Genomic TCF12 (HEB) locus accession number: NC_000075.6

**HEBCan Annotated Protein Sequence:**

**HEBCan-specific exons**

AD1 (1-99)

```
MFSPPVNSGKTRPTTLGSSQFSGSGMNPPQQRHAAIGTDKELSDLDFSAMDERRGTTSGQPSYD
SSRGFTDSPHYSDHNLDSRLGTHEGLSPTPFMNSNLIGKTSESFGSFLYSDSGLSGCQSSLLRDQLGLG
PAQLSSSGKPGTPYYSFSATSSRRPLHDSVALDPLQAKKVRPPGLPS
```

**HEBal-specific exon**

Alt exon (1-23)

```
MYCAYPVPGMNNSLMYYYNGKT
```

**Exons shared by HEBCan and HEBal**

AD3

```
VYAPSPNSDDFNRESPSYPSPKPTSMFASTFFMQDGHSSSDWLSSNGMSQPGFGLGTSTSHSMQSS
SYGSLHSHDRLSYSPPHSVPDINTSLLPPMSSFHRGSTSSSPYVAAHSTPINSDSLGLTRGNAAGSSQT
GDALGKALASIYSPDHTSSFSNFSPNSTPVGPSPLTGTSQMPRAGQAPSSPSYENSLHLQSRMEDRLDR
LDDAIHLRNHAGPVSTLSLPTSHDHLGSSNNLNSIGNLNSNYGGSLVTVNSRASVMTHEQDVSNLN
GNHSVLSSVTASSNTNHLKTPENFRGGSQSVQVESPEIKKENKDENLHEPPSSDDMSDESSQKD
bHLH (557-682)
IKVSSRGRSTSTTEDDLNPQKIEREKERRMANNARERLRVDINEAFKELGRMCQVHLKSEKPKQTLLI
LHQAVAVISLEDQQVRENNLNPKAAACLKRREEEKVSAASAEPPNTLPAGHPGLSESTNPMGHL
```

Fig. S1. Annotated amino acid sequence of HEB showing domain structures. Purple = activation domain 1 (AD1), light blue = activation domain 3 (AD3), pink = Alt exon, green = activation domain 2 (AD2), orange = basic helix-loop-helix (bHLH) DNA binding and dimerization domain. The amino acid numbers designated above each domain is based on the amino acid sequence of HEBCan exons, except for the Alt domain. Note that only the second half of the AD3 domain is contained within HEBal, and that the Alt exon is not contained within HEBCan.
Fig. S2. Validation of the specificity of a rabbit-derived antibody that detects a peptide within the Alt domain of HEBAlt. HEK293T cells were transfected with constructs expressing HEBAlt (ALT), HEBCan (CAN) or empty vector (HA). Anti-ALT was used to detect (A) overexpressed HEBAlt in HEK293T cells or (B) endogenous HEBAlt in Rag2-/- thymocytes using a secondary rat anti-rabbit antibody conjugated to horseradish peroxidase, followed by visualization using chemiluminescence.