```
SAP domain          hpl.pLb..h..b.h...G.K..L...l

S.c.   1   ----MADYSSLTVVQLKDLLTKRNLSVGGLKNELVQRLIK   36
C.e.   1   MVLSVADAKKLTVPRLKEELASRGLDNTGNKPELLARLTE   40
M.m.   1   MAAETVELHKLKLAELKQECLARGLETKGIKQDLINRLQA   40
H.s.   1   MATETVELHKLKLAELKQECLARGLETKGIKQDLIHRLQA   40

S.c.  37   DDEESKGESEVSP-----------QEQNQEQGSEPAAIEE   65
C.e.  41   SIEQAADELLGDAGITSDTHLLSDDILNDDILDAPSVGDE   80
M.m.  41   YLEDHAEEEANEE-----------------DVLGDETEEEE   64
H.s.  41   YLEEHAEEEANEE-----------------DVLGDETEEEE   64

S.c.  66   PASQNITEKKEVSSEPKETNEPKEENKD---VQKPSDGPS   102
C.e.  81   SEADKLLGSVDGAEEHKTDKSEGEEKHEVHLIDKKIDEVT   120
M.m.  65   PKPIELPVKEEEPPEKAVDMASEKK------VVKITSGIP   98
H.s.  65   TKPIELPVKEEEPPEKTVDVAAEKK------VVKITSEIP   98

S.c. 103   ATASENEQAAASTAAPALSPE------EIKAKALDLLNKK   136
C.e. 121   AETDKRARAIRFGLPVTAEDLGSDGAKAARAKRFDLPEDS   160
M.m.  99   QTERMQKRAERFNVPVSLESK-----KAARAARFGISSVP   133
H.s.  99   QTERMQKRAERFNVPVSLESK-----KAARAARFGISSVP   133

S.c. 137   LHRANKFGQDQ---------ADIDSLQRQINRVEKFGVDL   167
C.e. 161   KRLTSNDAKASRAERFGITPKTPAADDKLAARAARFGIPV   200
M.m. 134   TKGLSSDTKP-----------MVNLDKLKERAQRFGLNV   161
H.s. 134   TKGLSSDNKP-----------MVNLDKLKERAQRFGLNV   161

S.c. 168   NSKLAEELGLVSRKNEPESGNNGKFKNRNKN--ANNRSRV   205
C.e. 201   GGAASASASKNAPNSEGKAAKLAARAKRFGGEIDDEEMEA   240
M.m. 162   SSISRKSEDDEKLKKRKERFGIVTSSAGTG---TTEDTEA   198
H.s. 162   SSISRKSEDDEKLKKRKERFGIVTSSAGTG---TTEDTEA   198

S.c. 206   SKNRRGNRSGYRR   218
C.e. 241   KKKARMERFGTSN   253
M.m. 199   KKRKRAERFGIA-   210
H.s. 199   KKRKRAERFGIA-   210
```

Figure S2. Tho1 is a conserved nuclear protein that contains a SAP domain. Sequence alignment of *S. cerevisiae* Tho1 protein. Database searches revealed that Tho1 is conserved in a variety of higher eukaryotes: *S. cerevisiae* (P40040); *C.elegans* (Hypothetical protein Y53G8AR.6), *M. Musculus* (Q9D1J3) and *H. sapiens* (Hcc-1 protein P82979) and others not included here. The SAP-domain consensus is shown. h indicates hidrofobic residues); l, aliphatic; b, bulky; p, polar. Gray boxes indicate similarity, and open boxes identities.