGAATGACTATGCATAGATCACCCCTATTTGTGTGGTCCG   +562
2&4       .TAGAT.TG..ATA.ACGTCCT.CGCGTC.CC.C.CCTG.CGAGA.GGC..CACAAAT.C

1&3       TTCTAGTGACAGCATTCCTACTTTTATTATCACTTCCGGTACTGGCGGGGGCAATTACCA   +622
2&4       AATCTT.T.TTAACC..GA.TCGGCT.CGG..AC..AAAAGG.TATAATT.ACCAGTG..

1&3       TGTTATTAACCGATCGAAACTTTAATACAACCTTTTCTGATCCCGCTGGAGGGGGAGACC   +682
2&4       AAACTA.TTTTTT.GA...AGAAG..T.G.A...G.T.AT...T.AATTCTACTTT..TG

1&3       CCATATTATACCAGCATCTCTTTCGGTTCTTTGGTCATCCAGAAGTGTATATTCCCATTC   +742
2&4       TT..GG.TAGTGCATTGT.AGAC.AAC..AAC.TAG.ATTTA.TCAAGGAGCA.T.TCAA

1&3       TGCCTGGATCCGGTATCATAAGTCATATCGTTTCGACTTTTTCGGGAAAACCGGTCTTCG   +802
2&4       GCAT.CTCATATC.C..TCC.
                          P4
1&3       GGTATCTAGGCATGGTTTATGCCATGATCAGTACAGGTGTTCTTGGATTTCTTGTTTGGG   +862

1&3       CTCATCATATGTTTACTGTGGGCTTAGACGTTGATACCCGTGCCTACTTTACCGCAGCTA   +922

1&3       CCATGATCATAGCCGTCCCCACTGGAATAAAAATCTTTAGTTGGATCGCTACCATGTGGG   +982

1&3       GGGGCTCGATACAATACAAAACACCCATGTTATTTGCTGTAGGGTTCATCTTTTTGTTCA   +1042

1&3       CCATAGGAGGACTCACTGGAATAGTCCTGGCAAATTCTGGGCTAGACATTGCTCTACATG   +1102

1&3       ATACTTATTATGTGGTTGCACATTTCCATTATGTACTTTCTATGGGAGCCGTTTTTGCTT   +1162

1&3       TATTTGCAGGATTTCACTATTGGGTGGGTAAAATCTTTGGTCGGACATACCCTGAAACTT   +1222

1&3       TAGGGCAAATACATTTTTGGATCACTTTTTTCGGGGTTAATCTGACCTTCTTTCCAATGC   +1282

1&3       ATTTCTTAGGCCTTTCGGGTATGCCACGTCGCATTCCAGATTATCCGGATGCTTACGCTG   +1342

1&3       GGTGGAATGCCATTAGCAGTTTTGGCTCTTACATATCCGTAGTTGGGATTCGTCGTTTCT   +1402

1&3       TCGTGGTCGTAACAATCACTTCAAGCAGTGGAAATAACAAAAGATGCGCTCCAAGTCCTT   +1462

1&3       GGGCTGTTGAACAGAATTCAACCACACTGGAATGGATGGTACAAAGTCCTCCAGCTTTTC   +1522
```


Cox1
1&3 ATACTTTTGGAGAACTTCCAGCTATCAAGGAAACCGTGAAGTAAAAGAAGAAAAGGTAGA +1582
1&3 CGACTTCTACTAAGAACCTAACAGAACAAGCAATGCCCCATGCCTTTCTTGGTTGGACCA +1642
P6
1&3 GCCCAACCGGCGACTTCCGTCTTCCTGAATTGGGAGAGCAAGAACAAGTCTCTCTTCTTT +1702
1&3 TTCTTCTTATTACGGGCGACCCGTTTTCTCTTTTTACTTTTTTTTTCAGTGGGTTTTTAC +1762
1&3 CCTTTATTGGAACTTTTAAAATTCTTCCTACTGAACGGGAGCTACTAGATTTTTATTTGA +1822
1&3 ATCTTCCTTACCAAGACCCCGAATGTGTCAATACCACGCTGCTCAGTAGTCCACTGCTGA +1882
1&3 GCAGGATTCCGCTTTCCAGTCGTTTATGCGGCTAGAATTCTCCGTATCCACCCCGGAAAT +1942
P2
1 GGTGATTTCTCTATTCAAAACCATTTTTTTCGGGCGGGAGGACCCCCTTTTGTTTATTCC +2002
1 AGCGGAGGATCTGGATTTGCTACTCTGTAG +2062

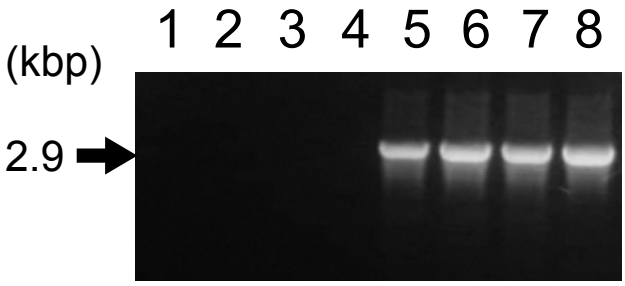
P3/P2 (*G-cox1*)



P1/P4 (*G-φcox1*)



P1/P2 (*D-cox1*)



P3/P4 (*D-φcox1*)