J Mol Evol (1993) 37:296–310

Journal
of Molecular
Evolution

Molecular Characterization and Evolution of a Duck Mitochondrial Genome

V. Ramirez, P. Savoie, R. Morais

Département de Biochimie, Faculté de Médecine, Université de Montréal, Québec, Canada, H3C 3J7

Received: 3 September 1992/Revised: 17 December 1992

Abstract. We sequenced 6,478 bp of mitochondrial DNA from Peking duck (*Anas platyrhyncos*). Eight protein genes, 11 tRNAs, part of the small and large ribosomal subunits, and the control region sequences were compared to homologous chicken sequences. The gene organization in duck and chicken is identical but differs from other vertebrates in the juxtaposition of the tRNA^{Glu}-ND6 genes next to the control region and in the lack of a hairpinlike structure between the genes for tRNA^{Asn} and tRNA^{Cys} used for light-strand replication. Protein, tRNA, and rRNA genes evolved mainly through base substitutions and small insertions and deletions. Transitions greatly outnumber transversions in the tRNA and rRNA genes, but this bias is not evident in protein genes; the control region has a higher proportion of transversions. The duck and chicken control regions show a high frequency of length mutations. Large A-T-rich nucleotide stretches dispersed across the region between the bidirectional transcription promoter and the heavy-strand replication origin in the chicken are absent in the duck. Sequence elements for heavy-strand replication in mammals are conserved in the duck and chicken control regions. Estimates of divergence for ribosomal RNAs and proteins based on total substitutions, transversions, and amino acid replacements show that all the duck/chicken values are lower than the corresponding mammal/mammal (cow, human, mouse) values. If paleontological data suggesting that avian and eutherian ordinal radiation occurred at approximately the same time are correct, this suggests that at great evolutionary distance, rate of mitochondrial DNA evolution in birds is somewhat decelerated compared to mammals.

Key Words: Mitochondrial DNA — *Anas platyrhynchos* — Gene order — Rate of evolution

During the past decade, analysis of vertebrate mitochondrial DNA (mtDNA) nucleotide sequences has provided valuable information on phylogeny and organismal evolution (Brown 1983, 1985; Kraus and Miyamoto 1991; Irwin et al. 1991). In mammals, where complete and partial nucleotide sequences from a number of species are available, the mean rate of initial sequence divergence over the whole mtDNA molecule has been estimated as being about 2%/Myr (Brown et al. 1979, 1982; Ferris et al. 1981; Miyata et al. 1982). A similar rate has been inferred for frogs, salmonid fishes, and geese (Wilson et al. 1985). Peculiar features of mtDNA sequence changes among closely related species include a high incidence of transitions in relation to transversions and a higher proportion of silent replacement substitutions (Brown and Simpson 1982; Ferris 1983). Rate of sequence divergence within the mtDNA molecule is variable, being higher in noncoding regions versus coding regions and heterogeneous within lineages at different nucleotide positions and genes. When distantly related species are compared, the transition/transversion ratio falls as the time of divergence increases, probably as a result of mutational saturation by multiple substitutions at the same nucleotide site (Brown et al. 1982; DeSalle et al. 1987). In primates, cow, and mouse the apparent rate of substitution slows substantially

---

Correspondence to: R. Morais

after 15% overall sequence divergence and is reduced by an order of magnitude at approximately 30% divergence (Moritz et al. 1987).

We have recently sequenced the complete chicken (*Gallus gallus domesticus*) and part of the Japanese quail (*Coturnix japonica*) mitochondrial genomes (Desjardins and Morais 1990, 1991). We found that the genome organization and modes of replication and transcription (L'Abbé et al. 1991) of gallinaceous mtDNAs differed from that seen in mammals and amphibians. In contrast, sequence comparisons of large mtDNA fragments suggests that gallinaceous mitochondrial genomes evolve in a similar manner to mammalian mtDNAs, in accordance with previous observations (Wilson et al. 1985). To obtain further information on genome organization, evolution, and expression of bird mtDNAs, we have been engaged in sequencing selected mtDNA fragments from other extant bird orders. Here, we report on the chemical characterization and evolution of large mtDNA segments of the Peking duck (*Anas platyrhyncos*), a distant relative of galliforms. These data will be used to address the question of tempo of evolution of avian mtDNAs.

Materials and Methods

Mitochondrial DNA was extracted from Peking duck liver following the procedure described for chicken (Morais et al. 1988). Five mtDNA fragments were generated by endonuclease digestion and cloned into pBluescript SK (clone pMtD4) and pUC18 (clones pMtD1-3,5). The relative position of the fragments on a linear representation of the duck mtDNA has been reported elsewhere (Desjardins et al. 1990). Clones pMtD1 and 2 were transferred into M13mp18-mp19 phages and subclones were obtained by exonuclease digestion of the 3′ end according to a published procedure (Dale et al. 1985). All the cloned mtDNA fragments were found stable after multiple passages through *Escherichia coli*. Single- and double-stranded DNA fragments were sequenced according to the dideoxynucleotide chain termination method (Sanger et al. 1977) using T7 DNA polymerase (Pharmacia) and either universal primers or synthetic oligomers. All sequences shown were sequenced at least twice in both orientations. The gene content was determined by comparison with mtDNA sequences available in the GenBank databases. Alignment of the sequences was performed using the program package of Corpet (1988). Alignments were maximized for sequence similarity by visual inspection.

Results and Discussion

DNA Sequence and Genomic Organization

The nucleotide sequence of several cloned regions of the duck mitochondrial genome is shown in Fig. 1. A total of 6,478 nucleotides was surveyed and the overall base composition of the light (L) strand is 30.2% A, 16.8% G, 21.8% T, and 31.2% C. Se-

```
A) COI to COIII.

     --COI--▸                                    *---------------
        P  P  Y  H  T  F  E  E  P  A  F  V  Q  V  Q  E  *
   1 CCCCTCCATACCACACCTTCGAGGAGCCAGCTTTCGTTCAAGTACAAGAAAGGAAGGAAT

     -----------tRNA(Ser)------------------------------------*  *
  61 CGAACCTCCATACACTGGTTTCAAGCCAGCTGCATTAACCACTCATGCTTCTTTCTCATG

     --------------------tRNA(Asp)-------------------------------
 121 AAATGTTAGTAAACCAATTACATAGCCTTGTCAAGGCTAAATCACAGGTGAAAGCCCTGT

              *--COII--▸
     -------*  M  A  N  H  S  Q  L  G  F  Q  D  A  S  S  P  I  M
 181 ACATCTCACGTGGCCAACCACTCCCAACTAGGATTCCAAGACGCCTCATCACCCATTATA

      E  E  L  V  E  F  H  D  H  A  L  I  V  A  L  A  I  C  S  L
 241 GAAGAGCTCGTTGAATTCCACGACCACGCTCTGATTGTTGCCTTAGCTATCTGCAGCCTA

      V  L  Y  L  L  A  H  M  L  M  E  K  L  S  S  N  A  V  D  A
 301 GTCCTATACCTCTTAGCCCACATGCTAATAGAAAAACTATCATCCAACGCAGTAGACGCC

      Q  E  V  E  L  I  W  T  I  L  P  A  I  V  L  V  L  L  A  L
 361 CAAGAAGTAGAACTAATCTGAACAATCCTACCCGCCATCGTCCTAGTACTCCTCGCCCTC

      P  S  L  Q  I  L  Y  M  M  D  E  I  D  E  P  D  L  T  L  K
 421 CCATCCCTACAAATCCTGTACATAATAGACGAAATCGACGAGCCAGACCTCACACTAAAA

      A  I  G  H  Q  W  Y  W  S  Y  E  Y  T  D  F  K  D  L  S  F
 481 GCCATTGGCCACCAGTGATACTGAAGCTACGAATACACAGACTTCAAGGACCTCTCATTC

      D  S  Y  M  I  P  T  T  D  L  P  N  G  H  F  R  L  L  E  V
 541 GACTCCTACATAATTCCCACCACAGACCTGCCAAATGGGCACTTCCGACTCCTAGAAGTT

      D  H  R  V  V  V  P  M  E  S  P  I  R  V  I  I  T  A  G  D
 601 GACCACCGCGTAGTCGTACCCATAGAATCACCGATCCGCGTAATTATTACTGCCGGAGAC

      V  L  H  S  W  A  V  P  T  L  G  V  K  T  D  A  I  P  G  R
 661 GTACTTCACTCATGAGCAGTTCCAACGCTCGGAGTTAAAACAGATGCAATCCCAGGCCGA

      L  N  Q  T  S  F  I  T  T  R  P  G  I  F  Y  G  Q  C  S  E
 721 CTAAACCAAACCTCATTCATTACCACCCGGCCTGGGATTTTCTACGGCCAGTGCTCAGAA

      I  C  G  A  N  H  S  Y  M  P  I  V  V  E  S  T  P  L  P  Y
 781 ATCTGCGGGGCTAACCACAGCTACATGCCTATTGTAGTAGAATCTACCCCACTCCCATAC

      F  E  A  W  S  S  L  L  S  S  S  *  *----------------------
 841 TTTGAAGCCTGATCATCCCTCCTATCGTCATCCTAATCATTAAGAAGCTATGCAACAGCA

                                                    *--ATPase8--▸
     tRNA(Lys)------------------------------------*  M  P  Q  L
 901 CTAGCCTTTTAAGCTAGCTAAAGAGGAATTATCCCCTCCTTAATGGCATGCCTCAACTCA

     N  P  A  P  W  F  S  I  M  V  M  T  W  L  T  L  A  L  L  I
 961 ACCCTGCACCATGATTCTCAATCATAGTCATAACCTGACTAACCCTCGCACTCCTAATCC

     Q  P  K  L  L  T  F  T  T  T  N  P  P  S  K  K  P  S  L  I
1021 AGCCAAAACTGCTAACCTTCACCACAACAAATCCCCCATCAAAAAAACCATCACTCATCA
                              *--ATPase6--▸
                               M  N  L  S  F  F  D  Q  F  S  S  P
     T  K  P  T  P  W  A  W  P  W  T  *
1081 CCAAACCCACACCATGAGCCTGACCATGAACCTAAGTTTCTTTGACCAATTCTCAAGCCC

       H  L  L  G  H  P  L  I  L  L  S  L  L  L  P  A  L  L  F  P
1141 CCACCTACTTGGTCATCCCCTGATCCTACTATCCCTGCTTCTTCCAGCCCTATTGTTCCC

       S  P  G  N  R  W  I  N  N  R  L  S  T  I  Q  L  W  L  L  H
1201 ATCCCCAGGCAACCGATGAATCAACAACCGACTATCCACCATCCAACTGTGACTCCTACA

       L  I  T  K  Q  L  M  I  P  L  N  K  N  G  H  K  W  A  L  M
1261 CCTAATCACAAAACAACTAATAATCCCATTAAACAAAAACGGCCACAAATGAGCCCTGAT

       L  T  S  L  M  T  M  L  L  T  I  N  L  L  G  L  L  P  Y  T
1321 GCTAACATCACTAATAACCATACTCCTAACAATCAACCTTCTAGGACTTCTCCCATATAC

       F  T  P  T  T  Q  L  S  M  N  M  A  L  A  F  P  L  W  L  A
1381 ATTCACCCCAACCACCCAGCTATCCATAAACATGGCCCTAGCTTTCCCCCTGTGGCTTGC

       T  L  L  T  G  L  R  N  K  P  S  A  S  L  A  H  L  L  P  E
1441 TACCCTACTAACAGGCCTGCGAAACAAACCATCAGCCTCCTTGGCTCACTTACTGCCAGA

       G  T  P  T  P  L  I  P  A  L  I  L  I  E  T  T  S  L  L  I
1501 AGGAACCCCAACACCCCTGATCCCCGCACTAATCCTGATCGAAACAACCAGCCTGCTGAT

       R  P  L  A  L  G  V  R  L  T  A  N  L  T  A  G  H  L  L  I
1561 CCGGCCCTTAGCTCTAGGAGTCCGCCTCACAGCTAACCTCACAGCAGGCCACCTACTTAT

       Q  L  I  S  T  A  S  I  A  L  K  P  I  L  P  T  V  S  I  L
1621 TCAACTCATCTCCACAGCCTCCATCGCACTCAAGCCCATCCTTCCCACAGTATCAATCCT

       T  M  A  I  L  L  L  L  T  I  L  E  V  A  V  A  M  I  Q  A
1681 AACAATAGCCATCCTACTACTCCTCACCATCCTAGAAGTAGCAGTGGCCATAATCCAGGC

                                                     *--COIII--▸
       Y  V  F  V  L  L  L  S  L  Y  L  Q  E  N  I   *M  A  H  Q
1741 CTACGTTTTCGTCCTCCTCCTAAGCCTGTACTTACAAGAAAACATCTAATGGCACACCAA

      A  H  S  Y  H  M  V  D  P  S  P  W  P  I  F  G  A  A  A  A
1801 GCACACTCCTACCACATAGTCGACCCCAGCCCCTGACCAATCTTTGGAGCTGCCGCCGCC

      L  L  T  T  S  G  L  V  M  W  F  H  Y  N  S  S  I  L  L  A
1861 TTACTCACAACCTCAGGGCTAGTCATGTGATTCCACTACAACTCATCTATCCTGCTAGCC

      A  G  L  L  S  M  L  L  V  M  L  Q  W  W  R  E  I  V  R  E
1921 GCCGGCCTCTTATCAATGCTCCTAGTGATACTCCAATGATGACGGGAGATTGTCCGAGAG

      S  T  F  Q  G  H  H  T  P  T  V  Q  K  G  L  R  Y  G  M  I
1981 AGCACCTTCCAAGGCCACCACACACCTACAGTCCAAAAAGGCCTACGATACGGCATAATC

      L  F
2041 CTCTTC
```

Fig. 1. Continued on next page.

B) ND6 to 12SrRNA.

```
       A  E  P  F  P  E  A  A  L  A  V  S  Y  V  F  V  V  L  M  G
   1 AAGCTTCCGGAAAAGGCTCCGCGGCCAGAGCTACAGAGTAGACAAATACCACCAACATGC

       G  F  Y  V  M  F  L  V  L  A  V  F  S  V  G  L  S  L  L  W
  61 CCCCAAAATACACCATAAACAGCACCAGGGCTACAAACGAAACCCCGAGGCTCAATAACC

       G  C  G  A  V  S  A  L  V  L  G  V  V  G  Y  Y  P  S  P  N
 121 ACCCACACCCAGCCACAGATGCTAAAACTAAGCCAACTACCCCATAGTACGGCGAAGGAT
                                                          ◄---ND6
       S  A  V  G  L  V  G  V  V  F  C  I  G  L  F  F  V  F  Y  T
 181 TCGATGCTACACCTAAAACACCCACAACAAAGCAGATTCCTAGAAAAAATACAAAATAAG
     ---*
       M *------------------------tRNA(glu)----------------------
 241 TCATTATTCCTGCTCGGCCACTACCCGAGACCTACGGCTCGAAAAGCCGTTATTGTTATT
     --------------**--- Control region ---►
 301 TTCAACTACCAGAACAGCTAGAATAGCCTAATAATGCTCTCAGGACCCCCCCCCCCTTCC
                                                 TAS
 361 CCCCCCAGGGGTTGCGGGGTTATTTGGTTATGCATATCGTGCATACATTTATATTCCCCA
      TAS
 421 TATATTAACCTATGGTCCCGGTAATAAACACTATTAACCAACTATCCTACATGCACGGAC

 481 TAAACCCATCACATGTCAACGGACATACCCTACTATCGGACTACCCTCCCAACGGACCCA

 541 GAGTGAATGCTCTAATACCCAACACCTCAACACCACATAACATGCCCCCAACCAGAACAA

 601 GGCCCCATAATGATGAATGCTTGACAGACATACCCTACCAACACTCCAAATTCCTCTCCA
             29 bp sequence
 661 CCCACCCATTACTCATGAAGCTGCGTACCAGATGGATTTATTAATCGTACACCTCACGTG

 721 AAATCAGCAATCCTTGCACATAATGTCCGACGTGACTAGCTTCAGGCCCATACGTTCCCC

 781 CTAAACCCCTCGCCCTCCTCACATTTTTGCGCCTCTGGTTCCTCGGTCAGGGCCATCAAT

 841 TGGGTTCACTCACCTCTCCTTGCCCTTCAAAGTGGCATCTGTGGAATACTTCCACCATCT

 901 CAATGCGTAATCGCGGCATCTTCCAGCTTTTTGGCGCCTCTGGTTCCTTTTATTTTTTCC

 961 GGGGTTACCTCACAGCTGGCCCTTCCCAGTGACTTCGGGGGTCCCACAATCTAAGCCTGG

1021 ACACACCTGCGTTATCGCGCTATCCTATATCTCAGGGATTACTCAATGAGACGGTTGGCG

1081 TATATGGGGAATCACCTTGACACTGATGCACTTTGACCACATTCAGTTAATGCTCTCTCC
                                    CSB-1
1141 ACAGCTCTATATAATAGGGCTATTTAGTGAATGCTCGATGGACATACTTTAAAAACAAAA
     CSB-2  _____CSB-3______     LSP/HSP
1201 CCACCCCAACCACAACCCCACAATATATATATACATACAAAACGAAATGCATAACATGAC

1261 CTAAATTTATTAGAGAAACTCCAGCACTAAAGACGATCCAAACCCGATGACAATCATTAC
                                                 *---------------
1321 TTTGACCTAACAAACATTACCCAATTAACCAGCCACCTGCCCCGTCCACATAGCTTACCA
     ------tRNA(phe)------------------------------------**-------
1381 CAAAAGCATGGCACTGAAGCTGCCAAGACGGCACACGAACATGCCTGCGGACAAAAGACT
     -------------------------12SrRNA----------------------------
1441 TAGTCCTAACCTTACAGTTGGTTTTTGCTAGACATATACATGCAAGTATCCGCGCCCCAG
     ------------------------------------------------------------
1501 TGTAAATGCCCTCAATAGCCTTCACCCCAGGCTTAAGGAGCGGGTATCAGGCACACCAAG
     --------------------►
1561 CAGTAGCCCAAGACGCTTGCT
```

C) 12SrRNA.

```
     --------------------------12SrRNA---------------------------
   1 CCACCGCGGTCATACAAGAGACCCAAATCAACTGTCCTACAAGCGGCGTAAAGAGTGGTA
     ------------------------------------------------------------
  61 AGATGCCTATCCTACCTAACTAAGATCAAAATGCAACTAAGCTGTCGCAAGCACAAGATG
     ------------------------------------------------------------
 121 CACCTAAACACACCAATCAAGATGATCTTAGAACTAGCGATTAATTTTAACCCACGAAAG
     ------------------------------------------------------------
 181 CCAGGGCCCAAACTGGGATTAGATACCCCACTATGCCTGGCCCTAAATCTTGATACATTA
     ------------------------------------------------------------
 241 CCCTACCGAAGTATCCGCCAGAGAACTACGAGCACAAACGCTTAAAACTCTAAGGACTTG
     ------------------------------------------------------------
 301 GCGGTGCCCTAAACCCACCTAGAGGAGCCTGTTCTGTAATCGATGATCCACGATCAACCC
     ------------------------------------------------------------
 361 AACCGCCCCTTGCCAAGCACAGCCTACATACCGCCGTCGCCAGCCCACCTCGAATGAGAG
     ------------------------------------------------------------
 421 CGCAACAGTGGCGCAACAGCACCCCGCTAATAAGACAGGTCAAGGTATAGCCTATGGGAC
     -----------------------------------------------------------►
 481 GGAAGAAATGGGCTACATTCCCTATGCATAGGGCAACACGGGAAAGAAGTATGAAACTG
```

D) 16SrRNA to ND1.

```
     ------------------------16SrRNA-----------------------------
   1 ACCTCGACTAGAGGATAAATCTAATCTAATCACCAACCTTACTGTGGGCCTTAAAGCAGC
     ------------------------------------------------------------
  61 CTTCAACAAAGAGTGCGTCAAAGCTCCACACTCAAAAATGCCAAAACAAGATGAATCCCT
     ------------------------------------------------------------
 121 TACCACAAACAGGTTAACCTATGAATATAGGAGAATTAATGCTAAAATGAGTAACTTGGG
     ------------------------------------------------------------
 181 GCCACACCCACCCCTCTAGCGGCGCAAGCTTACATGAGAACATTATTAACAGACCCAGAC
     ------------------------------------------------------------
 241 ATATACAAAAACTCCTACAAGACCAGGTATAAACTCACCCTGTTAACCCGACTCAGGAGC
     ------------------------------------------------------------
 301 GCCCATAAGAGAGATTAAAATCTGTGAAAGGAACTCGGCAAAACAAGGCCCGACTGTTTA
     ------------------------------------------------------------
 361 CCAAAAACATAGCCTTCAGCAAACAAACAAGTATTGAAGGTGATGCCTGCCCAGTGACCT
     ------------------------------------------------------------
 421 AGGTTAAACGGCCGCGGTATCCTAACCGTGCAAAGGTAGCGCAATCAATTGTCCCATAAA
     ------------------------------------------------------------
 481 TCGAGACTTGTATGAATGGCTAAACGAGGTCTTAACTGTCTCTCACGGATAATCAGTGAA
     ------------------------------------------------------------
 541 ATTGATCTCCCCGTGCAAAAGCGGGATGTGAACATAAGACGAGAAGACCCTGTGGAACTT
     ------------------------------------------------------------
 601 AAAAATCAACGGCCACCGCGAACCTAAGACTAAACCCACCGGGCTACAGACATCGCAGAG
     ------------------------------------------------------------
 661 CATGGCCGATATTTTTCGGTTGGGGCGACCTTGGAGAACAACAGATCCTCCAAAAACAAG
     ------------------------------------------------------------
 721 ACCACACCTCTTTACTTAGAGCCACCCCTCAAAGTGCTAATAGTGACCAGACCCAATATA
     ------------------------------------------------------------
 781 ATTGATTAATGGACCAAGCTACCCCAGGGATAACAGCGCAATCCCCCTCAAGAGCCCCTA
     ------------------------------------------------------------
 841 TCGACAGGGGGGTTTACGACCTCGATGTTGGATCAGGACATCCTAATGGTGCAGCCGCTA
     ------------------------------------------------------------
 901 TTAAGGGTTCGTTTGTTCAACGATTAATAGTCCTACGTGATCTGAGTTCAGACCGGAGCA
     ------------------------------------------------------------
 961 ATCCAGGTCGGTTTCTATCTATGAACTACTCTCCCCAGTACGAAAGGACCGGGAAAGTAA
     ------------------------------------------------------------
1021 GGCCAATACTACAAGCACGCCTTCCCTCTAAGTAGTGAAACCAACTCAACTATGAAGAGG
     -----------------------------------**-----------------------
1081 ACTCCCCCCCACCACCCCAACTCCTAGAAAAGGATCAGCTAGAGTGGCAGAGCCGGCAAA
                                                            *----
     ---tRNA(Leu)--------------------------------------*     M  P
1141 ATGCAAAAGGCTTAAGCCCTTTACCCAGAGGTTCAAATCCTCTCCCTAGCTACACATGCC
     ND1--►
       Q  T  T  M  V  S  Y  L  I  M  A  L  L  Y  I  I  P  I  L  I
1201 ACAAACAACAATAGTAAGCTACCTCATTATAGCCCTGCTATACATCATCCCAATCTTAAT

       A  V  A  F  L  T  L  V  E  R  K  I  L  S  Y  M  Q  S  R  K
1261 TGCCGTGGCTTTCTTGACTCTAGTCGAACGAAAAATTCTAAGCTACATGCAATCCCGTAA

       G  P  N  I  V  G  P  F  G  L  L  Q  P  I  A  D  G  I  K  L
1321 AGGCCCCAACATCGTGGGCCCTTTTGGCCTGCTCCAACCCATTGCAGACGGAATCAAACT

       F  I  K  E  P  I  R  P  S  T  S  S  P  L  L  F  I  M  M  P
1381 ATTCATTAAAGAGCCCATTCGACCTTCCACCTCCTCACCGCTCCTCTTCATCATAATGCC

       M  L  A  L  L  L  A  L  T  A  W  V  P  L  P  L  P  F  S  L
1441 CATACTAGCCCTCCTCCTAGCCCTCACCGCCTGAGTGCCCCTCCCCCTCCCGTTCTCACT

       V  D  L  N  L  G  V  L  F  M  V  A  M  S  S
1501 AGTAGACCTGAACCTCGGGGTCCTCTTTATAGTAGCCATATCAAGCTT
```

E) ND2 to COI.

```
     ---ND2--►
      M  K  Q  W  Y  T  S  K  P  P  S  T  P  T  A  I  L  A  S  L
   1 ATAAAACAGTGGTACACTAGCAAACCCCCAAGCACGCCCACCGCAATCCTAGCCTCACTA

      S  I  L  L  L  P  L  S  P  M  V  H  A  I  V  *-------------
  61 TCAATCCTCCTACTCCCCCTCTCCCCCATAGTCCACGCTATTGTCTAGAAACTTAGGATA

     ----------------tRNA(Trp)-----------------------------------
 121 ACACCCACCTAAACCGAAGGCCTTCAAAGCCTTAAATAAGAGTTAAACCCTCTTAGTTTC

     -*   *--------------------------tRNA(Ala)-------------------
 181 TGCGCTAAGACCAACAGGACAGTAACCTGTATCTCCTGGATGCAAACCAGACGCTTTAAT

     -------------*  *----------------------------tRNA(Asn)------
 241 TAAGCTAAAGCCTTTGCTAGACAGACGGGCTTCGATCCCGCAAAATTTAGTTAACAGCT

     ----------------------------**---------------------tRNA(Cys)
 301 AAACGCCCAAACCTACTGGCCTCTGCCTAAGGCCCCGGTACACTCTCGTGCACATCGATG

     ----------------------------------*-------------------------
 361 AGCTTGCAACTCAACATGAACTTCACTACAGGGCCGATAAGAAGAGGAATTGAACCTCTG

                                                    *--COI--►
     -tRNA(Tyr)-----------------------------------*  M  T  F  I
 421 TAAAAAGGACTACAGCCTAACGCTTTAAACACTCAGCCATCTTACCCGTGACCTTCATCA

     N  R  W  L  F  S  T  N  H  K  D  I  G  T  L  Y  L  I  F  G
 481 ATCGATGACTATTTTCTACCAATCACAAAGACATCGGTACTCTATACCTTATCTTCGGGG

     A  W  A  G  M  I  G  T  A  L  S  L  L  I  R  A  E  L  G  Q
 541 CATGAGCCGGAATAATTGGCACAGCACTCAGCCTACTGATCCGGGCAGAACTAGGCCAGC

     P  G  T  L  L  G  D  D  Q  I  Y  N  V  I  V  T  R  H  A  F
 601 CAGGGACCCTCCTGGGCGACGACCAAATTTATAACGTGATCGTCACCCGTCACGCCTTCG

     V  M  I  F  F  M  V  M  P  I  M  I  G  G  F  G  N  W  L  V
 661 TAATAATCTTCTTCATGGTAATGCCCATCATAATTGGAGGGTTCGGCAACTGATTGGTCC

     P  L  M  I  G  A  P  D  M  A  F  P  R  M  N  N  M  S
 721 CCCTGATAATCGGTGCCCCCGACATAGCATTCCCACGAATAAACAACATAAGCTT
```

Fig. 1. Continued. Nucleotide sequence of various cloned duck mitochondrial DNA fragments **(A–E)**. Sequences shown are those of the L-strand and are numbered commencing from the first nucleotide at the 5′ end of each mtDNA fragment. All coding regions except those for ND6 and six tRNA (Glu, Ser, Ala, Asn, Cys, Tyr) genes are transcribed from the H-strand. The encoding tRNA and rRNA genes are indicated by *broken lines* above the nucleotide sequence and are delimited by *asterisks* to indicate the putative 5′ and 3′ encoded nucleotide. Anticodons are *underlined*. Translation of the protein genes is indicated above the sequence using the one-letter amino acid code. Stop codons are designated by *asterisks*. Also *underlined* in the control region **(B)** are the positions for the conserved sequence blocks (CSB-1, -2, -3), the putative bidirectional promoter (LSP/HSP), the termination-associated sequences (TAS), and the 29-bp CSB-1-like sequence. CSB-3 is *overlined*.

```
           AA         D    D             D      AC    AC      AC    V     T     T       T     AA
          stem       stem loop          stem   stem  loop    stem  loop  stem  loop    stem  stem

ALA      :AAGGCTT TA GCTT AATTA         AAGC G TCTGG TTTGCAT CCAGG AGAT  ACAGG TTAGTGT CCTGT TGGTCTTA
          G****C* ** **** *****         **** * ****A ******* T**** ****  ***** ***A*** ***** ********

ASN      :TAGGCAG AG GCCA GTAGGTTT      GGGC G TTTAG CTGTTAA CTAAA ATTTT GCGGG ATCGAAG CCCGT CTGTCTAG
          ******* *A **** A*T***G*      T*** A ***** ******* ***** ***G* AT*** ******* ***A* ********

ASP      :GAAATGT TA GTAA ACCAA         TTAC A TAGCC TTGTCAA GGCTA AATC  ACAGG TGAAAGC CCTGT ACATCTCA
          **G*C** ** **** *****         **** * ***A* C****** *A*** ****  ***** **C**A* ***** ********

CYS      :GGCCCTG TA GTGA AGT           TCAT G TTGAG TTGCAAG CTCAT CGAT  GTGCA CGAGAG  TGTAC CGGGGCCT
          *A*T*** ** **** ***           **** A A**** ******A ***G* T***  **A** *T*A**  ***G* *****T**

GLU      :GTTCTGG TA GTTG AAAATAA       CAAT A ACGGC TTTTCGA GCCGT AGGT  CTCGG GTAGTGG CCGAG CAGGAATA
          ****CC* ** **** *G**---       ***C * *T*** *****A* ***** **TC  **T** *GTC*AA **A** *G******

LEU(UUR) :GCTAGAG TG GCAG AGC-CGGCAA    ATGC A AAAGG CTTAAGC CCTTT ACCC  AGAGG TTCAAAT CCTCT CCCTAGCT
          *****C* ** **** ***T******    **** * ***** ******* ***** *T**  ***** ******* ***** ********

LYS      :CATTAAG AA GCTA TGCAA         CAGC A CTAGC CTTTTAA GCTAG CTAAA GAGGA ATTATC- CCCTC CTTAATGG
          ******* ** **** ****C         **** * ***** ******* ***** --*G* ****G GAC*C*C T**C* *******A

PHE      :GTCCACA TA GCTT ACCA-CAA      AAGC A TGGCA CTGAAGC TGCCA AGAC  GGCAC ACGAAC  ATGCC TGCGGACA
          *C**C** ** **** *A*CC***      **** * ***** ******A ***** ***T  **T** -*T*CT  **A** **T**G**

SER(UCN) :GAGAAAG AA GCAT -GAGTGGTTA-   ATGC A GCTGG CTTGAAA CCAGT GTAT  GGAGG TTCGATT CCTTC CTTTCTTG
          A****** ** **** TA*******TG   **** G *T*** ******* ***AC A*G*  *AG** ******* ***** ********

TRP      :AGAAACT TA GGAT -AACACCCACCTA AACC G AAGGC CTTCAAA GCCTT AAAT  AAGAG TTAAACC CTCTT AGTTTCTG
          ******* ** **** T**TGT*****-* **** A ***** ******* ***** ****  ***** ******T ***** ********

TYR      :GGTAAGA TG GCTG AGTGTTTA      AAGC G TTAGG CTGTAGT CCTTT TTAC  AGAGG TTCAATT CCTCT TCTTATCG
          ******* ** **** ******-G      **** * ***** ******* ***** ****  ***** ******* ***** ********
```

Fig. 2. Sequence comparison of the duck tRNAs (sense strand) with those from chicken (Desjardins and Morais 1990). In each case, the duck sequence is shown in full. Numbering of the equivalent sequence region in duck is the following: Fig. 1A, tRNA^{Ser} (45–117); tRNA^{Asp} (120–188); tRNA^{Lys} (875–946). Fig. 1B, tRNA^{Glu} (245–315); tRNA^{Phe} (1364–1433). Fig. 1D, tRNA^{Leu} (1118–1191). Fig. 1E, tRNA^{Trp} (107–182); tRNA^{Ala} (186–254); tRNA^{Asn} (257–329); tRNA^{Cys} (330–395); tRNA^{Tyr} (395–466). Residues in chicken homologous to those of duck are indicated by *asterisks* and missing nucleotides by *dashes*.

quenced fragments contain the entire control region; the protein genes COII, ATPase6, and ATPase8; 11 tRNA genes; and partial sequences of the genes encoding ND1, ND2, ND6, COI, COIII, and the small and large ribosomal subunits. As in chicken and other vertebrates, the duck mitochondrial genome is organized in an economical fashion. Some contiguous genes are butt-joined: ATPase6-COIII, ND6-tRNA^{Glu}, and tRNA^{Asn}-tRNA^{Cys}. Others are separated by short noncoding sequences of a few nucleotides or overlapped: COI-tRNA$^{Ser(CUN)}$, ATPase8-ATPase6, tRNA^{Cys}-tRNA^{Tyr}, ND2-tRNA^{Trp}. All the structural genes are punctuated by one or more tRNA genes at either their 5′ or their 3′ or both their ends.

The relative position and the orientation of all genes and the control region in the duck mitochondrial genome are identical to those found in homologous regions of chicken and quail (Desjardins and Morais 1990, 1991). Compared to other vertebrates, duck and gallinaceous birds mtDNAs display two peculiar characteristics. First, the contiguous tRNA^{Glu}-ND6 genes are located immediately upstream of the control region in duck (Fig. 1B) and gallinaceae (Desjardins and Morais 1990, 1991), and thus are transposed with respect to the organization in mammals (Anderson et al. 1981, 1982; Bibb et al. 1981; Gadaleta et al. 1989; Arnason et al. 1990; Arnason and Johnsson 1992), *Xenopus laevis* (Roe et al. 1985), and fishes (Johansen et al. 1990). Second, duck (Fig. 1E) and gallinaceae lack a hairpinlike structure located between the genes for tRNA^{Asn} and tRNA^{Cys} in vertebrates, which serves as start site for the light (L)-strand replication in mammals (Hixson et al. 1986). Since both Galliformes and Anseriformes are ancient lineages (Brodkorb 1964), our observations suggest that the molecular events causing these changes took place in an early bird ancestor. Polymerase chain reaction and DNA sequencing showed that the gene organization seen in duck and gallinaceae is also characteristic of many other extant bird orders (Morais et al., unpublished results). These observations, and those of Pääbo et al. (1991) showing that tRNA genes in the vicinity of the putative L-strand replication origin in marsupials have been rearranged, clearly suggest that the arrangement of vertebrate mtDNAs is more fluid than previously thought.

Transfer and Ribosomal RNA Genes

We have identified 11 tRNA genes from their location and nucleotide similarity with corresponding sequences in chicken and other vertebrates. Sequence comparisons indicate that these genes are highly homologous to their chicken counterparts (Fig. 2). Nucleotide identity ranges from 76.5% for tRNA^{Glu} to 98.6% for tRNA^{Tyr}, with a mean of 88.2% (Table 1). This value is higher than those for

Table 1. Sequence comparisons between duck and chicken mitochondrial DNA genes

Gene	Length (nt)	% Nucleotide identity	% Amino acid identity	% Amino acid similarity	Transition			Transversion				
					A-G	C-T	Total	A-C	A-T	G-C	G-T	Total
ND1	353	76.2	80.3	95.7	15	26	41	27	8	6	2	43
ND2	105	67.6	54.3	85.7	8	11	19	7	3	4	1	15
ND6	244	77.9	82.7	98.8	13	14	27	13	9	5	0	27
COI	358	81.6	97.5	98.3	14	26	40	12	7	6	1	26
COII	684	83.5	93.8	98.7	18	37	55	36	14	7	1	58
COIII	258	81.5	90.7	98.8	16	9	25	12	7	4	0	23
ATPase 8	165	73.9	72.2	94.4	3	11	14	16	8	2	3	29
ATPase 6	681	80.0	88.5	96.0	21	40	61	52	13	8	2	75
16S rRNA	1117	84.1	—	—	41	55	96	44	21	10	7	82
12S rRNA	687	87.4	—	—	23	28	51	15	10	6	3	34
tRNA (11)	781	88.2	—	—	28	39	67	7	6	9	3	25
Control region	1048	71.0	—	—	40	73	113	55	78	24	34	191

genes encoding rRNAs and proteins, in agreement with the situation in mammals when distantly related species are compared (Brown 1985; Cantatore and Saccone 1987). The high degree of conservation of the tRNA genes likely depends on functional constraints associated with protein synthesis on mitochondrial ribosomes, processing of the H- and L-strand transcripts (Ojala et al. 1981) and regulatory roles such as transcription termination (Christianson and Clayton 1988; Kruse et al. 1989). Evolution of the avian mitochondrial tRNA genes mainly involves point mutations, but length variations of one to three nucleotides also contribute to this process. Transitions, which greatly outnumber transversions (Table 1), are found mainly in stem regions and nearly half of them are compensated by transitions in the complementary strand such that base pairing is maintained (Fig. 2). Transversions are mainly confined to the DHU, TYC, and variable loops, along with addition/deletion events.

We have sequenced about two-thirds of the 12S and 16S rRNA duck genes (Fig. 1B–D). Percent nucleotide identities between the duck and chicken sequences are respectively 87.4 and 84.1 for 12S and 16S rRNAs (Table 1). Our comparisons revealed that avian rRNAs genes evolve similarly to tRNA genes (Table 1). Both 12S and 16S rRNAs can be folded by base pairing into domains which involve conserved structural elements similar to those found in mammals (Glotz et al. 1981; Zwieb et al. 1981) and *Xenopus laevis* (Roe et al. 1985; Dunon-Bluteau and Brun 1986). Most base substitutions and length mutations are found in nonpaired segments. In stem regions, substitutions account for about 40% of the total and consist mainly of transitions, whereas transversions greatly outnumber transitions in nonpaired segments. Similar observations have been made in the comparison of the chicken and quail 12S rRNA sequences (Desjardins and Morais 1991) and in those of the 12S and 16S rRNA genes of various mammals (Mindell and Honeycutt 1990).

Codon Usage and Protein Genes

We have described 952 codons and more than 80% of them are specified either as hydrophobic (L, I, C, M, V, F, Y, W) or weakly neutral hydrophobic (P, A, G, S, T) amino acids (Table 2). All proteins contained about the same percentage of hydrophobic residues, which is consistent with the fact that these proteins are located within the mitochondrial inner membrane (Attardi and Schatz 1989). The overall codon usage has a strong bias against the use of triplets ending in G. Nearly all codon families show a net preference for A and C at the silent position. More than 70% of all codons end in A and C and about 11% in T or G. The infrequent use of T and G at the third position has previously been reported for similar genes in chicken and quail (Desjardins and Morais 1990, 1991) and for closely and distantly related birds for a part of the cytochrome b gene (Kocher et al. 1989; Edwards and Wilson 1990). A similar tendency to exclude G and T from silent positions, which goes beyond the bias in base composition of the coding strand, was also observed in other vertebrates, including fishes (Johansen et al. 1990), and likely corresponds to a codon strategy elaborated by mitochondria over evolutionary time (Cantatore and Saccone 1987).

The protein genes encoded by duck mtDNA show a relatively high degree of sequence homology with their chicken counterparts at both the nucleotide and amino acid levels (Table 1). The relative order of nucleotide and amino acid sequence conservation among the eight protein genes appears to be the same as in species representing different

Table 2. Codon usage in Peking duck mitochondria[a]

TTT	(phe)	15	TCT	(ser)	5	TAT	(tyr)	5	TGT	(cys)	1
TTC	(phe)	31	TCC	(ser)	18	TAC	(tyr)	22	TGC	(cys)	4
TTA	(leu)	13	TCA	(ser)	25	TAA	(ter)	3	TGA	(trp)	21
TTG	(leu)	6	TCG	(ser)	4	TAG	(ter)	1	TGG	(trp)	3
CTT	(leu)	10	CCT	(pro)	10	CAT	(his)	1	CGT	(arg)	2
CTC	(leu)	48	CCC	(pro)	29	CAC	(his)	24	CGC	(arg)	3
CTA	(leu)	60	CCA	(pro)	30	CAA	(gln)	21	CGA	(arg)	11
CTG	(leu)	26	CCG	(pro)	5	CAG	(gln)	7	CGG	(arg)	4
ATT	(ile)	21	ACT	(thr)	5	AAT	(asn)	5	AGT	(ser)	1
ATC	(ile)	50	ACC	(thr)	30	AAC	(asn)	23	AGC	(ser)	16
ATA	(met)	32	ACA	(thr)	26	AAA	(lys)	19	AGA	(ter)	0
ATG	(met)	17	ACG	(thr)	2	AAG	(lys)	2	AGG	(ter)	1
GTT	(val)	12	GCT	(ala)	15	GAT	(asp)	1	GGT	(gly)	5
GTC	(val)	17	GCC	(ala)	36	GAC	(asp)	20	GCC	(gly)	20
GTA	(val)	23	GCA	(ala)	19	GAA	(glu)	20	GGA	(gly)	11
GTG	(val)	13	GCG	(ala)	1	GAG	(glu)	8	GGG	(gly)	13

[a] Frequency of codon usage is calculated from all protein coding genes. Amino acids are indicated using the standard three-letter code

mammalian orders (Brown 1985). The cytochrome subunits (COI-III) show the highest and the ATPase8 and ND2 genes the lowest degree of identity. The genes evolve mainly by substitutions but an internal codon-size deletion occurs in the ND1 and ATPase8 duck genes while the COI gene has an extra codon at its 3′ end. Point mutations involve primarily C-T transitions and A-C transversions. Transversions slightly outnumber transitions (Table 1). The evolutionary rate is the highest at the third codon position (71%), compared with the first (20%) and second (9%), and changes involve mainly synonymous bases (Table 3). Consequently, a high proportion of base substitutions are silent. Except for the ND6 and ATPase8 genes, transitions and transversions give rise to amino acid replacements at about the same frequency. When changes to chemically similar amino acids are considered, similarities range from 86% for ND2 to 99% for COII, COIII, and ND6 (Table 1).

The GTG codon has been proposed to serve as translational initiator of the COI gene in chicken and quail (Desjardins and Morais 1990, 1991) and in cod (Johansen et al. 1990). The duck COI gene is also initiated by GTG (Fig. 1E). This unusual start codon is also found at the 5′ end of the duck COII gene (Fig. 1A).

Control Region

The control region of all vertebrate mtDNAs analyzed thus far is a noncoding sequence of variable length encompassing the heavy (H)-strand replication origin and the promoter for the transcription of both the H- and L-strands. The region spans the area between the genes for tRNA^{Pro} (tRNA^{Glu} in birds) and tRNA^{Phe} and can be divided into three subdomains: a central, more conserved segment, with a reduced L-strand adenine content, flanked on both sides by more variable adenine-rich regions, the left and right domains (Brown et al. 1986; Saccone et al. 1991). Small conserved sequences which regulate mtDNA replication, CSBs (conserved sequence blocks) and TAS (termination associated sequences), are contained within the right and left domain, respectively. They are located upstream and downstream of relatively stable cloverleaf-like secondary structures of low-primary-sequence homology among vertebrates (Dunon-Bluteau and Brun 1987).

The duck control region (Fig. 1B) shows 71.0% nucleotide identity with its chicken counterpart (Table 1). Most of the substitutions are transversions (63%) and their distribution across the control region is uneven, being more frequent in the adenine-rich left (nt 1-348) and right (nt 827-1048) domains where the transition/transversion ratio is respectively 0.51 and 0.47, and 0.75 in the central domain. Comparisons of the two avian sequences with homologous control regions from mammals, *Xenopus laevis*, and cod reveal a rather poor degree of primary sequence conservation in the right and left domains, while relatively long stretches of nucleotides dispersed across the central domain are conserved.

Duck and chicken control regions show a high frequency of length mutations. Size differences between the two avian species (179 nucleotides) are due mainly to the absence of rather large DNA segments in the duck right and left domains. Length variations of a few nucleotides were also noted in these two domains and in the central region. De-

Table 3. Sequence differences between duck and chicken protein genes

Gene	Size (codons)	Number of Differences				Substitution events	
		Position 1	Position 2	Position 3		Transition	Transversion
ND1	117	20	8	56	Silent	32.1%	33.3%
					Replacement	16.7%	17.9%
ND2	35	11	8	15	Silent	26.5%	11.8%
					Replacement	29.4%	32.3%
ND6	81	13	4	37	Silent	42.6%	27.8%
					Replacement	7.4%	22.2%
CO1	118	6	2	58	Silent	56.1%	36.4%
					Replacement	4.5%	3.0%
CO2	228	13	6	96	Silent	40.7%	43.3%
					Replacement	8.0%	8.0%
CO3	86	7	3	38	Silent	39.6%	39.6%
					Replacement	12.5%	8.3%
ATPase 8	55	14	5	24	Silent	20.0%	35.6%
					Replacement	13.3%	31.1%
ATPase 6	227	25	12	99	Silent	33.1%	39.7%
					Replacement	11.8%	15.4%
All genes	947	109	48	423	Silent	38.0%	37.1%
					Replacement	11.0%	13.9%

leted segments in the duck right domain consist of A-T-rich sequences dispersed across the region between the chicken bidirectional transcription promoter (LSP/HSP; L'Abbé et al. 1991) and CSB-1 (Fig. 3a). The pintail duck (*Anas acuta*) right domain region, which is of the same size as in Peking duck (*Anas platyrhynchos*), is devoid also of A-T-rich genomic tracts between the putatives LSP/HSP and CSB-1 (data not shown), suggesting that this feature is common to many or all birds of the genus *Anas*. Both avian CSB-1 sequences are highly homologous to their human counterpart (Fig. 3b). The 3′ end of CSB-1 in chicken (Glaus et al. 1980, Morais et al. unpublished results) and in other vertebrates lies in close proximity to the H-strand replication origin. The rather short nucleotide stretch between the duck putative transcriptional promoter 3′ end and the CSB-1 5′ end contains sequence elements that are highly homologous to human CSB-2 (Fig. 3c) and CSB-3 (Fig. 3d). These two putative conserved-sequence blocks overlap in duck (Fig. 1B), in contrast to the situation in mouse (Bibb et al. 1981) and human (Anderson et al. 1981), where they are separated from each other by short nucleotide stretches. In these mammalian species, CSB-2 and -3 are recognized by RNase RMP, an endoribonuclease which cleaves control region L-strand transcripts at specific RNA-to-DNA transition site positions in vitro (Karwan et al. 1991). Sequence elements similar to those found in duck CSB-2 and -3 are also detected in the chicken right domain but they are separated from each other by A-T-rich nucleotide stretches (Fig. 3a, double underlined). A chicken mtDNA binding protein that exhibits sequence-specific interaction with most of these sequence elements has been recently characterized (D'Agostino and Nass 1992). The protein, a site-specific endodeoxyribonuclease, interferes with mtDNA replication in vitro and may be assisted in its activity by HMG1-like protein-induced DNA bending (Wu and Crothers 1984; Bianchi et al. 1989; Fisher et al. 1992). Taken together, these observations suggest that the conserved sequence elements identified in the duck and chicken control regions serve as recognition sites for molecular complexes involved in mtDNA replication and transcription. Their high-sequence homology with mammalian CSBs makes them likely to interact with a similar set of regulatory molecules. The compact organization of the CSBs within the relatively short CSB-1/LSP/HSP region in duck provides an attractive molecular model to further investigate nucleo-mitochondrial interactions in vertebrates.

Although the duck and chicken left-domain regions are about the same size, a rather large segment that encompasses one of the two 29-bp direct repeats detected in the chicken sequence (Fig. 3e) is absent in the duck sequence. That same segment is also deleted in the Japanese quail (Desjardins and Morais 1991). The duck and chicken 29-bp monomers are well conserved and homologous to their putative CSB-1 (Fig. 3e). Repeated sequences have

```
a) CSB-1 - LSP/HSP region.

             CSB-1
C: TATTTAGTGAATGCTTGTCGGACATATTTTTATCAATTTTCACTTCCTCTATTTTCTTCAC
D: ***************C**T********C***A*------------------------AA**

C: AAAACTAGGAAATTCACCACAATTTTTCTTTGTTATTTTTAATTTTTTTTTTATTTTTT
D: *****---------****C***--------------------------------------

C: AAAAACATTTTTTAAAAAACTAAATTACATACAAACTACCGCATAAAATCCCTCAAACTAT
D: -------------------------------------**A**----*C***A***-----
                          LSP/HSP
C: ACAAACGTTTATCGTATAATATATATACAT
D: -----------------T************

b) CSB-1.

D: TATTTAGTGAATGCTCGATGGACATA
C: ***************T**C*******
H: **A***A*T******T*TA*******

c) CSB-2.

D: AAAACCACCCCAACCAC
H: C*****C****TC**C*

d) CSB-3.

D: AACCACAACCCCACAA
H: TG***-*******A**

e) Homologous sequences.

            _________29bp repeat_________
       C: CACC-TAACTATGAATGGTTACAGGACATAAATCTCACTCTCATGTTCTCCCCCCAACAA
       D: *C**A***TG*******C**GACA******C-C**AC*AACAC*CCAAATT**T*TC**C
   CSB-1:     TATTTAG******C*CGAT*******

            _________29bp repeat_________
       C: GTCACCTAACTATGAATGGTTACAGGACATACATTTAACTACCATGTTCTAACCCATTTG
       D: CC-------------------------------------------------*******AC
```

Fig. 3. **a** Comparison of the chicken sequence spanning the region between CSB-1 and LSP/HSP (Desjardins and Morais 1990; nucleotides 867–1079) to the equivalent region in duck (nucleotides 1161–1237). *Underlined* are the positions for CSB-1 and LSP/HSP. The CSB-2 and CSB-3 sequence elements are *double underlined.* **b** Comparison of the putative duck CSB-1 sequence (nucleotides 1161–1186) with those from chicken (Desjardins and Morais 1990; nucleotides 867–892) and human (Anderson et al. 1981; nucleotides 209–234). **c** Comparison of the putative duck CSB-2 sequence (nucleotides 1196–1212) with that from human (Anderson et al. 1981; nucleotides 299–315). **d** Comparison of the putative duck CSB-3 sequence (nucleotides 1208–1220) with that from human (Anderson et al. 1981; nucleotides 346–360). **e** Comparisons of the chicken sequence spanning the region between the two 29-bp repeats (Desjardins and Morais 1990; nucleotides 259–376) to an equivalent sequence in duck starting at position 603 and to the putative duck CSB-1 (nucleotides 1161–1187). Species compared: D, duck; C, chicken; H, human.

been detected in vertebrate left domains (Doda et al. 1981; Walberg and Clayton 1981; MacKay et al. 1986) and are believed to assist in termination of nascent H-strand DNAs at the end of the control region. In chicken, nascent H-strand DNAs end farther to the downstream 29-bp repeat (Glaus et al. 1980; Morais et al. unpublished results), in the vicinity of conserved TAS elements (5′—TACAT—3′), and encompass a stable cloverleaflike structure similar to those reported for various vertebrates (Dunon-Bluteau and Brun 1987). Such putative TAS elements have been identified in the duck control region (Fig. 1B).

Rate of Evolution and Divergence Times

Pairwise estimates of divergence between specific duck, chicken, and mammalian mtDNA genes are shown in Table 4. Estimates based either on total substitution events or on transversions indicate that each gene evolves at its own rate, and that rates for the ribosomal genes are lower than those for the protein genes. As expected, intraclass comparisons give estimates substantially lower than interclass comparisons, except for the COII gene, which has undergone rapid evolution in the human lineage compared to the bovine and murid lineages (Brown 1985). Interclass comparisons indicate that duck and chicken genes are about equally distant from those of mammals, and vice versa. These observations suggest that the overall rate of mtDNA evolution within the two bird orders and the three mammalian orders has been comparable since those species diverged from a common ancestor. This view is further supported by estimates derived from pairwise comparisons of amino acid sequences (Table 5).

The fossil record for birds is notoriously poor, but it is generally believed that the respective ancestor of modern ducks and chickens was distinct in the late Cretaceous to early Tertiary periods (Brodkorb 1964), suggesting that these lineages originated well into the Cretaceous (Cracraft 1986). The radiation of the major eutherian lineages seemingly occurred at about the same time (Romer 1966; Li et al. 1990; Novacek 1992). Thus, assuming that the relative mtDNA rate of evolution of bird and mammals has been similar over time, as suggested by mtDNA-calibrated clock studies of closely related geese (Shields and Wilson 1987a,b; Quinn et al.

Table 4. Pairwise divergence estimates (%) for total substitutions (above the diagonal) and transversions (below the diagonal)[a]

	16SrRNA						AtPase6						AtPase8				
	D	C	H	B	M		D	C	H	B	M		D	C	H	B	M
D	—	17	29	32	33	D	—	20	37	38	38	D	—	27	48	48	46
C	8	—	31	32	32	C	11	—	39	36	38	C	18	—	52	47	48
H	17	18	—	22	25	H	22	24	—	27	29	H	33	33	—	37	38
B	19	20	13	—	24	B	23	23	14	—	23	B	36	32	22	—	31
M	19	19	16	14	—	M	23	23	15	13	—	M	33	32	23	20	—
	12SrRNA						COII						ND6				
	D	C	H	B	M		D	C	H	B	M		D	C	H	B	M
D	—	10	24	27	27	D	—	16	32	31	31	D	—	22	44	40	42
C	4	—	23	24	23	C	8	—	31	30	30	C	11	—	42	39	43
H	10	12	—	14	13	H	18	16	—	30	28	H	28	28	—	26	28
B	13	13	5	—	12	B	17	16	14	—	22	B	23	25	13	—	28
M	12	12	6	8	—	M	18	16	15	10	—	M	27	31	20	17	—

[a] MtDNA sequences used to calculate the above values are available in the Genbank databases. Deletions and insertions were excluded from the analysis. Species compared: D, duck; C, chicken, H, human; B, bovine; M, mouse. Multialignments are given in Appendix A.

Table 5. Pairwise divergence estimates (%) for total amino acid replacements[a]

Gene	C/D	C/H	C/B	C/M	D/H	D/B	D/M	H/B	H/M	B/M
COII	6	34	32	33	34	33	33	27	28	9
ATPase6	12	46	43	46	44	41	45	23	25	22
ATPase8	28	69	69	61	59	69	61	46	56	44
ND6	17	59	54	56	57	52	53	20	33	26

[a] Protein sequences used to calculate the above values are available in the Genbank databases. Deletions and insertions were excluded from the analysis. Abbreviations of animal species are as in Table 4. Multialignments are given in Appendix B.

1991) and mammals (Brown et al. 1979; Ferris et al. 1983; Wilson et al. 1985), the level of divergence between the duck and chicken on one side, and the different mammals on the other, should be about the same. This is clearly not the case (Table 4): the duck/chicken estimates based on total substitutions or transversions are less than the corresponding mammal/mammal values by a mean factor of 1.38 and 1.48, respectively ($P < 0.01$). Estimates based on amino acid sequence homology of the mitochondrially encoded proteins also indicate a smaller genetic distance between duck and chicken than between human, mouse, and bovine (Table 5).

The overall divergence between duck and chicken sequences for protein and tRNA genes is 25%, a value which corresponds to a divergence time of about 30 Myr in mammals (Moritz et al. 1987). A similar divergence time is estimated when the duck and chicken rRNA sequences are compared. Our results suggest that, unless the Anseriform/Galliform split is drastically more recent than generally thought or factors such as differences in generation time (Li et al. 1987) and age at first breeding (Sibley et al. 1988) are under-estimated, the apparent rate of nucleotide substitutions for mtDNA in chicken and duck is somewhat decelerated relative to that of mammals of the same taxonomic levels. This has been previously suggested to account for low divergence estimates among waterfowl based on restriction-enzyme analyses of mtDNA (Kessler and Avise 1985). Anatomical and chromosomal characters, the loss of the potential for interspecific hybridization, and genetic evolution at nuclear protein-coding loci are traits which also appear to evolve slowly in birds (Prager et al. 1974; Prager and Wilson 1975, 1980; Barrowclough and Corbin 1978; Avise et al. 1980; Gutierrez et al. 1983; Patton et Avise 1986).

Very little is known about efficiency of mtDNA replication (Kunkel and Soni 1988) and repair mechanisms in birds. Moreover, birds have a significantly higher body temperature than other vertebrates, and this is likely to influence the composition of the mtDNA and protein residues as regards thermodynamic stability and activity (Avise and Aquadro 1982). In chicken, the G-C content (46%) of the complete mtDNA nucleotide sequence is 2% higher than that of any other complete mtDNA vertebrate sequence reported thus far, and that of the partial sequence for duck presented above is 4% higher. Point and length mutations occurring either in dispensable sequences (intergenic sequences, tRNA and rRNA nonpaired segments, control regions) or at codon positions where they do not cause amino acid replacements may account for the apparently similar rate of molecular evolution in closely related geese and mammalian species. As divergence times increase, constraints on nucleotide composition and amino acid replacements may be more severe in birds, such that saturation is achieved faster in birds than in mammals. It is worth noting that low rates of mtDNA evolution have been reported recently in salmonids (Thomas and Beckenbach 1989), sharks (Martin et al. 1992), and turtles (Avise et al. 1992). Evidence suggests that rates of mitochondrial and nuclear DNA evolution in those species, and in other vertebrates, is related to metabolic rate (Martin et al. 1992). If so, further molecular mechanisms underlying rates of mtDNA evolution could be operating in birds. Additional molecular data are needed to shed further light on these issues.

Acknowledgments. This work was supported by the Medical Research Council of Canada. We thank Nicole Robinson for her excellent secretarial assistance in preparing the typescript, Matthew A. Cronin for providing pintail duck mtDNA, and Robert Cedergren and Gilles Seutin for reviewing the manuscript.

References

Anderson S, Bankier AT, Barrell BG, DeBruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlick DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–464

Anderson S, DeBruijn MHL, Coulson AR, Eperon IC, Sanger F (1982) The complete sequence of bovine mitochondrial DNA: conserved features of mammalian mitochondrial genome. J Mol Biol 156:683–717

Arnason U, Gullberg A, Widegren B (1990) The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Baleonoptera physalus*. J Mol Evol 33:556–568

Arnason U, Johnsson E (1992) The complete mitochondrial DNA sequence of the harbor seal *Phoca vitulina*. J Mol Evol 34:493–505

Attardi G, Schatz G (1988) Biogenesis of mitochondria. Ann Rev Cell Biol 4:289–333

Avise JC, Patton JC, Aquadro CF (1980) Evolutionary genetics of birds. III. Conservative protein evolution in North American sparrows and relatives. Syst Zool 29:323–334

Avise JC, Aquadro CF (1982) A comparative summary of genetic distances in the vertebrates. Evol Biol 15:151–185

Avise JC, Bowen BW, Lamb T, Meylan AB, Bermingham E (1992) Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the testudines. Mol Biol Evol 9:457–473

Barrowclough GF, Corbin KW (1978) Genetic variation and differentiation in the Parulidae. Auk 95:691–702

Bianchi ME, Beltrame M, Paonessa G (1989) Specific recognition of cruciform DNA by nuclear protein HMG1. Science 243:1056–1059

Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and gene organization of mouse mitochondrial DNA. Cell 26:167–180

Brodkorb P (1964) Catalogue of fossil birds. Part 2 (Anseriformes through Galliformes). Bull Florida State Mus Biol Sci 8:195–335

Brown GG, Simpson MV (1982) Novel features of animal mtDNA evolution as shown by sequences of two rat cytochrome oxidase subunit II genes. Proc Natl Acad Sci USA 79:3246–3250

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

Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 76: 1967–1971

Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates: tempo and mode of evolution. J Mol Evol 18:225–239

Brown WM (1983) Evolution of mitochondrial DNA. In: Nei M, Koehn RK (eds) Evolution of genes and proteins. Sinauer, Sunderland, MA, pp 62–88

Brown WM (1985) The mitochondrial genome of animals. In: MacIntyre RJ (ed) Molecular evolutionary genetics. Plenum, NY, pp 95–130

Cantatore P, Saccone C (1987) Organization, structure, and evolution of mammalian mitochondrial genes. Int Rev Cytol 108: 149–208

Christianson RW, Clayton DA (1988) A tricademer DNA sequence supports human mitochondrial RNA 3′ end formation in vitro. Mol Cell Biol 8:4502–4509

Corpet F (1988) Multiple alignment with hierarchical clustering. Nucleic Acids Res 16:10881–10890

Cracraft J (1986) The origin and early diversification of birds. Paleobiology 12:383–399

D'Agostino MA, Nass MMK (1992) A new chick mitochondrial DNA-binding protein exhibits sequence-specific interaction near heavy-strand replication origin: cleavage activity, stimulation of mtDNA synthesis, and enhancement in transformed fibroblasts. Exp Cell Res 199:191–205

Dale R, McClure BA, Houchins S (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. Plasmid 13:31–40

DeSalle R, Freedman T, Prager EM, Wilson AC (1987) Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J Mol Evol 26:157–164

Desjardins P, Frost E, Morais R (1985) Ethidium bromide-induced loss of mitochondrial DNA from primary chicken embryo fibroblasts. Mol Cell Biol 5:1163–1169

Desjardins P, Morais R (1990) Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. J Mol Biol 212:599–634

Desjardins P, Ramirez V, Morais R (1990) Gene organization of the Peking duck mitochondrial genome. Curr Genet 17:515–518

Desjardins P, Morais R (1991) Nucleotide sequence and evolution of coding and noncoding regions of a quail mitochondrial genome. J Mol Evol 32:153–161

Doda JN, Wright CT, Clayton DA (1981) Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc Natl Acad Sci USA 78:6116–6120

Dunon-Bluteau D, Brun G (1986) The secondary structures of the *Xenopus laevis* and human mitochondrial small ribosomal RNA subunits are similar. FEBS Lett 198:333–338

Dunon-Bluteau DC, Brun GM (1987) Mapping at the nucleotide level of *Xenopus laevis* mitochondrial D-loop H strand: structural features of the 3′ region. Biochem Int 14:643–657

Edwards CV, Wilson AC (1990) Phylogenetically informative length polymorphism and sequence variability in mitochondrial DNA of Australian songbirds (*Pomatostomus*). Genetics 126:695–711

Ferris SD, Wilson AC, Brown WM (1981) Evolutionary tree for apes and human based on cleavage maps of mitochondrial DNA. Proc Natl Acad Sci USA 78:2432–2436

Ferris SD, Sage RD, Prager EM, Ritte U, Wilson AC (1983) Mitochondrial DNA evolution in mice. Genetics 105:681–721

Fisher RP, Lisowsky T, Parisi MA, Clayton DA (1992) DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. J Biol Chem 267: 3358–3367

Gadaleta G, Pepe G, De Candia G, Quagliariello C, Sbizà E, Saccone C (1989) The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. J Mol Evol 28:497–516

Glaus KR, Zassenhaus HP, Fechheimer NS, Perlman PS (1980) Avian mtDNA: structure, organization and evolution. In: Kroon AM, Saccone C (eds) Organization of the mitochondrial genome, vol 2. Elsevier-North-Holland Biochemical Press, New York, pp 131–135

Glotz C, Zwieb C, Brimacombe R (1981) Secondary structure of the large RNA from *Escherichia coli,* Zea mays chloroplast, and human and mouse mitochondrial ribosomes. Nucleic Acids Res 9:3287–3306

Gutierrez RJ, Zink RM, Yang SY (1983) Genic variation, systematic, and biogeographic relationships of some galliform birds. Auk 100:33–47

Hixson JE, Wong TN, Clayton DA (1986) Both the conserved stem-loop and divergent 5′-flanking sequences are required for initiation at the human mitochondrial origin of light-strand DNA replication. J Biol Chem 261:2384–2390

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome b gene of mammals. J Mol Evol 32:128–144

Johansen S, Guddal PR, Johansen T (1990) Organization of the mitochondrial genome of atlantic cod *Gadus morhua.* Nucleic Acids Res 18:411–419

Karwan R, Bennett JL, Clayton DA (1991) Nuclear RNase MRP processes RNA at multiple discrete sites: interaction with an upstream G box is required for subsequent downstream cleavages. Genes Dev 5:1264–1276

Kessler LG, Avise JC (1985) A comparative description of mitochondrial DNA differentiation in selected avian and other vertebrate genera. Mol Biol Evol 2:109–125

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Vilablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in mammals: amplification and sequencing with conserved primers. Proc Natl Acad Sci USA 86:6196–6200

Kraus F, Miyamoto M (1991) Rapid cladogenesis among the pecoran ruminants: evidence from mitochondrial DNA sequences. Syst Zool 40:117–130

Kruse B, Narasinham N, Attardi G (1989) Termination of transcription in human mitochondria: identification and purification of a DNA binding protein factor that promotes termination. Cell 58:391–397

Kunkel TA, Soni A (1988) Exonucleotic proofreading enhances the fidelity of DNA synthesis by chick embryo DNA polymerase-τ. J Biol Chem 263:4450–4459

L'Abbé D, Duhaime JF, Lang BF, Morais R (1991) The transcription of DNA in chicken mitochondria initiates from one major bidirectional promoter. J Biol Chem 266:10844–10850

Li W-H, Wolfe KH, Sourdis J, Sharp PM (1987) An evaluation of the molecular clock hypothesis using mammalian DNA sequences. J Mol Evol 25:330–342

Li W-H, Gouy M, Sharp PM, O'huigin C, Yang Y-W (1990) Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carniovora and molecular clocks. Proc Natl Acad Sic USA 87:6703–6707

Martin AP, Naylor GJP, Palumbi SR (1992) Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature 357:153–155

MacKay SL, Olivo PD, Laipis PJ, Hauswirth WW (1986) Template-directed arrest of mammalian mitochondrial DNA synthesis. Mol Cell Biol 6:1261–1267

Mindell DP, Honeycutt RL (1990) Ribosomal RNA in vertebrates: evolution and phylogenetic applications. Ann Rev Ecol Syst 21:541–566

Miyata T, Hayashida H, Kikuno R, Hasegawa M, Kobayashi M, Koike K (1982) Molecular clock of silent substitutions: at least six-fold preponderance of silent changes in mitochondrial genes over those in nuclear genes. J Mol Evol 19:28–35

Morais R, Desjardins P, Turmel C, Zinkewich-Péotti K (1988) Development and characterization of continuous avian cell lines depleted of mitochondrial DNA. In Vitro Cell Dev Biol 24:649–658

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

Novacek MJ (1992) Mammalian phylogeny: shaking the tree. Nature 356:121–125

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

Pääbo S, Thomas WK, Whitfield KM, Kumazawa Y, Wilson AC (1991) Rearrangements of mitochondrial transfer RNA genes in marsupials. J Mol Evol 33:426–430

Prager EM, Brush AH, Nolan RA, Nakanishi M, Wilson AC (1974) Slow evolution of transferrin and albumin in birds according to microcomplement fixation analysis. J Mol Evol 3:263–278

Prager EM, Wilson AC (1975) Slow evolutionary loss of the potential for interspecific hybridization in birds: a manifestation of slow regulatory evolution. Proc Natl Acad Sci USA 72: 200–204

Prager EM, Wilson AC (1980) Phylogenetic relationships and rates of evolution in birds. Proc 17th Int Ornithol Congr 17: 1209–1214

Quinn TW, Shields GF, Wilson AC (1991) Affinities of the Hawaiian goose based on two types of mitochondrial DNA data. Auk 108:585–593

Romer AS (1966) Vertebrate paleontology. University of Chicago Press, Chicago

Roe BA, Ma DP, Wilson RK, Wong JFH (1985) The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome, J Biol Chem 260:9759–9774

Saccone C, Pesole G, Sbizà E (1991) The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. J Mol Evol 33:83–91

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain termination inhibitors. Proc Natl Acad Sci USA 74: 5463–5467

Shields GF, Wilson AC (1987a) Calibration of mitochondrial DNA evolution in geese. J Mol Evol 24:212–217

Shields GF, Wilson AC (1987b) Subspecies of the Canada goose (*Branta canadensis*) have distinct mitochondrial DNAs. Evolution 41:662–666

Sibley CG, Ahlquist JE, Monroe BL (1988) A classification of the living birds of the world based on DNA-DNA hybridization studies. Auk 105:409–423

Thomas WK, Beckenbach AT (1989) Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitutions. J Mol Evol 29:233–245

Walberg MN, Clayton DA (1981) Sequence and properties of the human KB cell and mouse L-cell D-loop regions of mitochondrial DNA. Nucleic Acids Res 9:5411–5421

Wilson AC, Cann RL, Carr SM, George M, Gyllensten UB, Helm-Bychowsky KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. Biol J Linn Soc 26:375–400

Wu H-M, Crothers DM (1984) The locus of sequence-directed and protein-induced DNA bending. Nature 308:509–513

Zwieb C, Glotz C, Brimacombe R (1981) Secondary structure comparisons between small subunit ribosomal RNA molecules from six different species. Nucleic Acids Res 9:3621–3640

Appendix A.

Multialignments of nucleotide sequences. Identical residues between species are indicated by asterisks, missing nucleotides by dashes. Species compared: DUC, duck; CHI, chicken; HUM, human; BOV, bovine; MOU, mouse. Continued on pages 308–310.

16SrRNA

```
DUC  CTTACTGTGGGCCTTAAAGCAGCCTTCAACAAA.GAGTGCGTCAAAGCTCCACA.CTCAA
     *  ************ ********  ******* ***************** *  *   *
CHI  CGCACTGTGGGCCTTCAAGCAGCCACCAACAAAAGAGTGCGTCAAAGCTCC.CT.CATTA
       ** *** *****  ************* *** **  ****  ******  *   *  *
MOU  ACCATTGTAGGCCTAAAAGCAGCCACCAATAAA.GAAAGCGTTCAAGCTCAACA.TAAAA
        ********************** **** ** ********* **********  ****
BOV  CATATAGTAGGCCTAAAAGCAGCCATCAATTAA.GAAAGCGTTAAAGCTCAACAACAAAA
     ************************* ******* ********* ********** *  *
HUM  CATATAGTAGGCCTAAAAGCAGCCACCAATTAA.GAAAGCGTTCAAGCTCAACA.CCCAC

DUC  AAATGCCAAAA...CAAGATGAATCC....CTTACCACAAACAGG.TTAACCTATGA...
     ***   * ***   * *  *** ***    ** * *  ** **** ***********
CHI  AAAAATCTAAAACCCTATTTGACTCC....CTCAACCAAAGCAGG.TTAACCTATGA...
         *  ***  ** ** ** ** **     **    *    *   *      **
MOU  TTTCAATTAATTCCATAATTTACACCA.ACTTCCTAAACTTAAAATTGGGTTAATCTATA
      ** ***  ******  *   *      *  *****  *  ** * ***  ********
BOV  ATTAAATAGATTCCAACAACAAATGATTAACTCCTAGCCCCAATACTGGACTAATCTATT
          * * * *** * *   *    * *******  * ** *   ***** *******
HUM  TACCTAAAAAATCCCAAACATATAACTGAACTCCTCACACCCAAT.TGGACCAATCTATC

DUC  ...ATATAGGAGAATTAATGCTAAAATGAGTAACTTGGGGCCAC.ACCCACCCCTCTAGC
          **** ***** *****************  ***  **   * ** **       *
CHI  ...CAATAGAAGAATCAATGCTAAAATGAGTAATCTGGAACCT..ATCCTCC.....TAC
          ***** * * ** ** **  ********   * *  *   ** ****     ***
MOU  ACTTTATAGATGCAACACTGTTAGTATGAGTAACAAGAATTCCA.ATTCTCCAGACATAC
     *    ***** **** * ***** ***************       ******   ****
BOV  ATAGAATAGAAGCAATAATGTTAATATGAGTAACAAGAAAAATT..TTCTCCTTGCATAA
     *    ******* * ******** *** ******* *****  *  * ****     ***
HUM  ACCCTATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCCTCCGCA..TAA

DUC  GGCGCAAGCTTACATGAGAACATTATTAACAGACCCAGACA.TATACAAAAAC...TCCT
     **** ** ******* *  ************** *   **  *****    **    *
CHI  GGCGTAAACTTACATTAATACATTATTAACAGAACTCAACT.TATACCCCCAC...ACTA
     *     ***       *  **  *  * *   **    *** **  *    *    ****
MOU  GCGTATAACAACTCGGATAACCATTGTTAGTTAATCAGACTATAGGCAATAATCACACTA
     *  **   **   *     *  * * * *            *     ******    * *
BOV  GTCTAAGTCAGTGCCTGATAATACTCTGACCACTAACAGTCAATAAAAATAAT...CCAA
     * **  *****      * * *   ***** * ****** * *  *     *     ***
HUM  GCCTGCGTCAGATTAAAACACTGAACTGACAATTAACAGCCCAATATCTACAA...TCAA

DUC  ACAAGACCAGGTATAAACTCACCCTGTTAACCCGACTCAGGAGCGCCCATA.AGAGAGAT
     *****            ****  ******* ** ** ************ *  **  ***
CHI  ACAAGCAATACGTATTCCTCAATCTGTTAAGCCAACCCAGGAGCGCCCACA.GGA.TGAT
       **  *        *      ******** ***** * ***  ** *  * ***  ***
MOU  TAAATAATCCACCTATAACTTCTCTGTTAACCCAACACCGGAATGC.CTAAAGGAAAGAT
      **  **        * * *   *************** *** ***    **********
BOV  CAATAAACAATTTATTGATTATACTGTTAACCCAACACAGGAGTGCATCTAAGGAAAGAT
     * *  **    *****      ***** *************  ***   ********* *
HUM  CCAACAAGTCATTATTACCCTCACTGTCAACCCAACACAGGCATGCTCATAAGGAAAGGT

DUC  TAAAATCTGTGAAAGGAACTCGGCAAAACAAGG.CCCGACTGTTTACCAAAAACATAGCC
     ***** **    *************** *** * *********** **************
CHI  TAAAACCTACAGAAGGAACTCGGCAAACCAAAGACCCGACTGTTTCCCAAAAACATAGCC
       ***   * * **************** * *  **** ****** **********  **
MOU  CCAAAAAGATAAAAGGAACTCGGCAAACAAGAACCCCGCCTGTTTACCAAAAACATCACC
       *** *  ********************  *****************************
BOV  TAAAAGAAGTAAAAGGAACTCGGCAAACACAAACCCCGCCTGTTTACCAAAAACATCACC
     ***** *********************    *  **************************
HUM  TAAAAAAAGTAAAAGGAACTCGGCAAATCTTAC.CCCGCCTGTTTACCAAAAACATCACC

DUC  TTCAGCAAACAAACAAGTATTGAAGGTGATGCCTGCCCAGTGACC...TAGGTTAAACGG
     ****** ***** ********************************    * *** *****
CHI  TTCAGCTAACAA.CAAGTATTGAAGGTGATGCCTGCCCAGTGACCCCCAAAGTTCAACGG
     *  ***       ********  ***   ***************    ****** *****
MOU  TCTAGCATTA...CAAGTATTAGAGGCACTGCCTGCCCAGTGACT...AAAGTTTAACGG
     ** ******    * ****** ****** ***************    *  *********
BOV  TCCAGCATTC...CCAGTATTGGAGGCATTGCCTGCCCAGTGACA...ACTGTTTAACGG
     ** ***  **   ******** ******  ***************     **********
HUM  TCTAGC.ATCA..CCAGTATTAGAGGCACCGCCTGCCCAGTGACA...CATGTTTAACGG

DUC  CCGCGGTATCCTAACCGTGCAAAGGTAGCGCAATCAATTGTCCCATAAATCGAGACTTGT
     ******************** *********************** ***** *********
CHI  CCGCGGTATCCTAACCGTGCGAAGGTAGCGCAATCAATTGTCCCGTAAATTGAGACTTGT
     ************ ******* ********  ***** **** ** *** * * **** *
MOU  CCGCGGTATCCTGACCGTGCAAAGGTAGCATAATCACTTGTTCCTTAATTAGGGACTAGC
     ************************************ ******  *** ** ***** *
BOV  CCGCGGTATCCTGACCGTGCAAAGGTAGCATAATCATTTGTTCTCTAAATAAGGACTTGT
     ******** *** *********************** ******  ****** **** ***
HUM  CCGCGGTACCCTAACCGTGCAAAGGTAGCATAATCACTTGTTCCTTAAATAGGGACCTGT

DUC  ATGAATGGCTAAACGAGGTCTTAACTGTCTCTCACGGATAATCAGTGAAATTGATCTCCC
     *******************************     * *****  *******  * * **
CHI  ATGAATGGCTAAACGAGGTCTTAACTGTCTCCTGTAGGTAATCTATGAAATTAGTATTCC
     ***** ************    ********* * *   *****  *******    ** *
MOU  ATGAACGGCTAAACGAGGGTCCAACTGTCTCTTATCTTTAATCAGTGAAATTGACCTTTC
     ***** ***   ********   ***********  *  ******************* *
BOV  ATGAATGGCCGCACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTTCC
     *********  **********   *************  ** *************** **
HUM  ATGAATGGCTCCACGAGGGTTCAGCTGTCTCTTACTTTTAACCAGTGAAATTGACCTGCC

DUC  CGTGCAAAAGCGGGA.TGTGAACATAAGACGAGAAGACCCTGTGGAACTTAAAAATCAAC
     ********* ** ** ********************************** ******* *
CHI  CGTGCAAAAACGAGAATGTGAACATAAGACGAGAAGACCCTGTGGAACTTTAAAATCA.C
      *** * *  *   *** * *  ****************** **** *** **    *
MOU  AGTGAAGAGGCTGAAATATAATAATAAGACGAGAAGACCCTATGGAGCTTAAATTATA..
      ********** * ***  *  **************************** ** *
BOV  CGTGAAGAGGCGGGAATGCACAAATAAGACGAGAAGACCCTATGGAGCTTTAACT.....
     ************** **     *  **************************** *
HUM  CGTGAAGAGGCGGGCATAACACAGCAAGACGAGAAGACCCTATGGAGCTTTAATTTATTA

DUC  GGCCACCGCG.AACCTAAGACTAAACCCACCGGG.CTACAGACATCGCAGAGCATGGCCG
     * *****    ***** * ** *  ***** *** * **  ***    *   * *** **
CHI  GACCACCTTACAACCTTACAC.AGCCCCACTGGGTCCACCCACACATAAACCCCTGGTCG
      * *   *        *  *  *  ** * ***     **** * *  *     *
MOU  TAACTTATCTATTTAATTTATTAAACCTAATGG.....CCCAAAAACTATAGTATAAGTT
      * *    * *    * *   *  * * * *          * **        ** ****
BOV  AACCAACCCAAAGAGAATAGATTTAACCATTAAG..GAATAACAACAATCTCCATGAGTT
     *  ***  **      *       * ***            *    *     * **  **
HUM  ATGCAA.ACAGTACCTAACA...AACCCACAGG....TCCTAAACTACCAAACCTGCATT

DUC  ATATTTTTCGGTTGGGGCGACCTTGGAGAACAACAGATCCTCCAAAAACA.AGACCACAC
     * **************************** ** * **********  ** ********
CHI  ACATTTTTCGGTTGGGGCGACCTTGGAGAAAAAAAAATCCTCCAAACCCACAGACCACAA
       *  ************ ***** ****** ************ **     * * * * *
MOU  TGAAATTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAATGATTATAAC.CTA
      * * ****************************************  **** **     *
BOV  GGTAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAGCGATTTTA....AA
        * ************ ********** * **   * *********  *        **
HUM  AAAAATTTCGGTTGGGGCGACCTCGGAGCAGAACCCAACCTCCGAGCAGTACATGCT.AA

DUC  CTCTTTACTTAGAGCCACCCCTCAAAGTGCTAATAGTGACCAGACCCAATATAATT..GA
     ***** *** *** * ** ********* * ** *** ******************  **
CHI  CTCTTCACTAAGACCAACTCCTCAAAGTACCAACAGTAACCAGACCCAATATAATT..GA
       *** **         *     *** *    * *       ******     ***  **
MOU  GACTT.AC.....AAGTCAAAGTAAAATCAACATATCTTATTGACCCAGATATATTTTGA
     ****  **         ***   ***     *      *  *    *  * *  **  **
BOV  GACTAGAC......CCACAAGTCAAATCACTCTATCGCTCATTGATCCAAAAACTT..GA
     ****  **         ***  * **  *** **              ** *****  **
HUM  GACTTCAC.....CAGTCAAAGCGAACTACTATACTCAATTGATCC..AATAACTT..GA

DUC  TTAATGGACCAAGCTACCCCAGGGATAACAGCGCAATCCCCCTCAAGAGCCCCTATCGAC
       ************************************ **  ********* *******
CHI  GCAATGGACCAAGCTACCCCAGGGATAACAGCGCAATCTCCTCCAAGAGCCCATATCGAC
      *** ******** ***** ******************   *  *****  *********
MOU  TCAACGGACCAAGTTACCCTAGGGATAACAGCGCAATCCTATTTAAGAGTTCATATCGAC
     ******** ********************************** ****** *********
BOV  TCAACGGAACAAGTTACCCTAGGGATAACAGCGCAATCCTATTCAAGAGTCCATATCGAC
      ******************************************* ************ **
HUM  CCAACGGAACAAGTTACCCTAGGGATAACAGCGCAATCCTATTCTAGAGTCCATATCAAC

DUC  AGG.GGGGTTTACGACCTCGATGTTGGATCAGGACATCCTAATGGTGCAGCCGCTATTAA
     * * * ****************************** ************ **********
CHI  AAG.GAGGTTTACGACCTCGATGTTGGATCAGGACAACCTAATGGTGCAACCGCTATTAA
     **    ****************************** ** ******* *   ********
MOU  AATTAGGGTTTACGACCTCGATGTTGGATCAGGACATCCCAATGGTGTAGAAGCTATTAA
     *** ***********************************  ****** *   ***** **
BOV  AAT.AGGGTTTACGACCTCGATGTTGGATCAGGACATCCTGATGGTGCAACCGCTATCAA
     *** *********************************** ********* ******* **
HUM  AAT.AGGGTTTACGACCTCGATGTTGGATCAGGACATCCCGATGGTGCAGCCGCTATTAA

DUC  GGGTTCGTTTGTTCAACGATTAATAGTCCTACGTGATCTGAGTTCAGACCGGAGCAATCC
     *********************** ************************************
CHI  GGGTTCGTTTGTTCAACGATTAACAGTCCTACGTGATCTGAGTTCAGACCGGAGCAATCC
      ********************** ************************************
MOU  TGGTTCGTTTGTTCAACGATTAA.AGTCCTACGTGATCTGAGTTCAGACCGGAGCAATCC
      ********************** ****************************** *****
BOV  AGGTTCGTTTGTTCAACGATTAA.AGTCCTACGTGATCTGAGTTCAGACCGGAGTAATCC
     *********************** ************************************
HUM  AGGTTCGTTTGTTCAACGATTAA.AGTCCTACGTGATCTGAGTTCAGACCGGAGTAATCC

DUC  AGGTCGGTTTCTATCTATGAACTACT.CTCCCCAGTACGAAAGGACCGGGAAAGTAAGGC
     ******************* ** *** *  ** ****************  ****  **
CHI  AGGTCGGTTTCTATCTATGGAC.ACT.CCTCCTAGTACGAAAGGACCGGAGAAGTGGGGT
     ******************  ** * *  **** *************  ***** * * *
MOU  AGGTCGGTTTCTATCTATTTACGATT.TCTCCCAGTACGAAAGGACAAGAGAAATAGAGC
     ****************** ****  * *****************************  **
BOV  AGGTCGGTTTCTATCTAT.TACGTATTTCTCCCAGTACGAAAGGACAAGAGAAATAAGGC
     *****************  * *  *** ***** **************************
HUM  AGGTCGGTTTCTATCTACCTTCAAATTCCTCCCTGTACGAAAGGACAAGAGAAATAAGGC

DUC  CAATACTACAAG...CACGC.CTTCCCTCTAA
     ****** **      *** * *  ** *****
CHI  CAATACCACTGAGCACACCC.CAACCTTCTAA
     **             * * * **** *  *
MOU  CACCTTACAAATAAGCGCTCTCAACTTAATTT
     ** ***  **  **  ** * *     *
BOV  CAACTTTAAATCAA..GCGC.CTTAAGACAAC
     * **** * *      **** ***      *
HUM  CTACTTCACAAA....GCGC.CTTCCCCGTA
```

12SrRNA

```
DUC  AACCCACGAAAGCCAGGGCCCAAACTGGGATTAGATACCCCACTATGCCTGGCCCTAAAT
     ************* *** ******************************** *********
CHI  AACCCACGAAAGCTAGGACCCAAACTGGGATTAGATACCCCACTATGCCTAGCCCTAAAT
     *** *** * ***** ******************************** **********
HUM  AACACACAATAGCTAAGACCCAAACTGGGATTAGATACCCCACTATGCTTAGCCCTAAAC
      ** *** ****************************************************
BOV  GACGCACTATAGCTAAGACCCAAACTGGGATTAGATACCCCACTATGCTTAGCCCTAAAC
      *  *** * ******************************************** *****
MOU  AATACACGACAGCTAAGACCCAAACTGGGATTAGATACCCCACTATGCTTAGCCATAAAC

DUC  CTAAATAATTTACCC.TACCGAAGTATCCGCCAGAGAACTACGAGCACAAACGCTTAAAA
       * ******  *    **  * ********* ***************************
CHI  ACAGATAATTCCCAT.CACACATGTATCCGCCTGAGAACTACGAGCACAAACGCTTAAAA
         * * **       *** *  *   **** **  ********* ** * ********
HUM  CTCAACAGTTAAATC.AACAAAACTGCTCGCCAGAACACTACGAGC.CACA.GCTTAAAA
       **  * *** **  ******* *  ********  ***** ***  *   ********
BOV  CTAGATACCTACATA.AACAAAATTATTCGCCAGAGTACTACTAGC..AACAGCTTAAAA
     ** ***** ** **  ******* **** ******* *********  *   ********
MOU  CTTGATACATAAATTTAACAAAACTATTTGCCAGAGAACTACTAGC.CATA.GCTTAAAA

DUC  CTCTAAGGACTTGGCGGTGCCCTAAACCCACCTAGAGGAGCCTGTTCTGTAATCGATGAT
     ********************** ************************* ******** **
CHI  CTCTAAGGACTTGGCGGTGCCCCAAACCCACCTAGAGGAGCCTGTTCTATAATCGATAAT
     *** ****** *********  ** * **  ***************** **********
HUM  CTCAAAGGACCTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTGTAATCGATAAA
     ********** *********** ****** ****************** ***********
BOV  CTCAAAGGACTTGGCGGTGCTTTATATCCTTCTAGAGGAGCCTGTTCTATAATCGATAAA
     ****************** ********** ******************************
MOU  CTCAAAGGACTTGGCGGTACTTTATATCCATCTAGAGGAGCCTGTTCTATAATCGATAAA

DUC  CCACGATCAACCCAACCGCCCCTTGCCAAGCACAGCCTACATACCGCCGTCGCCAGCCCA
     *******  ******** ********** *******************************
CHI  CCACGATTCACCCAACCACCCCTTGCCA.GCACAGCCTACATACCGCCGTCGCCAGCCCA
     ** ****  ***  ****** *****      ******* ******** **  ****  *
HUM  CCCCGATCAACCTCACCACCTCTTGCT.....CAGCCTATATACCGCCATCTTCAGCAAA
     ******* **********  *******     *** ************************
BOV  CCCCGATAAACCTCACCAATTCTTGCTA.ATACAGTCTATATACCGCCATCTTCAGCAAA
     ***** *  *********  ******** ** *** ************************
MOU  CCCCGCTCTACCTCACCATCTCTTGCTA.ATTCAGCCTATATACCGCCATCTTCAGCAAA

DUC  CCTCGAATGAGAGCGCAACAGTG.GCGCAACAGCACCCCGCTAATAAGACAGGTCAAGGT
     ************************* ***** **  ******** ** *** *** **
CHI  CCTCTAATGAAAGAACAACAGTGAGCTCAATAGCCCCTCGCTAATAAGACAGGTCAAGGT
     ** ** ***** *    * *** *** ***   * ****  ** * *  **********
HUM  CC.CTGATGAAGGCTACAAAGTAAGCGCAAGTAC.CCACGTAAAGACGTTAGGTCAAGGT
     ** ** *  ****  * ********** ** **   * * **** ***************
BOV  CC.CTAA.AAAGGAAAAAAAGTAAGCGTAATTATGATACATAAAAACGTTAGGTCAAGGT
     ** **** * *****  *********  **  ** ************************
MOU  CC.CTAA.AAAGGTATTAAAGTAAGCAAAAGAATCAAACATAAAAACGTTAGGTCAAGGT

DUC  ATAGCCTATGGGACGG.AAGAAATGGGCTACATTCCCTATGCATAGGGCA
     ************  **  ****************  ***  *****  **
CHI  ATAGCCTATGGGGTGG.GAGAAATGGGCTACATTTTCTAA.CATAGAACA
      ***** *** *****  *********************  *  *****
HUM  GTAGCCCATGAGGTGGCAAGAAATGGGCTACATTTTCTAC.CCCAGAAAA
     *** ** ****  *** ***************** ***** ** *** **
BOV  GTAACCTATGAAATGGGAAGAAATGGGCTACATTCTCTACACCAAGAGAA
     *** ** ************************** ***     **** *
MOU  GTAGCCAATGAAATGGGAAGAAATGGGCTACATTTTCTTATAAAAGAACA
```

CO II

```
DUC  GTGGCCAACCACTCCCAACTAGGATTCCAAGACGCCTCATCACCCATTATAGAAGAGCTC
      ********************** ** ************** ***** ************
CHI  ATGGCCAACCACTCCCAACTAGGCTTTCAAGACGCCTCATCCCCCATCATAGAAGAGCTC
     ****** ***  * ****** **  * ********* ******* ** ***********
MOU  ATGGCCTACCCATTCCAACTTGGTCTACAAGACGCCACATCCCCTATTATAGAAGAGCTA
     ***** ** **  * ***** **  * ***** ** ***** ** ** ******** ***
BOV  ATGGCATATCCCATACAACTAGGATTCCAAGATGCAACATCACCAATCATAGAAGAACTA
     ****** ** *    *** ****  * ***** ** ** ** ** *********** **
HUM  ATGGCACATGCAGCGCAAGTAGGTCTACAAGACGCTACTTCCCCTATCATAGAAGAGCTT

DUC  GTTGAATTCCACGACCACGCTCTGATTGTTGCCTTAGCTATCTGCAGCCTAGTCCTATAC
     ******************** ***** ** **  **** ** ****** **** ** ***
CHI  GTTGAATTCCACGACCACGCCCTGATAGTCGCACTAGCAATTTGCAGCTTAGTACTCTAC
      *  * ***** ** *** * ** *** * *   *   **** **  ****** *****
MOU  ATAAATTTCCATGATCACACACTAATAATTGTTTTCCTAATTAGCTCCTTAGTCCTCTAT
      *  * ** ***** ***** *********** *** ********** ***** ** **
BOV  CTTCACTTTCATGACCACACGCTAATAATTGTCTTCTTAATTAGCTCATTAGTACTTTAC
      *   ********* *** * ** *****  * *** * **  ***   **** ** **
HUM  ATCACCTTTCATGATCACGCCCTCATAATCATTTTCCTTATCTGCTTCCTAGTCCTGTAT

DUC  CTCTTAGCCCACATGCTAATAGAAAAACTATCA...TCCAACGCAGTAGACGCCCAAGAA
     **  ** * *  ** ** ***************    * *** * ***** *********
CHI  CTTCTAACTCTTATACTTATAGAAAAACTATCA...TCAAACACCGTAGATGCCCAAGAA
      *  *  * ** *** * * *  ******* **    *** ***  ******* ******
MOU  ATCATCTCGCTAATATTAACAACAAAACTAACACATACAAGCACAATAGATGCACAAGAA
     ** ** ** ****** **** ***** ** ** *********** ***************
BOV  ATTATTTCACTAATACTAACGACAAAGCTGACCCATACAAGCACGATAGATGCACAAGAA
         ***  **** *** ** ***** ** **  **** * **    *** ** ** ***
HUM  GCCCTTTTCCTAACACTCACAACAAAACTAACTAATACTAACATCTCAGACGCTCAGGAA

DUC  GTAGAACTAATCTGAACAATCCTACCCGCCATCGTCCTAGTACTCCTCGCCCTCCCATCC
     ** ************** *********** ** ******** ** ** ******** ***
CHI  GTTGAACTAATCTGAACCATCCTACCCGCTATTGTCCTAGTCCTGCTTGCCCTCCCCTCC
     ******   ** ***** ** ***** *** *  ****  ** *  **** ******** 
MOU  GTTGAAACCATTTGAACTATTCTACCAGCTGTAATCCTTATCATAATTGCTCTCCCCTCT
     ** ** ** ** ***** ***** ** **  * *** * **  ********** ** ***
BOV  GTAGAGACAATCTGAACCATTCTGCCCGCCATCATCTTAATTCTAATTGCTCTTCCTTCT
      **** **  ******* ** *************** ** * ** ** ** ** ** ** 
HUM  ATAGAAACCGTCTGAACTATCCTGCCCGCCATCATCCTAGTCCTCATCGCCCTCCCATCC

DUC  CTACAAATCCTGTACATAATAGACGAAATCGACGAGCCAGACCTCACACTAAAAGCCATT
     ** ******** *********************** ** ** ***** *********** 
CHI  CTCCAAATCCTCTACATAATAGACGAAATCGACGAACCTGATCTCACCCTAAAAGCCATC
     ** *  ** ** ** *************** ** * ** *   * *** * *** **** 
MOU  CTACGCATTCTATATATAATAGACGAAATCAACAACCCCGTATTAACCGTTAAAACCATA
      **** ******** ******** ******** *****     * ** ** *********
BOV  TTACGAATTCTATACATAATAGATGAAATCAATAACCCATCTCTTACAGTAAAAACCATA
      **** ** ** ******* *** **  ****  * ** ** *****  * *** * ** 
HUM  CTACGCATCCTTTACATAACAGACGAGGTCAACGATCCCTCCCTTACCATCAAATCAATT

DUC  GGCCACCAGTGATACTGAAGCTACGAATACACAGACTTCAAGGACCTCTCATTCGACTCC
     ** ***** ********** *** ***************************** ******
CHI  GGACACCAATGATACTGAACCTATGAATACACAGACTTCAAGGACCTCTCATTTGACTCC
     ** **************** *** ***** ** ****   * ***** *  ***** ** 
MOU  GGGCACCAATGATACTGAAGCTACGAATATACTGACTATGAAGACCTATGCTTTGATTCA
     ** ** ** ************** ** ***** ** ***** *** ** **** ** ** 
BOV  GGACATCAGTGATACTGAAGCTATGAGTATACAGATTATGAGGACTTAAGCTTCGACTCC
     ** ** ** ** ******* *** ***** ** ** ** *  *   *** **** *****
HUM  GGCCACCAATGGTACTGAACCTACGAGTACACCGACTACGGCGGACTAATCTTCAACTCC

DUC  TACATAATTCCCACCACAGACCTGCCAAATGGGCACTTCCGACTCCTAGAAGTTGACCAC
     *******  ** ** ******** **    ** ******** ** ******** ***** 
CHI  TACATAACCCCAACAACAGACCTCCCCCTAGGCCACTTCCGCCTACTAGAAGTCGACCAT
     ** **** ********  *****    *  **  *** *** ** ******** **  * 
MOU  TATATAATCCCAACAAACGACCTAAAACCTGGTGAACTACGACTGCTAGAAGTTGATAAC
     ** ***** ******   **  **** ** ** ** ********  ******* ***** 
BOV  TACATAATTCCAACATCAGAATTAAAGCCAGGGGAGCTACGACTATTAGAAGTCGATAAT
     ****** ****  *** *    ** * ***** ** ** *****  * ** ** ** ***
HUM  TACATACTTCCCCCATTATTCCTAGAACCAGGCGACCTGCGACTCCTTGACGTTGACAAT

DUC  CGCGTAGTCGTACCCATAGAATCACCGATCCGCGTAATTATTACTGCCGGAGACGTACTT
     *** * **  * *********** ** ** ** ***** ** ** ** *  ***** ** 
CHI  CGCATTGTAATCCCCATAGAATCCCCCATTCGAGTAATCATCACCGCTGATGACGTCCTC
     **  * **  * ** ******   ** ** **  ** * **  *  **** *********
MOU  CGAGTCGTTCTGCCAATAGAACTTCCAATCCGTATATTAATTTCATCTGAAGACGTCCTC
     ***** ** ** ********* *  ******* ** *** * ** ***********  * 
BOV  CGAGTTGTACTACCAATAGAAATAACAATCCGAATGTTAGTCTCCTCTGAAGACGTATTA
     ***** ***** ** ** ***    * ** ** ** **  ** *  * **  ******* ** 
HUM  CGAGTAGTACTCCCGATTGAAGCCCCCATTCGTATAATAATTACATCACAAGACGTCTTG

DUC  CACTCATGAGCAGTTCCAACGCTCGGAGTTAAAACAGATGCAATCCCAGGCCGACTAAAC
     *********** ** **  * ***** ** ******** ******** ** ******** 
CHI  CACTCATGAGCCGTACCCGCCCTCGGGGTAAAAACAGACGCAATCCCTGGACGACTAAAT
     *********** ** *** **** **  * ***** ** ** ***** ** *********
MOU  CACTCATGAGCAGTCCCCTCCCTAGGACTTAAAACTGATGCCATCCCAGGCCGACTAAAT
     *********** ** ***** ******** ***** ** ** *********** ***** 
BOV  CACTCATGAGCTGTGCCCTCTCTAGGACTAAAAACAGACGCAATCCCAGGCCGTCTAAAC
     ************** *** *  ****  ********** ***** ** ** *********
HUM  CACTCATGAGCTGTCCCCACATTAGGCTTAAAAACAGATGCAATTCCCGGACGTCTAAAC

DUC  CAAACCTCATTCATTACCACCCGGCCTGGGATTTTCTACGGCCAGTGCTCAGAAATCTGC
     ******** ***** ***** ** ** **  * ** ***** ** ***************
CHI  CAAACCTCCTTCATCACCACTCGACCAGGAGTGTTTTACGGACAATGCTCAGAAATCTGC
     *** *  *  * * * * *  ********  * ** ** ** ******** ***** ** 
MOU  CAAGCAACAGTAACATCAAACCGACCAGGGTTATTCTATGGCCAATGCTCTGAAATTTGT
     *** ****  * * ***   *** ***** ****  ** ** ******** ******** 
BOV  CAAACAACCCTTATATCGTCCCGTCCAGGCTTATATTACGGTCAATGCTCAGAAATTTGC
     ***** **  * *   *  * ** ** **  **** ************** ***** ** 
HUM  CAAACCACTTTCACCGCTACACGACCGGGGGTATACTACGGTCAATGCTCTGAAATCTGT

DUC  GGGGCTAACCACAGCTACATGCCTATTGTAGTAGAATCTACCCCACTCCCATACTTTGAA
     ** ***************** ** *********** ******** **   * ********
CHI  GGAGCTAACCACAGCTACATACCCATTGTAGTAGAGTCTACCCCCCTAAAACACTTTGAA
     *** ******* ****  ** ********  ****       ** ****** * ** ***
MOU  GGATCTAACCATAGCTTTATGCCCATTGTCCTAGAAATGGTTCCACTAAAATATTTCGAA
     ** ** ***** ** ** ** *********** **  * ** ******** ** ** ***
BOV  GGGTCAAACCACAGTTTCATACCCATTGTCCTTGAGTTAGTCCCACTAAAGTACTTTGAA
     **  **************** ***** ***** ** *** * ** *****   *******
HUM  GGAGCAAACCACAGTTTCATGCCCATCGTCCTAGAATTAATTCCCCTAAAAATCTTTGAA

DUC  GCCTGATCATCCCTCCTATCGTCATCCTAA
     ******** ** ** ** ** **    ***
CHI  GCCTGATCCTCACTACTGTCATCT...TAA
       ******            *  *   ***
MOU  AACTGATCT...GCTTCAATAATT...TAA
     ** ******   ** ****** *    ***
BOV  AAATGATCT...GCGTCAATATTA...TAA
     * * *  *    *  *  *   **   ** 
HUM  ATAGGGCCC...GTATTTACCCTA...TAG
```

ATPase6

```
DUC  ATGAACCTAAGTTTCTTTGACCAATTCTCAAGCCCCCACCTACTTGGTCATCCCCTGATC
     *********** ***** ******************  ****** **    ** ** ***
CHI  ATGAACCTAAGCTTCTTCGACCAATTCTCAAGCCCCTGCCTACTAGGAATCCCTCTAATC
     ******  **   * ** * *  ****   * ***     ** ***** ****  *  * 
MOU  ATGAACGAAAATCTATTTGCCTCATTCATTACCCCAACAATAATAGGATTCCCAATCGTT
     ************ ***** **** ** ********   ***  ****  ****  **** 
BOV  ATGAACGAAAATTTATTTACCTCTTTTATTACCCCTGTAATTTTAGGTCTCCCTCTCGTA
     ************ * **  * ** *  *** ****   ***  **** ** **   ** *
HUM  ATGAACGAAAATCTGTTCGCTTCATTCATTGCCCCCACAATCCTAGGCCTACCCGCCGCA

DUC  CTACTATCCCTGCTTCTTCCAGCCCTATTGTTCCCATCCCCAGGCAACCGATGAATCAAC
     ** * *** ** **************  *  * ***** ***** ******** ******
CHI  CTCCCATCACTCCTTCTTCCAGCCCTCCTACTTCCATCACCAGGAAACCGATGGATCAAC
      *  *     *  *  ****  *  ***** * *****  **   ** **    ******
MOU  GTAGCCATCATTATATTTCCTTCAATCCTATTCCCATCCTCA...AAACGCCTAATCAAC
           *** *  **** **     * ********* * ***   ** ** *** * * *
BOV  ACCCTTATCGTACTATTCCCAAGCCTACTATTCCCAACATCA...AACCGACTAGTAAGC
        ** *** * ***** **    *** *  **** ** **    **    **  * * *
HUM  GTACTGATCATTCTATTTCCCCCTCTATTGATCCCCACCTCC...AAATATCTCATCAAC

DUC  AACCGACTATCCACCATCCAACTGTGACTCCTACACCTAATCACAAAACAACTAATAATC
     ***** ** ************** *** **   ************************* *
CHI  AACCGCCTCTCCACCATCCAACTCTGATTCACCCACCTAATCACAAAACAACTAATAACC
     ***** ***    *  ****** **** *     ** * ** *  ****** ****    
MOU  AACCGTCTCCATTCTTTCCAACACTGACTAGTTAAACTTATTATCAAACAAATAATGCTA
     ** **  *     *  ******* *** ** ** ***** *    ************   
BOV  AATCGCTTTGTAACCCTCCAACAATGAATACTTCAACTTGTATCAAAACAAATAATGAGT
     ** **  *  * ***  ********** ** *  ****    *********** ** *  
HUM  AACCGACTAATCACCACCCAACAATGACTAATCAAACTAACCTCAAAACAAATGATAACC

DUC  CCATTAAACAAAAACGGCCACAAATGAGCCCTGATGCTAACATCACTAATAACCATACTC
     **  *******    ** **************  * ** ** ***** **    ******
CHI  CCCCTAAACAAGGCAGGTCACAAATGAGCCCTCCTACTCACCTCACTTATCCTAATACTC
       **  * *     *** *  * **** ****  ** *    ** ** *** ** *  * 
MOU  ATCCACACACCAAAAGGACGAACATGAACCCTAATAATTGTTTCCCTAATCATATTTATT
     *******   * ******* *********  ***** *  * ** ** *** ********
BOV  ATCCACAATTCTAAAGGACAAACATGAACATTAATATTAATATCTCTGATCCTATTTATT
     ** *****  ********* *** *** *  * *** ** ****  * *** * ******
HUM  ATACACAACACTAAAGGACGAACCTGATCTCTTATACTAGTATCCTTAATCATTTTTATT

DUC  CTAACAATCAACCTTCTAGGACTTCTCCCATATACATTCACCCCAACCACCCAGCTATCC
     **  * ** ***** ***** ** ***** ** ** *********** ***** ***** 
CHI  CTCTCCATTAACCTCCTAGGCCTCCTCCCCTACACCTTCACCCCAACTACCCAACTATCA
        ** *  ** ***********  * **  * ** ** ** ** ************** 
MOU  GGATCAACAAATCTCCTAGGCCTTTTACCACATACATTTACACCTACTACCCAACTATCC
     *********** ** ******** ***** *** **** ***** ** ** ******** 
BOV  GGATCAACAAACCTACTAGGCCTATTACCCCATTCATTCACACCAACAACACAACTATCA
     *   **** ***** ** ** **  * ** ** ***** ******** ** ******** 
HUM  GCCACAACTAACCTCCTCGGACTCCTGCCTCACTCATTTACACCAACCACCCAACTATCT

DUC  ATAAACATGGCCCTAGCTTTCCCCCTGTGGCTTGCTACCCTACTAACAGGCCTGCGAAAC
     ************ ****  * ** ** ** ** ** *** ********************
CHI  ATAAACATGGCCTTAGCCCTGCCACTATGACTAGCCACCTTACTAACAGGCCTGCGAAAC
     *****  *    * **** * *********   *   ** ** * ****** * *** **
MOU  ATAAATCTAAGTATAGCCATTCCACTATGAGCTGGAGCCGTAATTACAGGCTTCCGACAC
     ***** *** * ******** ** ** ***** ***************** *****  * 
BOV  ATAAACCTAGGCATAGCCATCCCCCTGTGAGCAGGAGCCGTAATTACAGGATTCCGCAAT
     ********** *** ********* * ***** **  * ** **** *** ** ***  *
HUM  ATAAACCTAGCCATGGCCATCCCCTTATGAGCGGGCACAGTGATTATAGGCTTTCGCTCT

DUC  AAACCATCAGCCTCCTTGGCTCACTTACTGCCAGAAGGAACCCCAACACCCCTGATCCCC
      **** ** ******** * *** **** ** ***** ***** ** ** ***** ** 
CHI  CAACCCTCCGCCTCCTTAGGACACCTACTCCCTGAAGGCACCCCCACCCCACTGATTCCA
      ***       ***  * *  *** * ** **  **** ** ** *   **** ******
MOU  AAACTAAAAAGCTCACTTGCCCACTTCCTTCCACAAGGAACTCCAATTTCACTAATTCCA
     ***   ***   *********** *** * *********** ** * * ******* ***
BOV  AAAACTAAAGCATCACTTGCCCATTTCTTACCACAAGGAACACCCACTCCACTAATCCCA
     ** * ****    * ** ***** ************** ***** ** ** ** ***** 
HUM  AAGATTAAAAATGCCCTAGCCCACTTCTTACCACAAGGCACACCTACACCCCTTATCCCC

DUC  GCACTAATCCTGATCGAAACAACCAGCCTGCTGATCCGGCCCTTAGCTCTAGGAGTCCGC
     ** ****** * ***************** ** ** ***** ***** ******** ***
CHI  GCCCTAATCATAATCGAAACAACCAGCCTACTTATTCGGCCATTAGCCCTAGGAGTACGC
        ** ** ** ** *******  ****** ******  *** * **  *** *** ** 
MOU  ATACTTATTATTATTGAAACAATTAGCCTATTTATTCAACCAATGGCATTAGCAGTCCGG
     *****  * *********** ** ***** *********** ** **  * ** ** ***
BOV  ATACTAGTAATTATTGAAACTATCAGCCTTTTTATTCAACCTATAGCCCTCGCCGTGCGG
     ******** ***** ***** ********  * ******** ******** ***** ** 
HUM  ATACTAGTTATTATCGAAACCATCAGCCTACTCATTCAACCAATAGCCCTGGCCGTACGC

DUC  CTCACAGCTAACCTCACAGCAGGCCACCTACTTATTCAACTCATCTCCACAGCCTCCATC
     ** ***** *********** ** *********** ***** ***** ****** * ***
CHI  CTAACAGCAAACCTCACAGCTGGTCACCTACTTATCCAACTTATCTCTACAGCCACAATC
     ** ***** *** * ** ** ** *** ** * ** ** ** ***     *** **  * 
MOU  CTTACAGCTAACATTACTGCAGGACACTTATTAATACACCTAATCGGAGGAGCTACTCTA
      * *********** ************ ******* ******************** ** 
BOV  TTAACAGCTAACATCACTGCAGGACACCTATTAATTCACCTAATCGGAGGAGCTACACTT
      **** ******** ******** ****** * ** ******** *** * ** ** ** 
HUM  CTAACCGCTAACATTACTGCAGGCCACCTACTCATGCACCTAATTGGAAGCGCCACCCTA

DUC  GCACTCAAGCCCATCCTTCCCACAGTATCAATCCTAACAATAGCCATCCTACTACTCCTC
     ** **    ** **  * **  ** * **   ******   *  ******* * ** ** 
CHI  GCCCTACTACCAATAATGCCATCAATCTCCGCCCTAACGGCACTCATCCTATTCCTACTA
     *   ** **   ** *  *** ***   *  ** * **     * **  ** * ***** 
MOU  GTATTAATAAATATTAGCCCACCAACAGCTACCATTACATTTATTATTTTACTTCTACTC
     * * ******  ****** *  ********   ******** *  *** ** * ***** 
BOV  GCACTAATAAGCATTAGCACTACAACAGCTCTAATTACATTCACCATTCTAATCCTACTA
     *** **  ** ***** *  * *  *  * ** ** *  ***** ******** ***** 
HUM  GCAATATCAACCATTAACCTTCCCTCTACACTTATCATCTTCACAATTCTAATTCTACTG

DUC  ACCATCCTAGAAGTAGCAGTGGCCATAATCCAGGCCTACGTTTTCGTCCTCCTCCTAAGC
     ************** ***** ** ******** ******** ******************
CHI  ACCATCCTAGAAGTGGCAGTAGCAATAATCCAAGCCTACGTCTTCGTCCTCCTCCTAAGC
     ** ** ****** * ********* **** *********** ***  ******  *****
MOU  ACAATTCTAGAATTTGCAGTAGCATTAATTCAAGCCTACGTATTCACCCTCCTAGTAAGC
     *********** ***********  **** ******** ******** ******** ***
BOV  ACAATTCTAGAGTTTGCAGTAGCTATAATCCAAGCCTATGTATTCACTCTCCTAGTCAGC
     ** ** *****  * ** ** **  ************* ** ***** ** ***** ***
HUM  ACTATCCTAGAAATCGCTGTCGCCTTAATCCAAGCCTACGTTTTCACACTTCTAGTAAGC

DUC  CTGTACTTACAAGAAAACATCTAA
     ** ************** ** ***
CHI  CTCTACTTACAAGAAAATATTTAA
     ** **  **** ** ****  ***
MOU  CTATATCTACATGATAATACATAA
     ******** ***** ** ******
BOV  CTATATCTGCATGACAACACATAA
     ** ** ***** ************
HUM  CTCTACCTGCACGACAACACATAA
```

ATPase8

```
DUC   ATGCCTCAACTCAACCCTGCACCATGATTCTCAATCATAGTCATAACCTGACTAACCCTC
      ***** *** * *****    *********** ****** ** **** *** * *** **
CHI   ATGCCCCAATTAAACCCAAACCCATGATTCTCCATCATACTCCTAACTTGATTCACCTTC
      ***** *** ** *  **    *******   **  **  **  * *   * * *** *
MOU   ATGCCACAACTAGATACATCAACATGATTTATCACAATTATCTCATCAATAATTACCCTA
      ***** ******** ** ********* * *  *  **  * ***  * *  * *****
BOV   ATGCCGCAACTAGACACGTCAACATGACTGACAATGATCTTATCAATATTCTTGACCCTT
      ***** ****** * **  *   *** *  ** ** **     * *** ** * ** **
HUM   ATGCCCCAACTAAATACTACCGTATGGCCCACCATAATTACCCCCATACTCCTTACACTA

DUC   GCACTCCTAATCCAGCCAAAACTGCTAACCTTCACCACAACAAATCCCCCATCAAAAAAA
       * ** ** ***** ** ***** **  * *****   ******   ***  **** ***
CHI   TCTCTGCTTATCCAACCCAAACTTCTTTCATTCACTCTAACAAACAACCCTGCAAACAAA
      * * *  *  * ****  *** *    ***   **  *  **      *** **   ***
MOU   TTTATCTTATTTCAACTAAAAGTCTCATCACAAACATTCCCACTGGCACCTTCACCAAAA
      ****** * ************** ***  *** *  **    *     **   **  * *
BOV   TTTATCATCTTTCAACTAAAAGTTTCAAAACACAACTTTTATCACAATCCAGAACTGACA
      **  *****   ********* * * ***    ****   * *    ***   **  *
HUM   TTCCTCATCACCCAACTAAAAATATTAAACACAAACTACCACCTACCTCCCTCACCAAAG

DUC   CCATCACTCATCACCAAA...CCCACACCATGAGCCTGACCATGAACCTAA
          **   *  ** ***   ***** ** *** *****************
CHI   ATTACA...ACAACTAAA...CCCACCCCCTGAACCTGACCATGAACCTAA
           *   ** *  ***      ***** ***   * *  ******  **
MOU   TCACTAACAACCATAAAAGTAAAAACCCCTTGAGAATTAAAATGAACGAAA
       **  ** **   *****  *** ************  *************
BOV   CCAACAAAAATATTAAAACAAAACACCCCTTGAGAAACAAAATGAACGAAA
      ** * *******  ***  * ****  ** ****** **************
HUM   CCCATAAAAATAAAAAATTATAACAAACCCTGAGAACCAAAATGAACGAAA
```

ND6

```
DUC   AAGCTTCCGGAAAAGGCTCCGCGGCCAGAGCTACAGAGTAGACAAATACCACCAACATGC
      **** ** *** **** ***** ** *  *  ***** ** ***** *********** *
CHI   AAGCCTCAGGATAAGGATCCGCAGCTAATGACACAGAATAAACAAAAACCACCAACATCC
      *** *** *****    ** * *****  *     * ***  **** ** **********
MOU   AAGTCTCTGGATATTCCTCAGTAGCTATAGCAGTCGTATATCCAAACACAACCAACATCC
      * * *** ***** ****** **** ** ** ** *********** ******* *** *
HUM   ATGCCTCAGGATACTCCTCAATAGCCATCGCTGTAGTATATCCAAAGACAACCATCATTC
      *    **********  **  ******* ** ** ***** ***** ** ******** *
BOV   AAATTTCAGGATACTGTTCTGTAGCCATAGCCGTTGTATAACCAAAAACTACCATCATAC

DUC   CCCCAAAATACACCATAAACAGCACCAGGGCTACAAACGAAACCCCGAGGCTCAATAACC
      **** ********** *** * * *** **  ***** ***** ** *  **** *****
CHI   CCCCTAAATACACCAAAAATAACGCCAAGGAAACAAAAGAAACACCCAAACTCACTAACC
      ********** *  ***** *     **    * *** **  **** ** *  *    *
MOU   CCCCTAAATAAATTAAAAAAACTATTAAACCTAAAAACGATCCACCAAACCCTAAAACCA
      ******************************* * * *   *** *****    * **  *
HUM   CCCCTAAATAAATTAAAAAAACTATTAAACCCATATAACCTCCCCCAAAATTCAGAATAA
      * ** ************** ** ** ** **** * *   *** **********  * **
BOV   CTCCCAAATAAATTAAAAAGACCATCAACCCCAAAAAGGATCCACCAAAATTCAATACAA

DUC   ACCCACACCCAGCCACAGATGCTAAAACTAAGCCAACTACCCCATAGTACGGCGAAGGAT
      ******* **   ****** ****  ** ** ** ** ******** *************
CHI   ACCCACATCCTATCACAGACGCTACCACCAACCCCACCACCCCATAATACGGCGAAGGAT
          *** ** *  *        ** *  ** ** *   * ******   ** ***** *
MOU   TTAAACAACCAACAAACCCACTAACAATTAAACCTAAACCTCCATAAATAGGTGAAGGCT
      * * *** ** ** *  ** ******** **  ******* *********** *******
HUM   TAACACACCCGACCACACCGCTAACAATCAATACTAAACCCCCATAAATAGGAGAAGGCT
      *  **** ** *** **** ** ***** **  **** ************** ***** *
BOV   TTCCACAGCCAACCCCACCACTCACAATTAACCCTAACCCCCCATAAATAGGTGAAGGTT

DUC   TCGATGCTACACCTAAAACACCCACAACAAAGCAGATTCCTAGAAAAAATACAAAATAAG
      * ** ** *** ******* ****  * *** ** ** ** *******  **********
CHI   TAGACGCCACAGCTAAAACCCCCAGCATAAAACAAATCCCAAGAAAAATCACAAAATAAG
      *  * ** *   * * *   ** * ** ***  *    *  * ** **  * * *** *
MOU   TTAATGCTAACCCAAGACAACCAACCAAAAATAATGAACTTAAAACAAAAATATAATTAT
      *  * *  ***** * * * ** *  *  **      *** ** * *** * *  **
HUM   TAGAAGAAAACCCCACAAACCCCATTACTAAACCCACACTCAACAGAAACAAAGCATACA
      * ***************** ** ** ** ***   ** ** *  * *** * *   ** *
BOV   TCGAAGAAAACCCCACAAAACCTATCACGAAAATAACGCTTAGAATAAATACAATGTATA

DUC   T...CAT
      *   ***
CHI   T...CAT
      *   ***
MOU   T...CAT
      *   ***
HUM   T...CAT
          ***
BOV   GTATCAT
```

Appendix B.

Multialignments of amino acid sequences. Identical and similar residues between species are respectively indicated by asterisks and colons, missing amino acids by dashes. Abbreviations of vertebrate species are as in Annex I.

COII

```
CHI   MANHSQLGFQDASSPIMEELVEFHDHALMVALAICSLVLYLLTLMLMEKLS.SNTVDAQE
      ****************************:*************  ******* ** *****
DUC   MANHSQLGFQDASSPIMEELVEFHDHALIVALAICSLVLYLLAHMLMEKLS.SNAVDAQE
      **:  *:*:*** *******: ******:: : ** **** *   *  **   *  ****
HUM   MAHAAQVGLQDATSPIMEELITFHDHALMIIFLICFLVLYALFLTLTTKLTNTNISDAQE
      **   *:*:***********: **** ***:***  **** : * ******:*   ****
BOV   MAYPMQLGFQDATSPIMEELLHFHDHTLMIVFLISSLVLYIISLMLTTKLTHTSTMDAQE
      **** ***:***********::**************************************
MOU   MAYPFQLGLQDATSPIMEELMNFHDHTLMIVFLISSLVLYIISLMLTTKLTHTSTMDAQE

CHI   VELIWTILPAIVLVLLALPSLQILYMMDEIDEPDLTLKAIGHQWYWTYEYTDFKDLSFDS
      ************************************** *********************
DUC   VELIWTILPAIVLVLLALPSLQILYMMDEIDEPDLTLKSIGHQWYWTYEYTDFKDLSFDS
      :* :*******:***:***** **** **:::* **:***************:  * *:*
HUM   METVWTILPAIILVLIALPSLRILYMTDEVNDPSLTIKSIGHQWYWTYEYTDYGGLIFNS
      :**:*********:************ **:*:****:* :****** ******  * *:*
BOV   VETIWTILPAIILILIALPSLRILYMMDEINNPSLTVKTMGHQWYWSYEYTDYEDLSFDS
      **********:***:****************** ********************** ***
MOU   VETIWTILPAVILIMIALPSLRILYMMDEINNPVLTVKTMGHQWYWSYEYTDYEDLCFDS

CHI   YMTPTTDLPLGHFRLLEVDHRIVIPMESPIRVIITADDVLHSWAVPALGVKTDAIPGRLN
      ** ****** ***********:*:************ ********* *************
DUC   YMIPTTDLPNGHFRLLEVDHRVVVPMESPIRVIITAGDVLHSWAVPTLGVKTDAIPGRLN
      **:*   *  * :***:**:***:*:* ***::**  ************:**********
HUM   YMLPPLFLEPGDLRLLDVDNRVVLPIEAPIRMMITSQDVLHSWAVPTLGLKTDAIPGRLN
      **:*   * **:****:********:*  ***:: *:********* *************
BOV   YMIPTSELKPGELRLLEVDNRVVLPMEMTIRMLVSSEDVLHSWAVPSLGLKTDAIPGRLN
      ***** :********************: ****:**************************
MOU   YMIPTNDLKPGELRLLEVDNRVVLPMELPIRMLISSEDVLHSWAVPSLGLKTDAIPGRLN

CHI   QTSFITTRPGVFYGQCSEICGANHSYMPIVVESTPLKHFEAWSSLLSS
      **********:*************************  **********
DUC   QTSFITTRPGIFYGQCSEICGANHSYMPIVVESTPLPYFEAWSSLLSSS
      ** *  ****::*************:****:*  **  **    ::
HUM   QTTFTATRPGVYYGQCSEICGANHSFMPIVLELIPLKIFEMGP.VFTL
      ***:   ***:********** ***********:*** **       *
BOV   QTTLMSSRPGLYYGQCSEICGSNHSFMPIVLELVPLKYFEKWS.ASML
      * *: * ****:********************:******* ** ***:
MOU   QATVTSNRPGLFYGQCSEICGSNHSFMPIVLEMVPLKYFENWS.ASMI
```

ATPase8

```
CHI   MPQLNPNPWFSIMLLTWFTFSLLIQPKLLSFTLTNNPANKIT.TTK.PTPWTWPWT
      ****** ******::**:*: ******** ** ** *  *    ** **** ****
DUC   MPQLNPAPWFSIMVMTWLTLALLIQPKLLTFTTTNPPSKKPSLITK.PTPWAWPWT
      ****:   **  ::   :** :* * *:   *    ** *     *  ***   **
MOU   MPQLDTSTWFITIISSMITLFILFQLKVSSQTFPLAPSPKSLTTMKVKTPWELKWT
      *********:  *:*  :****:******  :*   *       ::  **** ***
BOV   MPQLDTSTWLTMILSMFLTLFIIFQLKVSKHNFYHNPELTPTKMLRQNTPWETKWT
      ****:*  * ***  *:****:* ***:   *:   *   * **   * *** ***
HUM   MPQLNTTVWPTMITPMLLTLFLITQLKMLNTNYHLPPSPKPMKMKNYNKPWEPKWT
```

ATPase6

```
CHI   MNLSFFDQFSSPHLLGHPLILLSLLLPALLFPSPGNRWINNRLSTIQLWLLHLITKQLMI
      ************ *** **** ********:******************: ********
DUC   MNLSFFDQFSSPCLLGIPLILPSLLLPALLLPSPGNRWINNRLSTIQLWFTHLITKQLMT
      **  :*  *  * ::* *:::  :::* :*:**   * *****   * *:  ** **:*
MOU   MNENLFASFITPTMMGFPIVVAIIMFPSILFPSS.KRLINNRLHSFQHWLVKLIIKQMML
      ****** ***** ::*:*:*  *::***:*** *  **: **:  :*:*:: *: ****
BOV   MNENLFTSFITPVILGLPLVTLIVLFPSLLFPTS.NRLVSNRFVTLQQWMLQLVSKQMMS
      ****** *** * *****   **:*** ** ***   *: **::* ***:: * *****
HUM   MNENLFASFIAPTILGLPAAVLIILFPPLLIPTS.KYLINNRLITTQQWLIKLTSKQMMT

CHI   PLNKNGHKWALMLTSLMTMLLTINLLGLLPYTFTPTTQLSMNMALAFPLWLATLLTGLRN
      **** ******:****: *** ************************:*************
DUC   PLNKAGHKWALLLTSLILMLLSINLLGLLPYTFTPTTQLSMNMALALPLWLATLLTGLRN
           *: * *:: ***: : * ******* ***********: :*:***   ::**:*:
MOU   IHTPKGRTWTLMIVSLIMFIGSTNLLGLLPHTFTPTTQLSMNLSMAIPLWAGAVITGFRH
      **  ** *****::***:************* *********** ***************:
BOV   IHNSKGQTWTLMLMSLILFIGSTNLLGLLPHSFTPTTQLSMNLGMAIPLWAGAVITGFRN
      :** ** ** ***:***:**  ********************* ******** ** ***
HUM   MHNTKGRTWSLMLVSLIIFIATTNLLGLLPHSFTPTTQLSMNLAMAIPLWAGTVIMGFRS

CHI   KPSASLAHLLPEGTPTPLIPALILIETTSLLIRPLALGVRLTANLTAGHLLIQLISTASI
       ***** ****************:********************************** *
DUC   QPSASLGHLLPEGTPTPLIPALIMIETTSLLIRPLALGVRLTANLTAGHLLIQLISTATI
          ** *:**:***  *** **:*** **:* *:** ******:******::**  **:
MOU   KLKSSLAHFLPQGTPISLIPMLIIIETISLFIQPMALAVRLTANITAGHLLMHLIGGATL
      * * ***********  *****:****************************:********
BOV   KTKASLAHFLPQGTPTPLIPMLVIIETISLFIQPMALAVRLTANITAGHLLIHLIGGATL
      * *  *************************:********************:**** ***
HUM   KIKNALAHFLPQGTPTPLIPMLVIIETISLLIQPMALAVRLTANITAGHLLMHLIGSATL

CHI   ALKPILPTVSILTMAILLLLTILEVAVAMIQAYVFVLLLSLYLQENI
      ** *::* :* **  **:*****************************
DUC   ALLPMMPSISALTALILFLLTILEVAVAMIQAYVFVLLLSLYLQENI
       *: : *    :* :**:****** ***:****** **:****::*
MOU   VLMNISPPTATITFIILLLLTILEFAVALIQAYVFTLLVSLYLHDNT
       ** **  ** *** **:**********:******************
BOV   ALMSISTTTALITFTILILLTILEFAVAMIQAYVFTLLVSLYLHDNT
      *:  *     ** *********** ***:******  **********
HUM   AMSTINLPSTLIIFTILILLTILEIAVALIQAYVFVTLVSLYLHDNT
```

ND6

```
CHI   M.TYFVIFLGICFMLGVLAVASNPSPYYGVVGLVVASVMGCGWLVSLGVSFVSLALFLVY
      * **** ******::*** ***************:*** *****:******* * **:**
DUC   M.TYFVFFLGICFVVGVLGVASNPSPYYGVVGLVLASVAGCGWLLSLGVSFVALVLFMVY
      *   ::* *   *:** **:*  *** ** :**::    **  :* :* **: *::*::*
MOU   M.NNYIFVLSSLFLVGCLGLALKPSPIYGGLGLIVSGFVGCLMVLGFGGSFLGLMVFLIY
      *    :*:** :*::* :*:  *************** *** :** **************
BOV   MMLYIVFILSVIFVMGFVGFSSKPSPIYGGLGLIVSGGVGCGIVLNFGGSFLGLMVFLIY
      * :* :*:*** :****************** ***** *** *:***** ::********
HUM   M.MYALFLLSVGLVMGFVGFSSKPSPIYGGLVLIVSGVVGCVIILNFGGGYMGLMVFLIY

CHI   LGGMLVVFVYSVSLAADPYPEA
      :*********** ***:*:***
DUC   FGGMLVVFVYSVALAAEPFPEA
      :******* *  *:* * :**
MOU   LGGMLVVFGYTTAMATEEYPET
      ****:************:***
BOV   LGGMMVVFGYTTAMATEQYPEI
      *************** *:***
HUM   LGGMMVVFGYTTAMAIEEYPEA
```