

MSLRFWYGRSGSKTTLILNSVQEELERRRPRGKTLYLVPEDMTFQTQYEIKTPDVRGSIRAQVFSSRLAWRVLQHVGGMASRE
GAELFYAGASKLTCTHETDEKQAFFEPITFLVDPEQMTFQTQQALIGSEDVRGSIRAQVFSSRLAWRVLQEVGGASRI
MSLRFVIGRASGSKSTLCHEVQEELKQRPRGETILYLVPEDMTFQTQQALIGSEDVRGSIRAQVFSSRLAWKVVLQEVGGASRI
MGAEFLVGRS GSGKTKLIINSIQDELRAPFGKPIIFLVDPDMGTFLMEYELAKTPDMMGGMIRAQVFSSRLAWRVLQHTGGMSRP

YRYRDYAVLLRQPESEYVDWRTVFTDYNIPYFIDGRSMNHLVEFRSSLEVLKGNWRYEAVFRCVKTELLYPLNVRKETVRE
 YRYRDIAVLLRNGESYYDVMRTFTDYNIPHFIDEKRPMSHHLVECFIRSALETTSGNWRYDAVFRCVKTELLYPLDVRKETMRE
 YRYKDVAILARQPEYDKDMVKEFADYEIPYFIDGKASMLNHLPIEFIRSLDLVLKGNWRYEAVFRCVKTELLFPLNEPKAKVRE

594
593

RTWRDLTKKQCELFSYVWVERLAPLQRQLSSNRHYVVKQLQKVTFRVSIILSEHAKSSGFVPVGLVLPFGMGGTGLPPMTLE
LDIADLSIEEHFADAIETAPLRLRKEILLSSNRHFLKQLKQKIFRTSIIILREHAKSSGFVPVDLEVFPFGMGGTGLSPPMTLE
RTWADLSIKECEHLSAVVIEEIIAPLLQRQILLSSNRHFYLLKQLKQKIFRTSIIILREHAKSSGFVPVDLEVFPFGMGGTGLSPPMTLE
LDWRDLTKKECELFSSYDAVERLAPKLQKEILLSSNRHYVVKLEKLQKIVTRVSGILSEHAKASSGFVPIGLE..FGGKGGLPLPLT

[illegible]

Sequence logo for the 1000bp upstream region of the HNF1A gene. The top track shows the consensus sequence HNDIMVQVNLPLSLDE|EQE|FKKFKMKGLYLGDQEWRLMDTT|STGSSN|SAGLKKDGS|SSRS|SVSEEFNVLTQYVRRTF. The bottom track shows the corresponding amino acid sequence HNPIVEVKGDASEAETEKETILKKFKMKGLVLGDA|DVVRLMDNKLSTGSSDITISAGLKKDGSFSARSSIAS|EQEFNVLTQKYVHHTF. The logo indicates the relative frequency of nucleotides at each position, with positions 1-1000. The logo is color-coded to match the amino acid sequence below it.

ENIGKDKDITGVITDIAPYKKGNKAACTFCNFKSVCQFDESLEDNQFRLTKDMKDSEAMEKTIREEVGGE... 1171
 QEAGEQITDGRVSIIEPYKMKNKTPCTYCAFKSVCQFDESLEENYRPLKAEKDKTILWIKKEADGNEHS 1166

 non conserved
 similar
 $\geq 50\%$ conserved