

Supplementary file S2

Annotated 5' flanking sequences of the human *MKLP1*, *Ect2* and *RacGAP* genes.

Nucleotide sequences, 5' of the transcription start sites, were analyzed with the genomatix suite (MatInspector) and TSSW programs. Shown in colour are the predicted binding sequences for CDP/cux (red), E2F (blue), MYB (purple) and CHR boxes (green).

The major transcription start sites identified by RLM-RACE are marked as ▼ **5'RACE**.

Position of the longest EST available from NCBI data banks are marked as ▲ followed by their accession numbers.

Oligonucleotides used for cloning are shown in italics and underlined with an arrow.

5'flanking sequence of the human MKLP1 gene

```
-1140 GAGGGGTGGG AGACAAGAGG GCAGGAGAAA GTCAGAGACA GACTTTGCTT TTAAGGCTTC
-1080 TCTGAGGCAC TTCAGTCTT CT CHR E2F2 CDP/Cux
TTAGTTCA AAGCA TCAG TGTGCCAAAG AGCCATACTC
CDP/Cux
-1020 TGGGGTGTCG TTCCTGAGC ACCAACAATA TTGGCCAAGG ACATTGTG AC TATCCTTACC
AGTGCTGTGG TCCTTTAGGA AGTTGAGAAG GGGCTAGAGT TCCATGAGAG CCCAGGTGTA
-960 GATGTGTATT TGGTAGTCAT CGGCTTAGGG TCAAAGACAG AAGTGATATA GAATCAGTTC
CHR
-840 ATCTAAGGCA AGAG TAGGTT CAAAAGGTTG CCCCTTAAGG GTAAGGACCA AATTTTATTC
-780 ATCTCCGTAA GGCCAGGATC AAACACAATC AAGATTTGAT AAAACAGGCC GGGCGCAGTG
-720 GCTCACGCCT GTAATCCCAG CACTTTGGGA GGCCAAGGTA GGCAGATCAC CTTAAGTCAG
-660 GAGTTCTAGA CCACCCTGGC CAATATGGTG AAACCTCATC TCTACTAAAA ATACAAAAAT
E2F3
-600 TAGCTGGGCG T GGTGGTGCG CGCCTGTAAT CCCAGCTACT CGGGAGGCTG AGGCAAGAGA
-540 ATTGCTTGAA CCTAGGAGGT GGAGGTTGCA GTGAGCTGAA GACTGCAGTC CAGGCTGGGC
-480 GACAGAGCAA CTCCACCTCA AAAAAATAAAA AAGGAAAAAAA AAAAGATTTG ATAAAAATAAT
-420 GTACATATAG GACAGAGCTG GAGGGAAACC CAGGAAAGGT GAATCAGAGA GAGAAAAAGAC
-360 GCCAAGATAA TGTGGTGA CT TGGGAGCCAG AGAGACGGAG CCTGAGACTT GGAGCAGGGC
E2F1
-300 ACAGCACAGG AATACAA CAC TTTTACCAG AACGACATTC CGACCTCTGA ATCAACCCTG
-240 GAAAGGCCCT TCCATTGAAC TTCTTGTGAT GGAGGAAAAG AAACCCTCAG TTTCGTATC
E2F1
-180 GCAAGGCGGA AGAGTTTAAA TACCATACCT CAAGATCTCC CTCCAGCCCT GACACGCGAC
-120 TCCGCGGCC TTTACGCGTC GGTCGCCTAG GAGACAAGCG CCACTTCCGG CAGCGAAGTC
-60 ACTTCCGCGG AGGTCCTGCC TCTCCCTGCG GACGGGCGCT GATTGGC CHR
5'RACE 5'RACE
CCG TTTGAAATGC
+1 GCCAGGCGGG AGCTCACCTG GGCTCTTAGC GTCGCCGCGG GCTTCGAGA GCACCG CGCC
BC051826.1
E2F1 MYB
+61 TTAGCCGCGA AGTTCTAGTT CTTGCTGCCG GT CCTAACGT CCCGCAGTCT TCGCCAGCCA
5'RACE
+121 GCCGTCCC GC ATGCGCGTTT GGGCGGCGTG GAGCCTGCTG CCATGAAGTC AGCGTGAGTA
```

5'flanking sequence of the human Ect2 gene

-2800 CCTTAGTCTC CCAAAGTGTCT GGAATTATAG GCATGAGCCC CCACACCTAG CCTAACCTAG
-2740 CGTATGTAC → **CDP/CUX** TTGTCTTCCA CTGTCTCTTT CATGTATATT ATGATCCAGC CACACCAAAC

-2680 TGTTTAC **TGT CTCCTATATA TATTGA** GCTT TTATGAGAGC CTAAAGTGCA CGTAAACAAG
-2620 CTAATATTAA GAATTTATGT TTTTAAAAAA GTTCAAGGTT AATATGCCAT TCAGTTCAAT
-2560 TGCCTATGCC GTATATAGTA TTTAAAAAATT TTTCATAATG CTGATCTTGA AACACTATAG
-2500 CACTTGAAAC AGCTTCTGAC TCAGTAAGAC TAAGTTACTT GTATCATTGT TTTGTAG **GTA**

CHR **E2F1**
-2440 **TTTGAATTCT** GTGAAAACCT ATGCATGCAA TAA **ATTTTTC ACCGCAAATA CAG** CTATCTG
-2380 CTAGAAGTGC CATGAGAACA AGGCCATAT TTGTCTTGTC ACACCTATAT TCTCAGTGCT
-2320 AGCATAGTAC CTAATGGGTG CTTAATAGTT ATTAATGAA TAACTATGA TTAGGTTAAT
-2260 CAGGACATCC AGTCTCTCAA ACATTATGAA TGAACCTGGA TCTAGCTCAT TCACTCTCGA
-2200 AATATTTTATT TGGTACCTAC ATGTGAAGAT TCCTGTGTTT ATTAGAGAGC AATGTGTAAC
-2140 ATTTGCCTTT GCAGTACTAG CTTTTTAAAA ATACCTGAAG ATGAGGGATC ATGTTTTACT
-2080 TTACCAGACA CCAGCCCTCT GCTAGGTGCA ATTAAGTGGC ATATGATCAC CTATGCACAC

CDP/CUX
-2020 GATTATCTGA AATACTCCTC TCCACATTGC TTTCTTTGAG CCTGCA **GACA GATCTATACT**
-1960 **AGACT** TTATT TACTCATC TATTATTTCT GGGTCTAAT GCGGGCTAAA ATATTTATTA
-1900 GGAGACTG GGTAAAGTTAC TTAACCTTT CCTACAGTTT CATGGTATCA TGGGACTATG
-1840 ATACCTACTT GTGGGTGTT TTAAGGATTA AATAAGGTAT AAAGTCATA TAAAGTATC
-1780 AGTATGTTCA TTCTCTTTCT TCTCTCCCTT ATACTGTAAT TATCTGTTTA TCTGCTAATT
-1620 TCCCTACCA GATTGTGTGG TCCAGACAGG CAGTTTTCTA TGCCTTATTC ATCTTTATAA
-1560 CCCCACAAG CACAATGGCT GTCTTAAAAAT TAGTAAATAT TTGCTGAGTA AATGAAGACC

E2F1 **MYB** **MYB**
-1500 **TAAGAGAGTG AAAAAGG** TTG **TATTACCCGT TTAT** TTGTTA A **TCTAACTGC CACA** CATCCC

MYB
-1440 TTAAAAGAAA TCAAGTACAG CTTTTGTTAG GGAAGCTTAA AATTCATGGC AATG **ATACCA**
-1380 **TACACAGTTA AT** CATTAGCT TTGGAAAGTC TGTGAGCAG AAAGCGGTGC ACAAAGGAAG
-1320 TAGGGCTGCT TCCACAATA GCTTAATATG TTTTCAGTGA TGAACGTTT TGGAAATGCTG
-1260 GGCACACCCG ACTTCCAAT ACGCAAGCAC ATTCCAATG GAAATTTCAA TFACTGTTTT

MYB
-1200 TCTTTTTTTC TGC **CGTAACA GTTATC** ACAG CATTTAATG AAGGTCAGAC TTAATACTG
-1140 ACAAGCAGGG AAACCAATAA TTACCAGATT GCTTCATTCA TTTTGTGGAC AGCATTATA **A**

E2F2
-1080 **ACTTTTCAGG CATTCT** AAGA TTGCCTGACA AAATTAGAAA AAATTGTTAA GTCCCCACAA
-1020 CCTGCTTGAG ATGCTCACAA GTGTTCCATC CCAAAGTACA CGTTTGCCTC TTCTTATTGG
-960 TATTTAGGTT CTGGCATCTG ATTCCCCTTT CTTCTTCTAT TTATATTAGC CTTCAGGTGG
-900 GAACTATGGA AAACCCACCT GGTGGGGGTG TATTTTTCTT AAGAGATTGT GTTATATGGT
-840 ATCATATTTA GAAAATAAAG AAATTATTTT CATAATAGCA ACGCTAATCA CAGGAAAAAC
-780 TTGATACATT CTTGACTAGA CTTTAGGGGA AAGTACAAGT TCTTAGTGGG CGGAGGGGGG

E2F4
-720 CAAAGACGAA **GCAATTTGGC TCCAAAAATT** CTAGAATCTC GTGGCATATT TTTAT **CAAAAT**

CHR
-660 **TCAAGCTT** AT CTTTAAAAAA ATCCGAGAAT TTTTCTAATT TCTAAGAATT GATTCTAACCC
-600 CGCGATACTC CCTAGATGGG AGCGCAGTGC TTTAAATAAT AGCATTTAAA TAAGCGGCTG

CHR
-540 **TGATTTGAA GTCC** ACCTGG GGAGTCCGAG TTCTAGGACC GGTCCTTGGC CCTGCCACAC
-480 GCCAATACAA CCATCAAAGT CAAAATCCCT TAACTCAAAT TTAAATACAT TCCTAAAGCA
-420 GAAAGCATCC CGATAGGAAA ATCCTACAGC TCACAGCAAG CAGCAGGGAA ACATTTATCT
-360 CTGGCGGCAC TGATTTAAGC TTAAGAGATT TGGTATCGTT TAAAAAATAA AAAAAGCTGG
-300 GTGGGGCCGA GGAGAATCCA CCCAACTTT CCTCAGGAAG TAGGATGTTT ACTGAGTAAA
-240 TATGCGGTAG TCACAGAAAC ACCTGCTACT CTTTAAAGAC CTCAGGAGCC ACCACCGCCA
-180 CCTCCTGGCC ACATCCGGGC ACTTCTCTGC AAATCCTGG CGTCCCAGC CCCGATCCT

CHR
-120 TGGGC **TTATT TGAAACAC** AC GCACAACCA TCAAGCGGCG CCTGCCGGGA CTTGGAGAGC

CDP/CUX **E2F4 CDP/Cux**
-60 GTCGACCAAT CACCGCCCT CT **GGAGACTC CTGCCATATG** GAAAGAGGTT **GACATCACGC**

▲ **5'RACE** **5'RACE** **CHR**
+1 **GCCAAATCGGC ATGGC** TCTTA GAGAGAGCAG CTTA **GTTTTT GAATCGG** TTG TGGCGGCCGC
▲ BC070038.1

CHR
+61 CGGCGAGGAA TGGCGGTATT **TGTGAGAGGA GTCC** **GCGTTT GAAGAGG** TTG AACTCCTAGG
+121 GCTTTTTTGA GAGTGGTAAG GACTCTGTGT GAGTGAGAGA GGTGCAAGTG GCGAGCGGC

5'flanking sequence of the human RacGAP gene

```

-2100 GGCTTCCCAA AGTGCTGGGA TTACAGGCAT GAGCCACTGT GCCTGGCCTT GTGTATTTAT
-2040 TAAATATATG CTTGCCTCAT TCTTCAGAGG TCACATGATC TTCTTATTGC CAGTAATTTT
-1980 CCATTCCCTT CTCTCCTGTA AATTCCCTATT CACACAAGAC CCAGCTCAGA TCAGTCTCTG
-1920 CAGGCTTTTT TAACACTCCC ACTCCCCAGA TAGGATTAAT TGCTTCCTTG TCTGTGCTCC
-1860 TAAAGTACTT TATACATGTC TTTGTCTTAC TACTTATTTT ATTTTATTAT AGGTATCTAA

                                CDP/Cux
-1800 GCAGGCACAG GCTA ATGCAA TCAATGAATG CTCCTGACCT TGAAAGGTTG TTAGGAGGAT
-1740 TAAATGGGAT AACAAATGCA CAGTTCATAT TTTCTTTCTT CATCCCTTTC TTAATAATAT

                                CDP/Cux
-1680 TAATAATAAT CTGTATTGCA GTTTACCAAG TATAAACTTT GATAGTGATC CACACTGTGT
-1620 CTCAGTCGAT CTGGGAATAA ACCTGTGAGA TAGAGATGGG ATTATTATTT CTACATTATA
-1560 CTTGGAGGAA GTGAAGTACA GAAGTATTAA AAAGTGATAA TATGGGTGGG TCAGCCGAAC
-1500 GCCTGCTATT GTTCAAGGCC AGGTGGAGTG CTACTGTCTG TAATCCAGT GCTTTGGGAG
-1440 GCTCAGGTGG GAGGATCACC TAAGGCCAGG AGTTCAAGAC TAGCCTGGAC AACATAGTGA
-1380 GACCCTTCTC TACAAAAGAT AAACAAACAA GCAAAAAGAA TACCTTGATC ATGTGCCCTGT
-1320 AGTCCTAGGT ATGAGGGAGG ATCCCTTGAG TGTAGGAGTT CATATTTGCA GTGAGCTATA
-1260 ATTGTGCCAC TGCACTCCAG CCTGCGTGAC AGAGTAAGAC CCTGTCTCTA AAATAAAAAGT

                                E2F1
-1200 GGAGTCA AAAT TTTGCGGTTA CAAAATTTGT GATGTGATTT TTGGATAAAA TACAATTTTA
-1140 TATATTTATG GTGATGTTTT GATGTATGTA TACTACTGTGG AATGACTAAA TCAAGCGAAT
-1080 TAATATATCC ATTACCTCAC ATACTTTATT GAGGTGTGTG GTGAGAGCAT TTAAGATCTA

                                CHR
-1020 CTCTTAGC AA TTTTCAAATG CAATTATACTT TGTATTAAAC TACAGTCAC T GTGTTGTACA

                                CDP/Cux                                MYB
-960 ATAGATCTAA ATTTTCCTCT AACTGAAAT TTATATCTCT TGACCAACAT CTCCCTAATC
-900 TCCCTCCCC AAATCTGTGA TCTTTCTTTT CTTTTTTTTT TTTTAAATTA AGGACCAATG
-840 TGACAGTAAG TGATCTTTCC TTTAAACCAC AGCGCTTCTT GTATGCTCCA GCCAGTATTT

                                E2F4
-780 TTCC AAATTG CGGGTAGCCA TCGCTTGTG AGTAGTGAGT TCTGGGTACT TTGCATGTAG
-720 TAAGGATGAG CAGAGTTTTT ATTACCCGGT TGCAACGTAA AATGTATTTT TCACTGTGGG
-660 TCCAGAGTTT TTAAAAAGCA TGAAAGCTCC TGTGCTAGCA CAATGCCAAG TTCTAAATAA
-600 CGTTTTATTAA ATGGAATTAT AGTATTTTAT TTTTCAATTTG TATTTTCTAG
-540 AGCTATCGCC TCCATTCGTA GCTACTGGAC AGTAGAGACC ATACCATCTT TATTTTCCG
-480 ATTCCCAGTG GAGTGCTTGT GATGATGTGG GCTCATTCAA TGTGTGTCAA ATAAAAGTCT

                                CDP/Cux                                MYB
-420 GAATG GATAC ATCAATGGAA GGATAATAT AGCTCAGTGA ATGTGCTTGA A ACTAACTGC

                                MYB
-360 TGTTCTAACT CCTGGTATTC TTCTGATTCT ACTGTGAGAG CTGACA TGTT TCAGTTAATC
-300 GTTTCCAAGG GCCCTAGTCC CGCAGACAGT GAGACCAAGA ATAATTCGCA TTCAGGAAGT

                                E2F2
-240 ACCACTTGCC CTTAGTCAAT ATGGCATCCT GAATGACTTC CGGAAA GGGC ACCGCAAAAA

                                MYB
-180 AAAACAAAA AAACCG AAAA ACCGTTGACA TTAGTCAAAA TGGCCACCTT TGGCTTCGCT

                                5'RACE
-120 TTGTCAGAAT GCGCATGCTC ATTCCCCTTC CCGTCACCTC CGCGTGGCCC TGACCGGTAC

                                E2F3
-60 GTCACACCGT AGAG GGACGG CGCGGGAGGA ATAAATTTCT CTGTGATTGG TTGGTGAA GG

                                5'RACE                                CHR                                E2F3
+1 TTTTCAAACC GGAGCTGTGG GCGCGGCGCT GCTCTGCCGT TGG GTGAGGC GCGGAGCGAA
                                ▲ AK000733.1

+61 GTGAAGGGTG GCCCAGGTGG GCGAGGCTGA CTGAAGTGAG TAGTGGGGTG CCAGACCAGG
+121 TGCGTCTGCC GCTGGATTGT GATAGGAAGC AGAGTGTTCC TGTGGTGAGC CGTGGCCGCC

```