

**Supplementary Table 1.** The positions of regulatory sequences in the *HOTAIR* locus

Position	Promoter (active)	CpGIs	Tandem repeat (strand+)	CTCF	Enhancer	DNase I hypersensitivity	Module and TSSs
54354858–54356533	No	No	No	No	No	No	<b>025604:</b> 54355865–54356303 <b>TSS:</b> <b>(chr12.11801):</b> 54354858–54354859
54356534–54359334	<b>HSMC cells:</b> 54359134–54359334	<b>Bonafide1432:</b> 54357217–54357921 <b>CpG3 (CpGProD):</b> 54357032–54358001 <b>CpG2 (CpGProD):</b> 54359256–54359334	<b>(AAGGGG)n:</b> 54358177–54358291	No	<b>HSMC cells:</b> 12 Weak enhancers <b>HMEC cells:</b> 20 Weak enhancers	<b>8:</b> 54358245–54358454 <b>HSMC cells DNase I hotspot: 66574:</b> 54357302–54358901	<b>025605:</b> 54358063–54358978
54359335–54362492	<b>HSMC cells:</b> 54359335–54361533  <b>NHEK cells:</b> 54362134–54362333  <b>Second active promoter based on Ensembl:</b> 54359491–54362492	<b>CpG18:</b> 54359659–54359906 <b>Bonafide1433:</b> 54359598–54360005 <b>CpG2.5 (WE):</b> 54360184–54360883 <b>CpG25:</b> 54360375–54360660 <b>Bonafide1434:</b> 54360202–54360827 <b>Bonafide1435:</b> 54362119–54362323 <b>CPG2 (CpGProD):</b> 54359334–54360945	<b>(AAAG)n:</b> 54359478–54359704  <b>(TCCCTCTC)n:</b> 54359986–54360112	54361413–54361642	<b>HMEC cells:</b> 19 Weak enhancers	<b>19:</b> 54359645–54359854 <b>16:</b> 54361465–54361654  <b>HSMC cells DNase I hotspot: 66575:</b> 54359619–54360710 <b>66576:</b> 54361172–54361793	<b>025606:</b> 54359632–54360527  <b>025607:</b> 54361760–54362456  <b>TSS:</b> <b>CHR12-M0409-R1:</b> 54361133–54361133
54362493–54363334	No	<b>Bona fide 1436:</b> 54362691–54362900	No	No	<b>HSMC cells:</b> 6 Weak enhancers	No	<b>025608:</b> 54362765–54363139
54363335–54364965	No	No	<b>L2C (within strand-):</b> 54363655–54363707	No	<b>HSMC cells:</b> 7 Weak enhancers	No	<b>025609:</b> 54364519–54364965
54364966–54370999	<b>HSMC cells:</b> 54365934–54370733  <b>NHEK cells:</b> 54367139–54369133 <b>First active promoter based on Ensembl:</b> 54365691–54370092	<b>Bona fide 1437:</b> 54366623–54367999 <b>CpG2 (WE):</b> 54366684–54366909 <b>CpG165:</b> 54366816–54369103 <b>CpG1 (CpGProD):</b> 54366456–54368740 <b>CpG2.4 (WE):</b> 54368334–54368964 <b>Bona fide 1438:</b> 54368166–54369840	<b>(ACCCC)n:</b> 54366647–54366670  <b>(GGCGGA)n:</b> 54367601–54367637  <b>(GGGA)n:</b> 54367731–54367801  <b>GAGGGAGGGAGC</b> <b>GAGA:</b> 54367742–54367783	54366799–54367314	<b>HEpG2 cells:</b> 6 Weak enhancers <b>HMEC cells:</b> 13 Weak enhancers <b>HSMC cells:</b> 20 Weak enhancers <b>NHEK cells:</b> 7 Weak enhancers, <b>4 Strong enhancers:</b> 54365934–54367133	<b>31:</b> 54366145–54366374 <b>41:</b> 54366785–54367814 <b>HSMC cells DNase I hotspot: 66579:</b> 54365947–54366518 <b>NHEK cells DNase I hotspot: 75095:</b> 54366045–54370999	<b>025615:</b> 54366091–54366249 <b>025610:</b> 54366634–54366977 <b>025613:</b> 54367707–54368584 <b>TSSs:</b> <b>CHR12-P0397-R1:</b> 54366912–54366912 <b>CHR12-P0397-R2:</b> 54367584–54367584

Exact position of this gene is chr12:54356092-54368740. For easiness, genomic region under analysis is divided into smaller portions. Positions are based on UCSC hg19.

TSS, transcription start site; WE, Weizmann evolutionary.