```
      +---------------------------------------PDB domain sequence [Gallus gallus]
      !
      !              +------------------------citrate synthase [Bos taurus]
      !              !
      !              !              +---------PREDICTED: citrate synthase [Pan troglodytes]
      !              !         +-90.0
      !         +-39.0         !    !    +----citrate synthase precursor, isoform a [Homo sapiens]
      !         !    !    +-97.0    +-78.0
      !         !    !    !    !         +----citrate synthase [Homo sapiens]
      !         !    !    !    !
+-53.0          !    +100.0    +--------------citrate synthase [Homo sapiens]
!     !         !         !
!     !         !         !              +----PREDICTED: similar to citrate synthase precursor, isoform a [Macaca mulatta]
!     !         !         +-----------92.0
!     !         !                        +----CISY_MACFA Citrate synthase, mitochondrial precursor
!     !    +-38.0
!     !    !    !
!     !    !    !
!     !    !    !               +---------ENSMUSP00000052373-->ENSMUSP00000005826->
!     !    !    !               !
!     !    !    !         +-100.0
!     +-79.0    !         !     !
!         !    !    +-89.0     +--------------citrate synthase [ENSMUSP00000005826] [Mus musculus]
!         !    !    !    !
!         !    +-44.0    +-------------------citrate synthase [Rattus norvegicus]
!         !         !
!         !         +------------------------citrate synthase [Xenopus tropicalis]
!         !
!         +---------------------------------PREDICTED: similar to citrate synthase [Canis familiaris]
!
+---------------------------------------------citrate synthase [Sus scrofa]
```

Supplementary Figure 3: Example of a PR of citrate synthase in Mouse. ENSMUSP00000052373 is a PR made from ENSMUSP00000005826, which is conserved in other mammals with the same intron-exon structure. Pictures of intron-exon structure were taken from the Ensembl website [http://www.ensembl.org]. The tree is a maximum-likelihood majority-rule consensus tree, that was made using PHYLIP v3.6b, based on CLUSTALW multiple alignments [ref.] of homologous proteins to ENSMUSP00000052373 in other species, that were collated using *blastp* [e-value $\leq 10^{-4}$] using the NCBI blast server and the *nr*

database [http://www.ncbi.nlm.nih.gov/BLAST/]. One hundred bootstrap replicates were generated using the PHYLIP *bootseq*, *proml* and *consense* programs [ref.].