

(A) Without Optimization

```

AW874468      180  gtcattttctttaactttgcaaactgggtcttcctgctggcacatactgtg
HsChr1      103833954  gacattttctttaactttgcaaactgggtcttcctgctggcacatactgtg

AW874468      atgtcatttttggagataaaaattaatggcaattgcacaggcattaaaatt
HsChr1      atgtcattttctggagataaaaattaatggcaattgcacaggcattaaaatc

AW874468      tacgtttctgatgatggcaaagctcatttttctattag      95
HsChr1      tacgtttctgacgatggcaaagctcatttttctatta      103834091

<GTxxxxxxxxx...xxxxxxxxxxxxAG>

AW874468      94  taactctgctgaagatccatttattgcaattcatgctgaatttaaattgt
HsChr1      103880338  taactctgctgaagatccatttattgcaattcatgctgaatctaaattgt

AW874468      aaaattttaaataaatgcatgtcctc      18
HsChr1      aaaattttaaataaatgcatgtcctc      103880414

```

(B) With Optimization

```

AW874468      180  gtcattttctttaactttgcaaactgggtcttcctgctggcacatactgtg
HsChr1      103880200  gtcatttactttaactgtgcaaactgggtcttcctgctggcacatactgtg

AW874468      atgtcatttttggagataaaaattaatggcaattgcacaggcattaaaatt
HsChr1      atgtcattttctggagataaaaattaatggcaattgcacaggcattaaaatt

AW874468      tacgtttctgatgatggcaaagctcatttttctattagtaactctgctga
HsChr1      tacgtttctgatgatggcaaagctcatttttctattagtaactctgctga

AW874468      agatccatttattgcaattcatgctgaatttaaattgtaaaattttaaatt
HsChr1      agatccatttattgcaattcatgctgaatctaaattgtaaaattttaaatt

AW874468      taaatgcatgtcct      18
HsChr1      taaatgcatgtcct      103880414

```

Additional File 1. Exon-factorization into 2 or 1 exons of the EST AW874468 (S1938294) (*S*) (Unigene cluster Hs.514463, build #176) related to the gene *AMY2A* (*G*) on Chromosome 1. The EST sequence *S* agrees with the genomic sequence *G* in two possible ways: either *S* is factorized into 2 exons of *G*, that is $S = \langle C_1, C_2 \rangle$ or into 1 exon of *G*, that is $S = \langle C \rangle$ where $C = C_1 C_2$. Indeed, *G* contains two distinct "identical" regions that align to factor *C* of EST *S*, where *C* is about 200 bp long. The alignment of the same region of the EST is reported without the application of optimization criteria (2 exons) (A) and using the minimization criteria (1 exon) (B). EST and Chromosome coordinates (Ensembl release 26.35.1) are reported at start and end alignment positions.