Figure S1. Primer walking for *RHOXF1* and *RHOXF2* identifies transcription start sites. (**A**) RNA from commercially available normal human testis was converted to cDNA and amplified using the same 3' primer and variable 5' primers designed to amplify *RHOXF1*. The mRNA is depicted by a red arrow and the amplicons are depicted by black lines. The names of each amplicon represent the annealing location of the variable 5' primer and are numbered relative to the translation start site. Data are normalized to beta-actin expression and plotted relative to the longest fragment. (**B**) Same as in part A but for *RHOXF2*.

Figure S2. Basal expression of *RHOXF1* and *RHOXF2* in non-testicular cell lines. *RHOXF1* (top panel) and *RHOXF2* (bottom panel) mRNA expression in the various cell lines (derived from the data in Fig. 2) plotted based on their delta Ct values x 10,000. Error bars represent SEM of a triplicate experiment.

Figure S3. *RHOXF2* but not *RHOXF1* expression increases proportionally with increased 5-AzaC treatment in MCF-7 cells. MCF-7 cells were treated with the indicated concentrations of 5-AzaC followed by qPCR analysis for *RHOXF1* and *RHOXF2*. Data are normalized to β-actin expression and plotted relative to the untreated (0μM) sample. Data are plotted on a log-scale.

Figure S4. *RHOXF1* is demethylated in response to 5-Aza treatment. Gene features and CpG residue locations relative to the translation start site are indicated for *RHOXF1* as in Fig. 4a. 293 and MCF-7 cells were treated 5-AzaC or not treated (Control) and subjected to bisulfite sequencing. Clonal data are shown as either methylated (black dots) or unmethylated (white dots) according to whether they had a C or a T at the CpG site, respectively. The regions analyzed are color-coded with lines under the gene and boxes around the dot-plots. Each region was assessed for percent methylation (number of methylated cytosines divided by total residues) and plotted as a color-coded pie-chart where the light color represents unmethylated cytosines and the dark color represents methylated cytosines. Values under each pair of pies represents the p-value according to Fishers exact test.

Figure S5. *RHOXF2* is demethylated in response to 5-Aza treatment. Gene features and CpG residue locations relative to the translation start site are indicated for *RHOXF2* as in Fig. 4b. 293 and MCF-7 cells were treated 5-AzaC or not treated (Control) and subjected to bisulfite sequencing. Clonal data are shown as either methylated (black dots) or unmethylated (white dots) according to whether they had a C or a T at the CpG site, respectively. The regions analyzed are color-coded with lines under the gene and boxes around the dot-plots. Each region was assessed for percent methylation (number of methylated cytosines divided by total residues) and plotted as a color-coded pie-chart where the light color represents unmethylated cytosines and the dark color represents methylated cytosines. Values under each pair of pies represents the p-value according to Fishers exact test.

Figure S6. *RHOX* methylation is elevated in sperm from infertility patients. (**A**) The average percent methylation as determined by pyrosequencing of two regions for *RHOXF1* (A and B) and two regions for *RHOXF2* (C and D) as well as *MEST* is plotted for the normal sperm group (white bars) and the abnormal sperm group (black bars). See Fig. 1A and Fig. 4 for region locations. Numbers represent p-values from a Mann-Whitney test comparing the normal group and the abnormal group. (**B**) Data from part A was normalized to the normal group to determine fold-change in the abnormal group (indicated above the bars).

Figure S7. Abnormal sperm patients have significantly higher *RHOX* methylation than normal sperm patients. (**A**) The abnormal sperm group (having at least one semen abnormality) was further subdivided as falling within the normal range or outside of the normal range for each individual semen parameter as defined by the WHO. Percent methylation as measured by pyrosequencing was plotted using the Tukey method. The p-values are indicated and are the result of a Mann-Whitney test. See Table 2 for N. (**B**) Same as in part A but for MEST.

Figure S8. Patients have vast hypermethylation of *RHOXF1*. (**A**) Two men from the normal sperm group and two men from the abnormal sperm group were analyzed by bisulfite sequencing for the degree of methylation within the *RHOXF1* locus. The gene features are indicated at the top and the CpG residues are indicated to-scale as red, vertical lines relative to the translation start site. Data from sequencing bacterial clones are shown as either methylated (black dots) or unmethylated (white dots). The regions analyzed are color-coded with lines

under the gene and boxes around the dot-plots. The regions that were detected by pyrosequencing are demarcated with black lines and the name of the region. The results of the pyrosequencing assay are indicated as percent-methylation (number of C's over number of T's incorporated into the pyrosequencing run) under the CpG residues that were assayed. (**B**) Pyrograms of the residues examined in the *RHOXF1* gene are shown with their peak strength and percent-calls and determined by the pyrosequencing software.

Figure S9. Patients have vast hypermethylation of *RHOXF2*. (**A**) Two men from the normal sperm group and two men from the abnormal sperm group were analyzed by bisulfite sequencing for the degree of methylation within the *RHOXF2* locus. The gene features are indicated at the top and the CpG residues are indicated to-scale as red, vertical lines. Data from sequencing bacterial clones are shown as either methylated (black dots) or unmethylated (white dots). The regions analyzed are color-coded with lines under the gene and boxes around the dot-plots. The regions that were detected by pyrosequencing are demarcated with black lines and the name of the region. The results of the pyrosequencing assay are indicated as percent-methylation (number of C's over number of T's incorporated into the pyrosequencing run) under the CpG residues that were assayed. (**B**) Pyrograms of the residues examined in the *RHOXF2* gene are shown with their peak strength and percent-calls and determined by the pyrosequencing software.

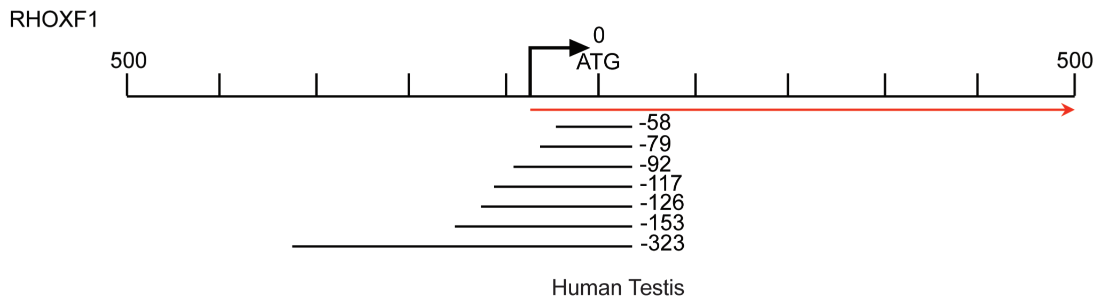
Figure S10. *NKAP* and *ZBTB33* expression in various cell lines after 5-AzaC treatment. Eight cell lines treated with 10μM 5-AzaC were analyzed for expression of *NKAP* (top panel) and *ZBTB33* (bottom panel) mRNA expression. Data were normalized to β-actin and plotted relative to the control-treatment. Error bars represent SEM of a triplicate experiment. Asterisk represents a p-value (students t-test between control and 5-AzaC treatment) $<.05$.

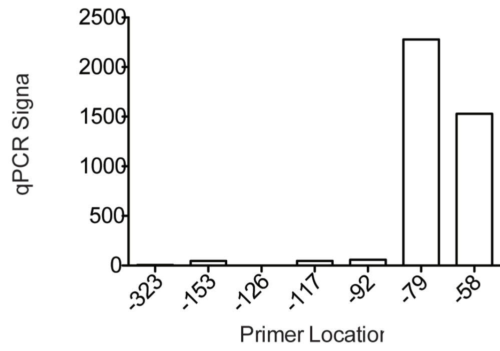
Figure S11. Genes neighboring the *RHOX* locus remain normo-methylated in infertile patients. The *NKAP* and *ZBTB33* genes are depicted at the top with the CpG residues represented by vertical red lines. The region analyzed by bisulfite sequencing is demarcated as horizontal black lines. Two patients from the normal group and two patients from the abnormal group were examined by bisulfite sequencing for the methylation status of *NKAP* and *ZBTB33*. The results of individual bacterial clones are shown as methylated (black dots), unmethylated (white dots) or undetermined (x) CpG residues.

Figure S12. Most infertile patients have hypermethylation along the entire *RHOX* cluster. (**A**) Overlap analysis was performed by categorizing patients as hypermethylated if their methylation scored higher than 95th percentile relative to the normal group. The number of patients exhibiting hypermethylation in each *RHOX* region is shown. Patients with hypermethylation at more than one region of the *RHOX* cluster are shown in the overlapping regions. (**B**) Regression analysis was performed by plotting the percent methylation, as determined by pyrosequencing, of the regions shown in all 140 patients.

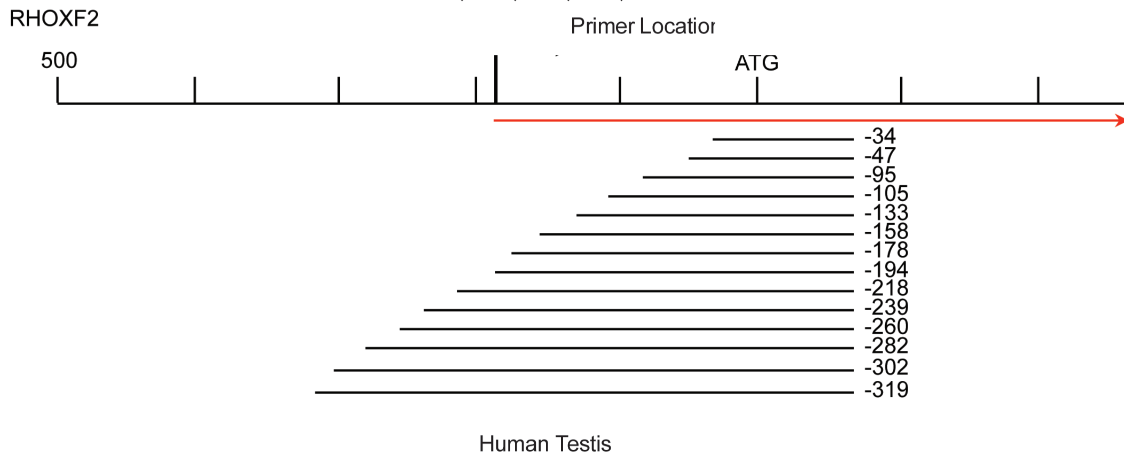
FIGURE S1

A





B



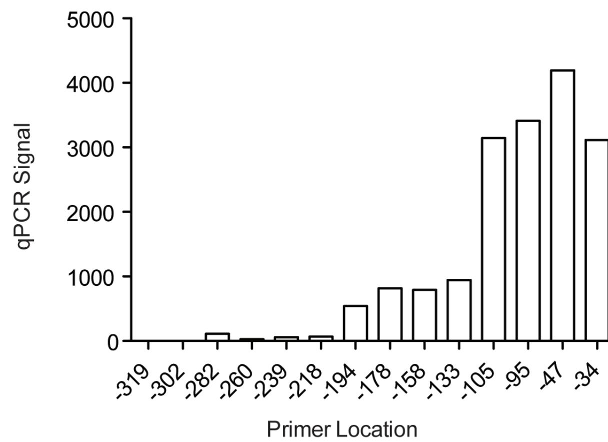
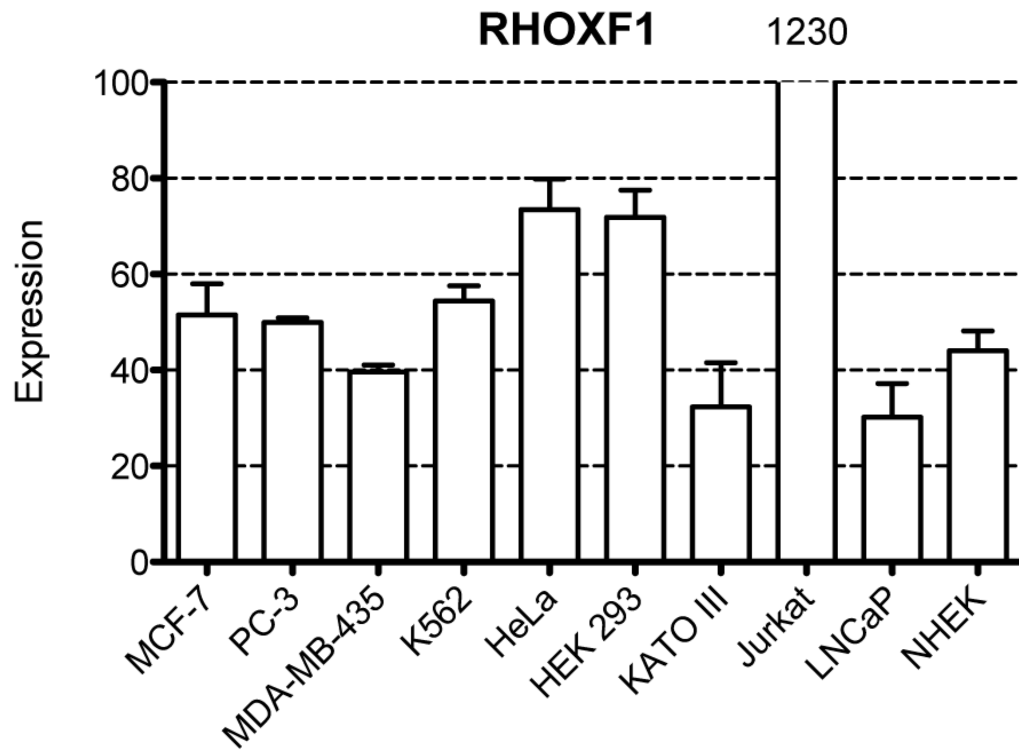


Figure S2



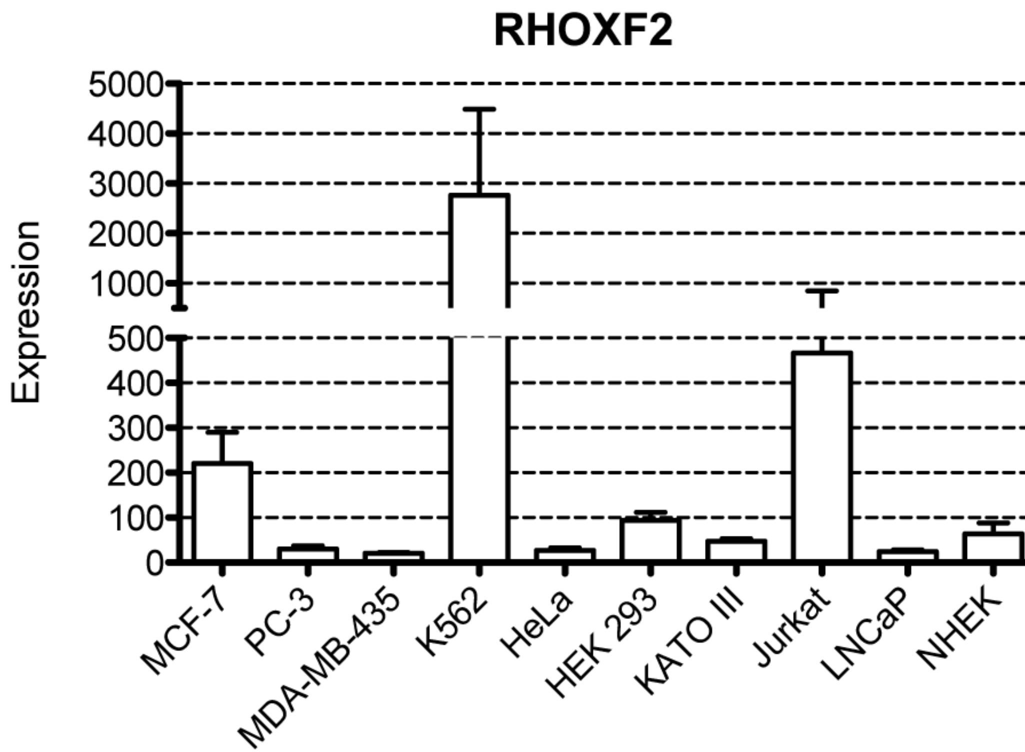


FIGURE S3

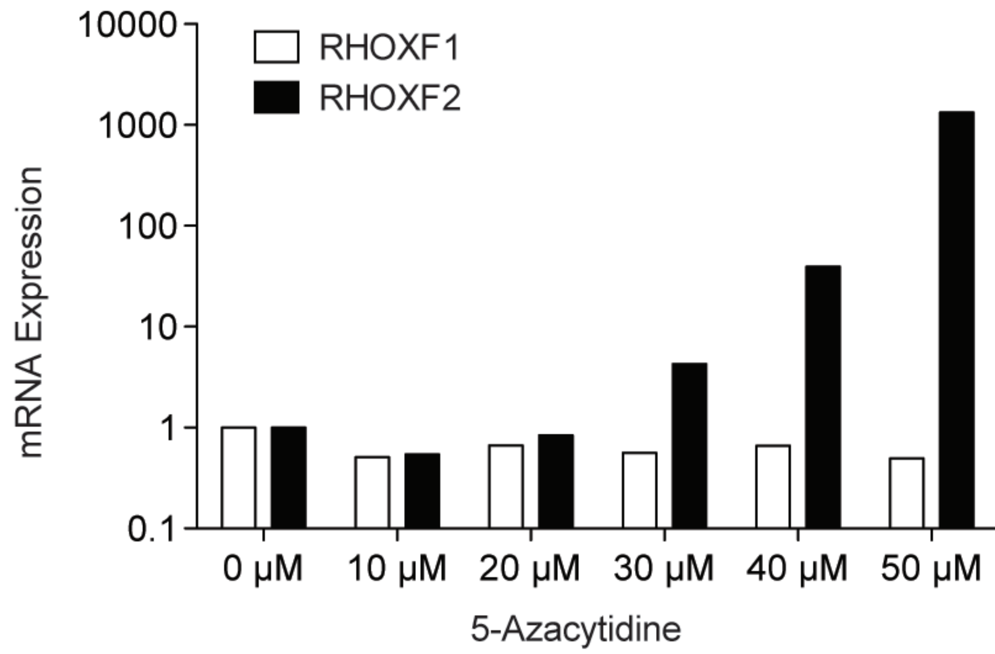


FIGURE S4

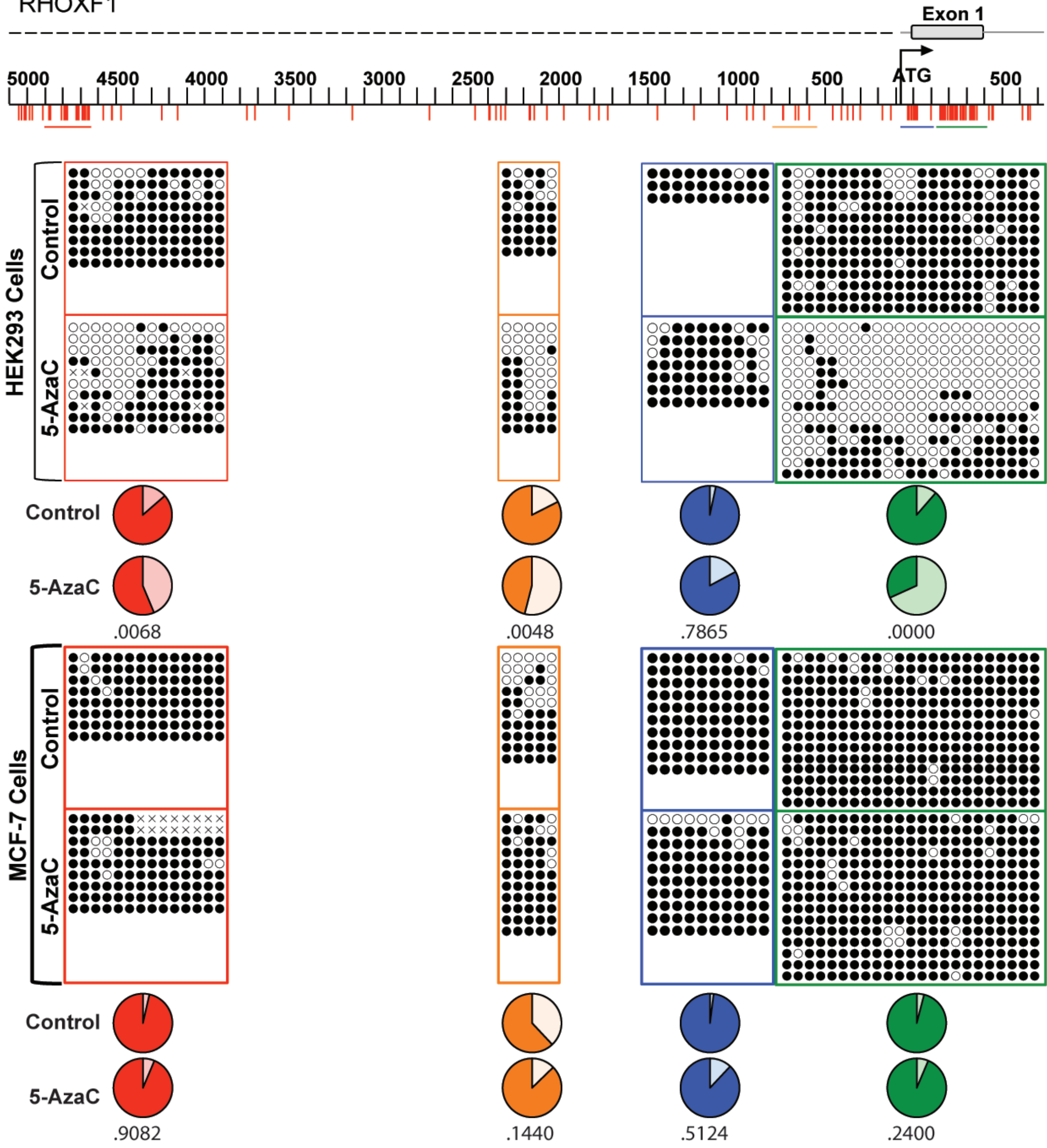
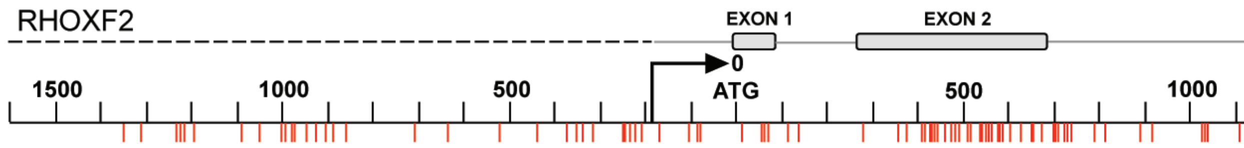
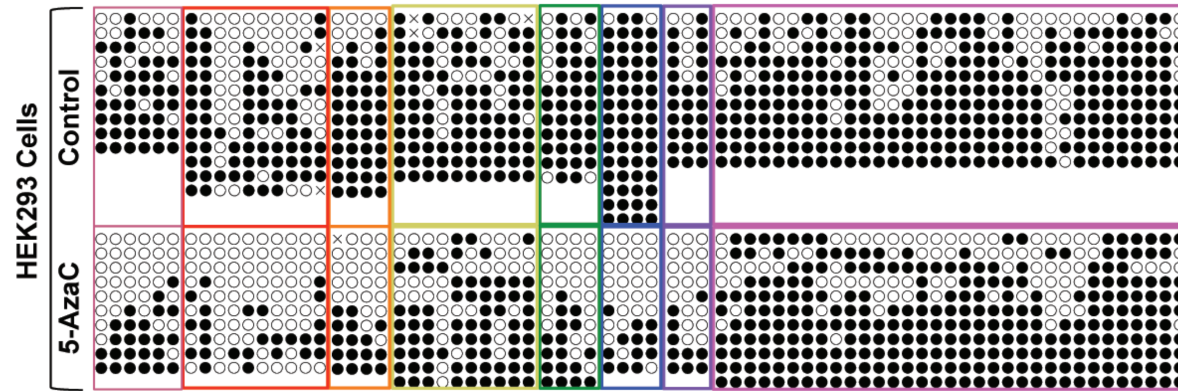
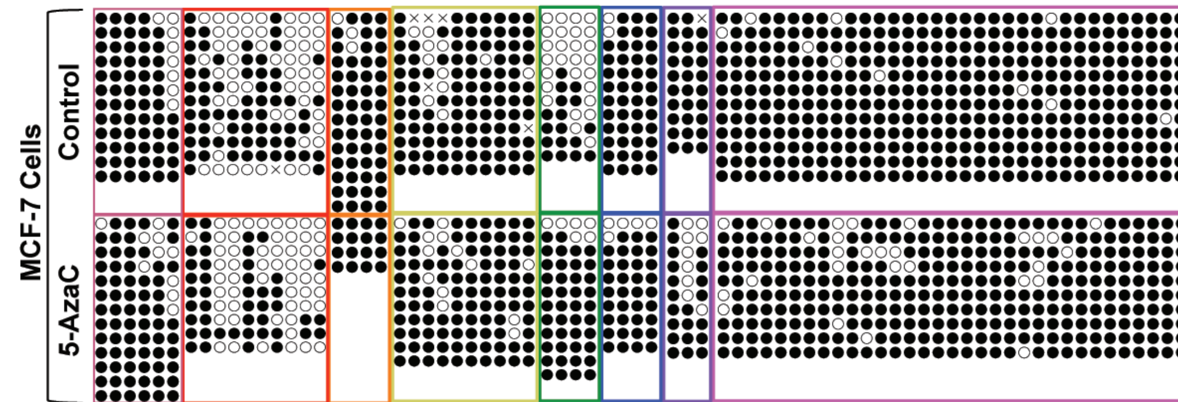


FIGURE S5







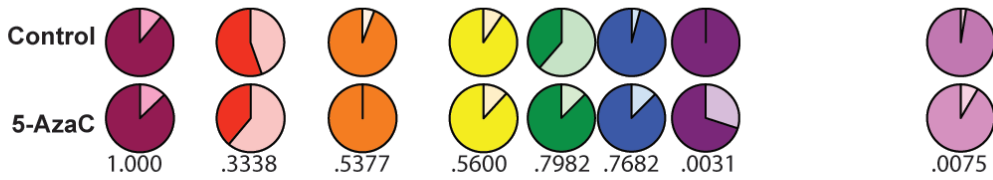
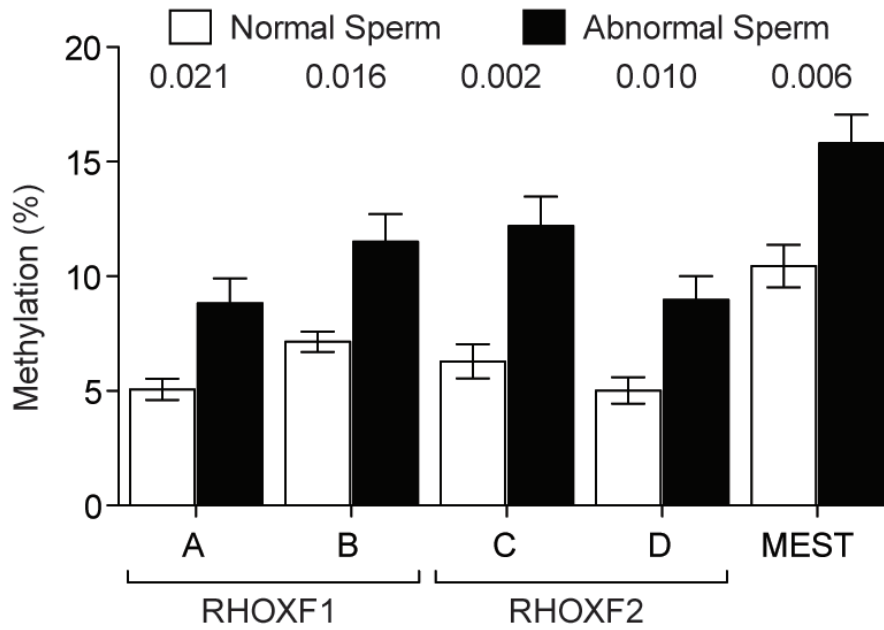


FIGURE S6

A



B

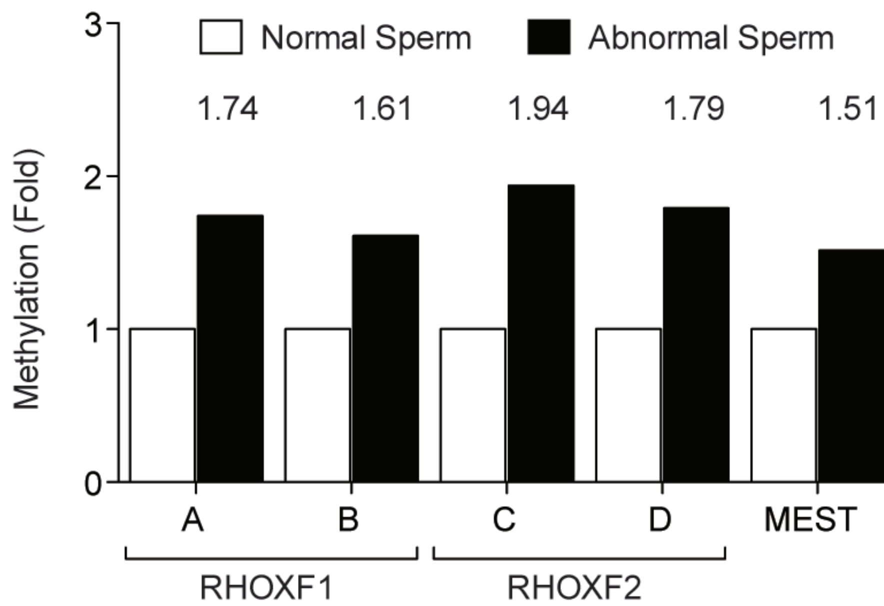
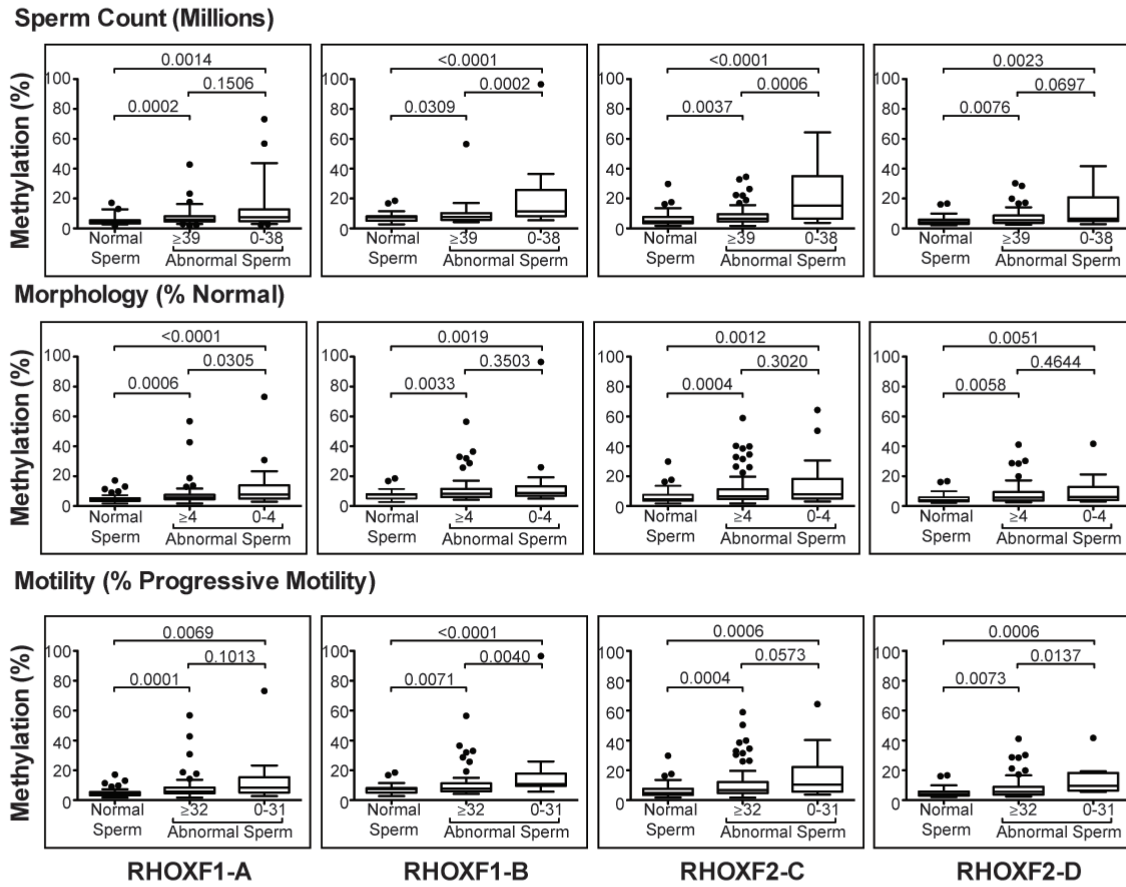


FIGURE S7

A



B

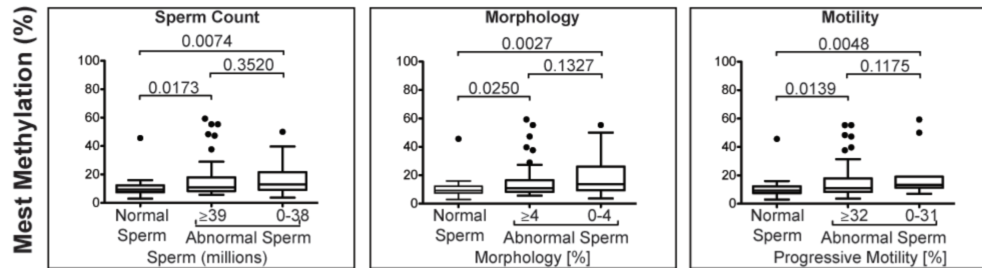
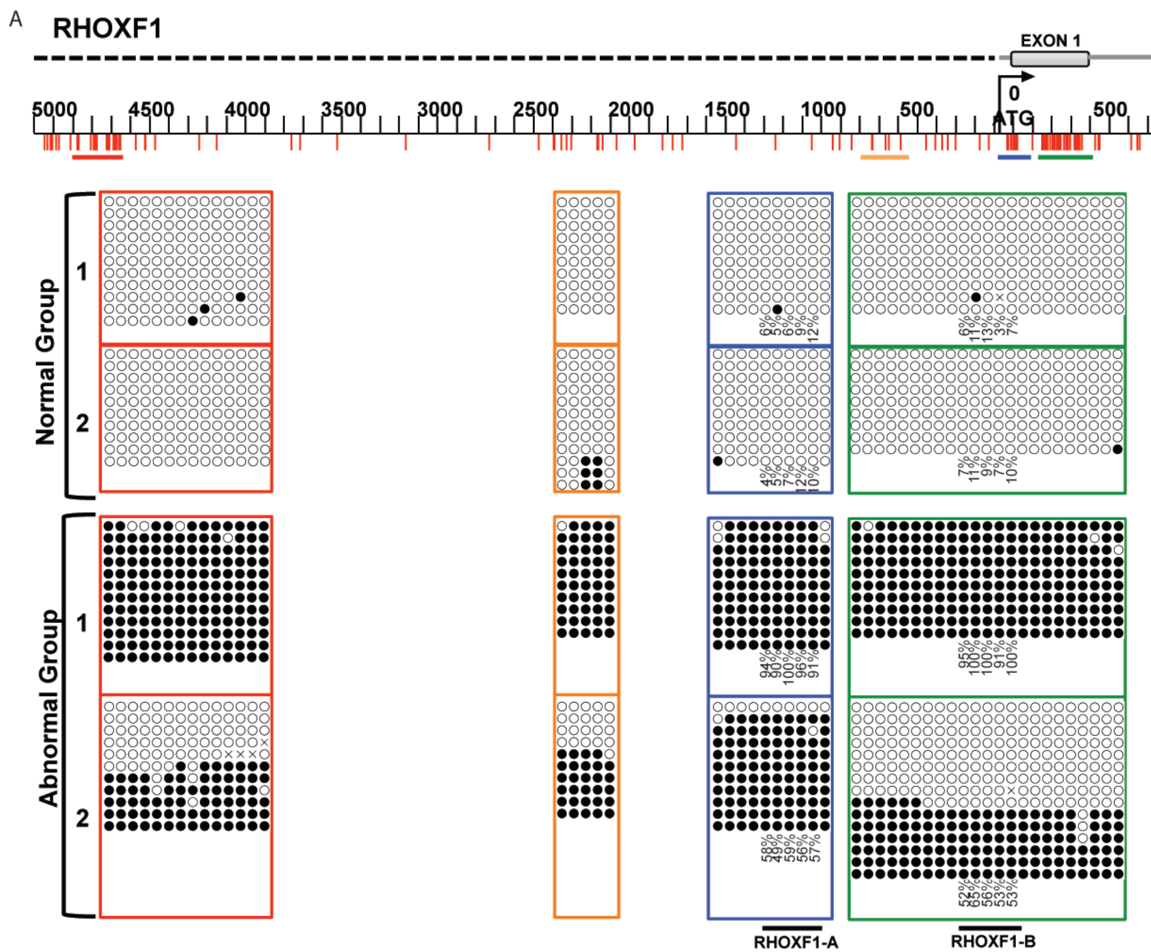


FIGURE S8

A



B

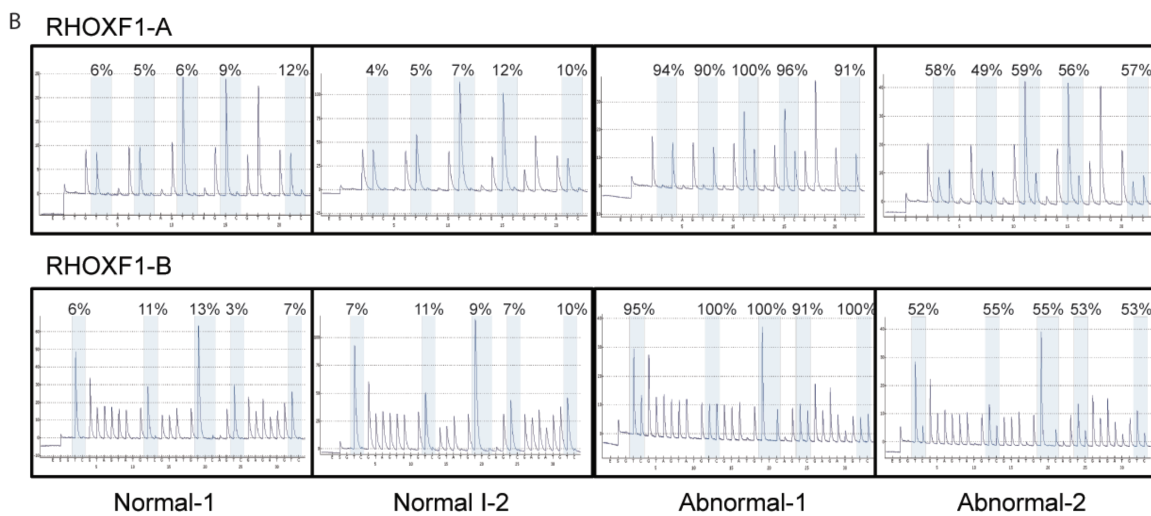


FIGURE S9

A **RHOXF2**

EXON 1 EXON 2

1500 1000 500 0 ATG 500 1000

Normal Group 1 2

4% 14% 4% 11% 3%

4% 14% 3% 14% 3%

Abnormal Group 1 2

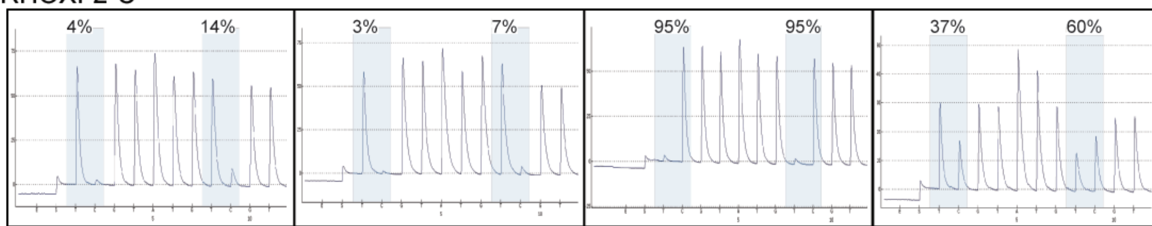
95% 95% 46% 84% 88%

37% 60% 36% 62% 60%

RHOXF2-C RHOXF2-D

B

RHOXF2-C



RHOXF2-D

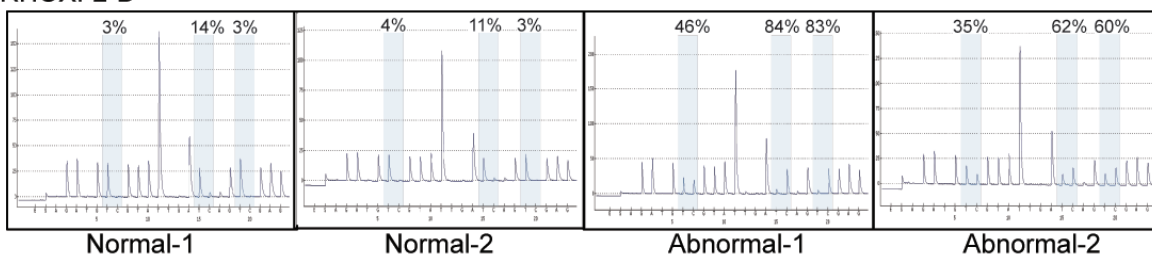
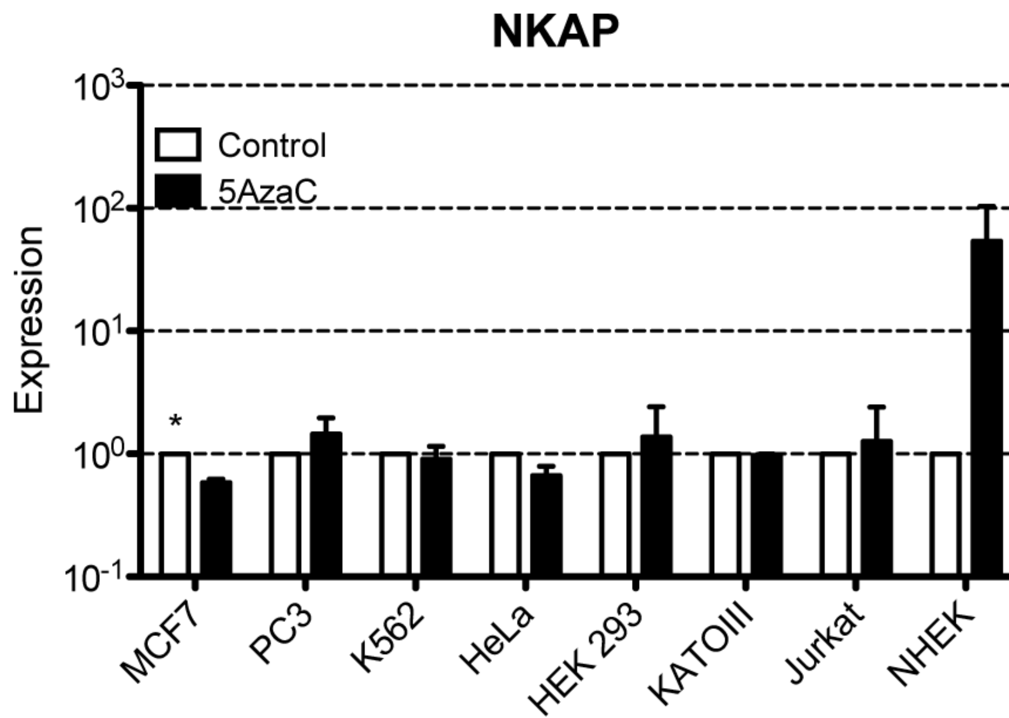


Figure S10



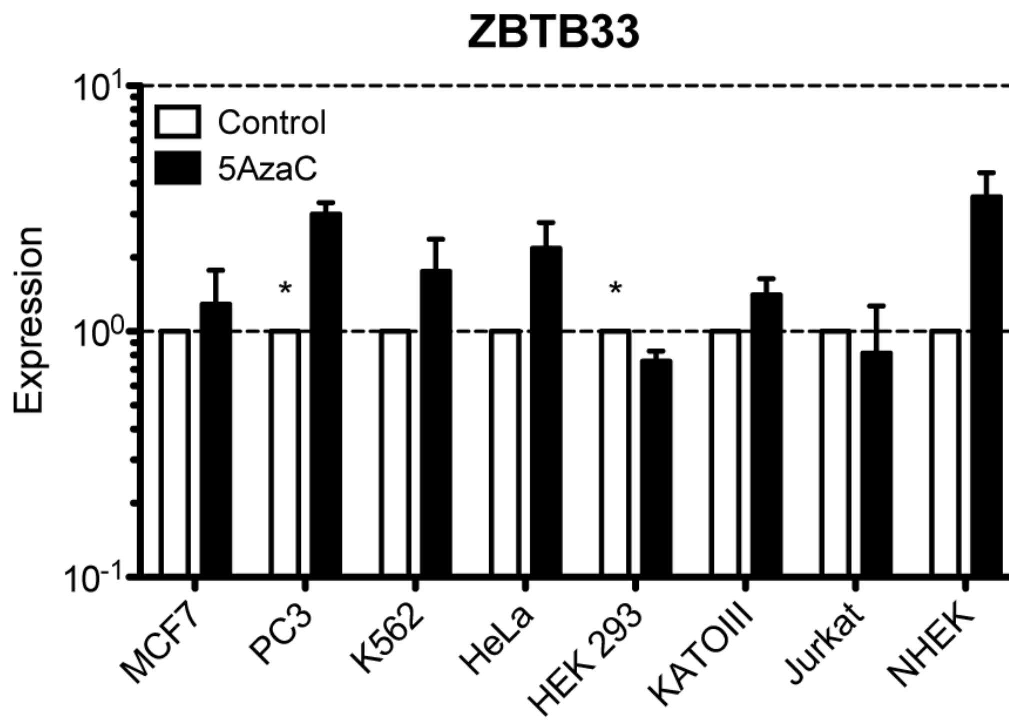
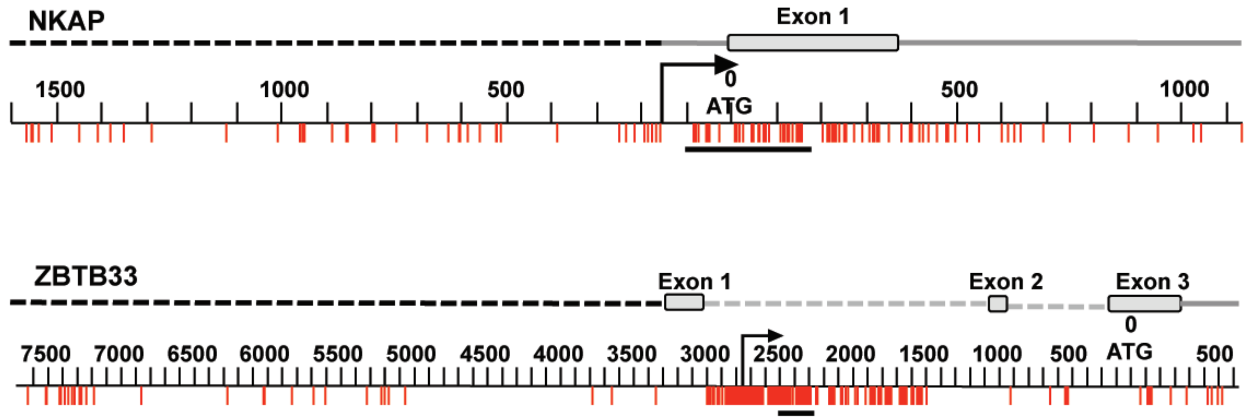


FIGURE S11



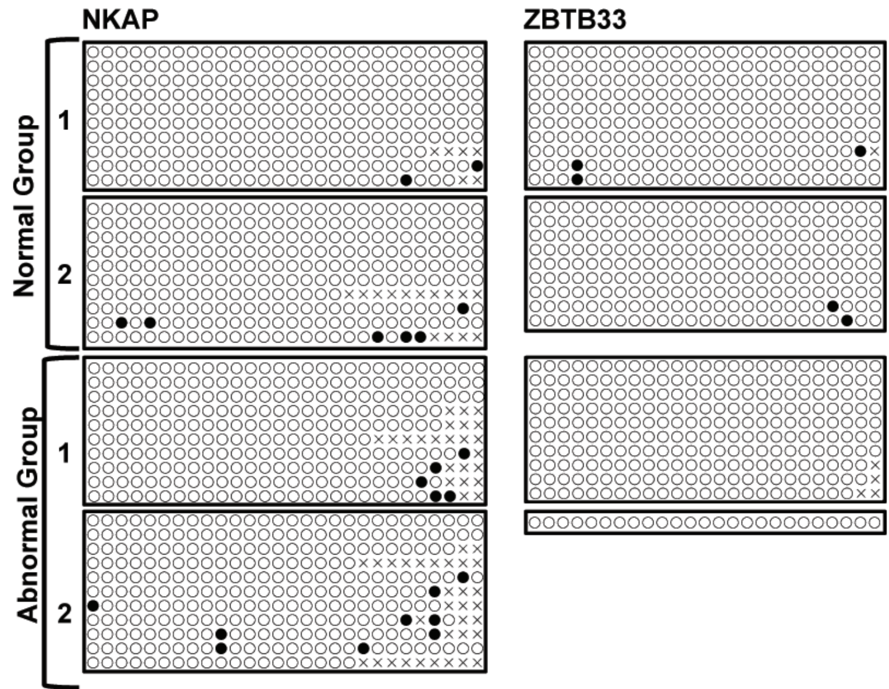
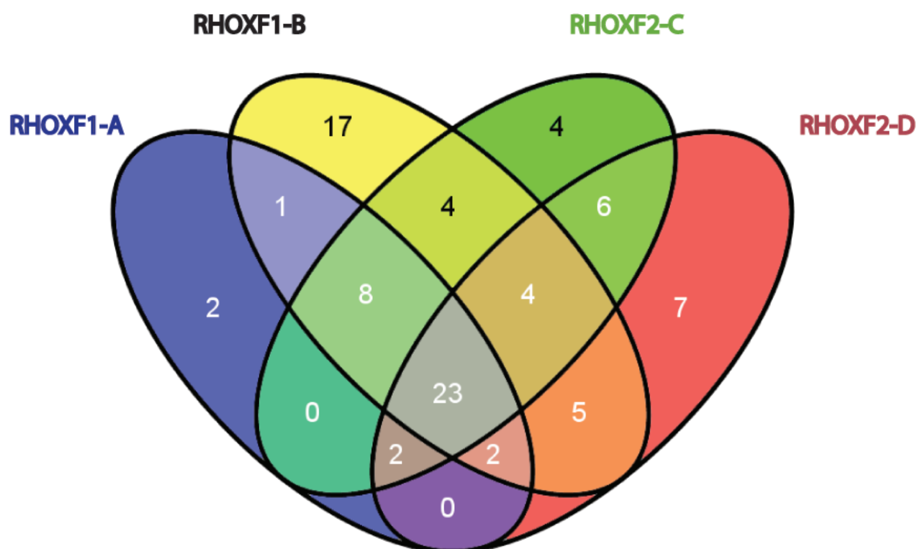


FIGURE S12



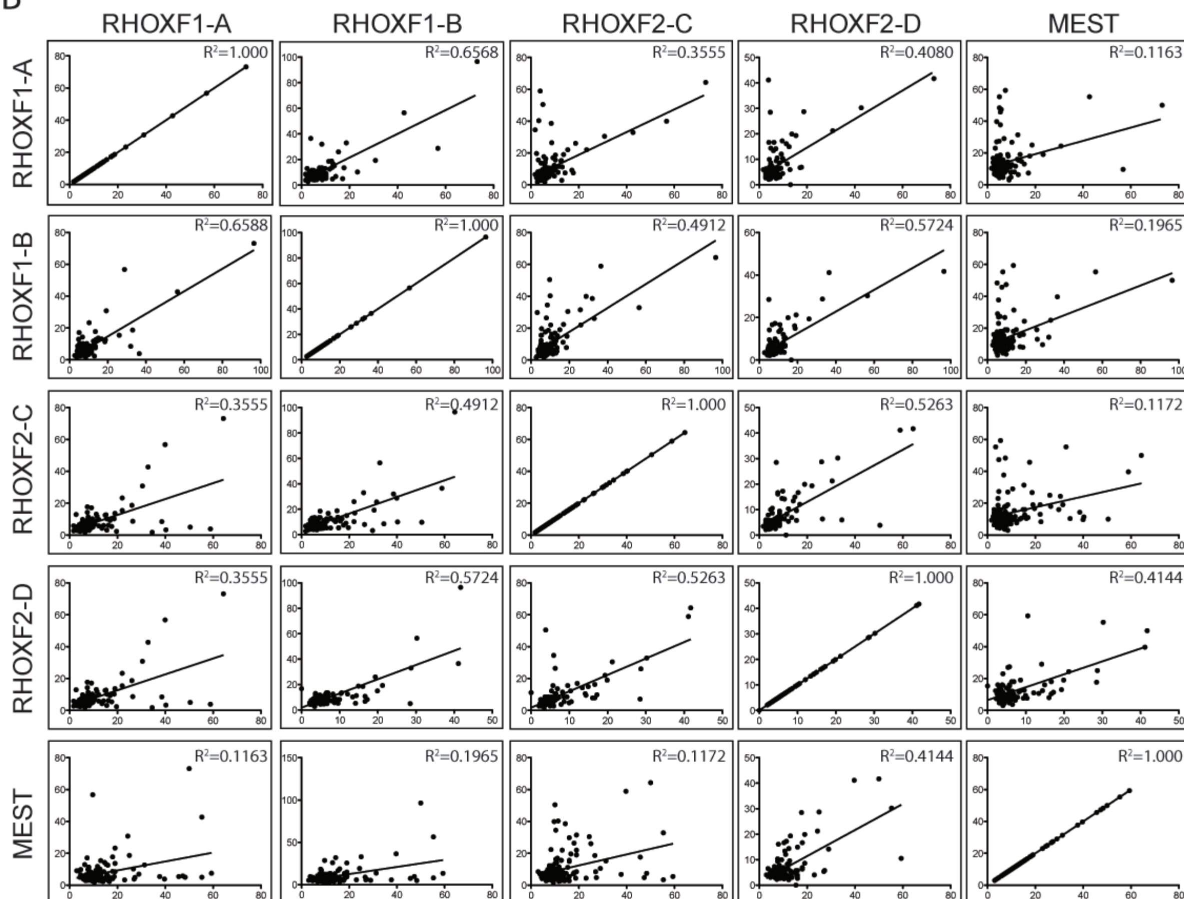


Table S1. Bisulfite Sequencing Primers

Assay	Gene	Location	Primer	Sequence	Amplicon Size (bp)	Number of CpGs Covered
	RHOXF1					
		-4910	Forward	AAATTGAAAGTTGTATAAAGATAAT	279	14
			Reverse	AATACTTAAACTTTTCTCAAAAAAC		
		-892	Forward	GAAGAGAGTATGGTAGTTTTGTTTTTAGTG	281	5
			Reverse	CAACATAAAACCCCAACTTACAACT		
		+83	Forward	ATATTAAGTATGAGAATAGGAATTAGTTT	199	10
			Reverse	CCTTCTACCCTTAATACTACCCCCAAC		
		+130	Forward	AAGGAGTTTTAGGTTTTATGGGTAAT	304	23
			Reverse	CTCCTAAAACAAACCACTTACCTTATA		
	RHOXF2					
		-1412	Forward	GTAGGAAAGAAGGGAAATTATTGTAG	248	6
			Reverse	AAACTACTAACCCCAAATAATCCAC		
		-1117	Forward	TAATAATTATTTAGGTGTGTTGGTG	229	10
			Reverse	AAAAACTTAACCTTACCTTAAACAC		
		-892	Forward	GTGTTTAAGGTAAGGTTAAGTTTTT	458	4
			Reverse	ATAAAAAAAATTTTACACTAACCCCA		
		-476	Forward	GGAGATTTAGGAAGTATGGGGTTAGTG	286	10
			Reverse	AAAACCTCCTCTCTTACTTTTCTACTTC		
		-171	Forward	GAAGTAGAAAAGTAAGAGAGGAGGTTTT	181	4
			Reverse	AAAACTCCATACCCTAAATAATATA		
		-15	Forward	TATATTATTTAGGGTATGGAGTTTT	103	4
			Reverse	AATACTAACCCTATAATTCTTTCTC		
		+64	Forward	AGAAAGAATTATAGGgttagtattttgatggg	266	3
			Reverse	CAAACTATACATCCTCTTCCTCCTC		
		+375	Forward	GAGGAGGAAGAGGATGTATAGTTTG	484	33
			Reverse	ACCAAACCCCAAATAAATTTTAATT		
	NKAP					
		-119	Forward	GTTTTTTTAGATTTGGGTTGTTT	300	28
			Reverse	CCAACTAATAAATAAATCCATTCC		
	ZBTB33					
		-2427	Forward	GGTTTTAATTAGGGAGGGGG	261	25
			Reverse	ATCTCTATAATAAATTCCCCAAACC		

Table S2-Primers for qPCR, Sequencing and Cloning

Gene	Assay	Location*	Primer	Sequence	Amplicon Size (bp)	Restriction Site
RHOXF1						
	Primer Walking	+35	Reverse	TAGAACACGGTGTCGTGGAC		
	Primer Walking	-58	Forward	AACCCCCACTCCAGCTGCA	93	
	Primer Walking	-79	Forward	CAGGAACTAGCTCCACCCTCT	114	
	Primer Walking	-92	Forward	CCAAGCATGAGAACAGGAACT	127	
	Primer Walking	-117	Forward	CCACTCATTGCAACTGTAACCAATACC	152	
	Primer Walking	-126	Forward	AAGCAAGTCCCACTCATTGC	164	
	Primer Walking	-153	Forward	TCCATTCCCTCTCTTGCTTG	188	
	Primer Walking	-323	Forward	CACAACCCCTACCCTCACC	358	
	Detection (qPCR)	+363	Forward	TGGAGGAGCTGGAAAGTGTT	75**	
	Detection (qPCR)	+2996	Reverse	ACCCGCACTTTGTCTTCAGT		
	Cloning	0	Reverse	GGATCCGGCTGGAGCGCTGCG		BamHI
	Cloning	-592	Forward	CCCAAGCTTAGTAGGGTGTAGTCAGCAAAGCT	592	HindIII
	Cloning	-1003	Forward	GGATCCATGAGTACAGGTTTCTCTTGCATAGCC	1003	BamHI
	Cloning	-2000	Forward	GGATCCAGTGAGTCATGCTTCGGATCTGTG	2000	BamHI
RHOXF2						
	Primer Walking	+16	Reverse	GGTCCGGAGGCTCCAT		
	Primer Walking	-34	Forward	AGAAGTGCACAGCTCCATCC	50	
	Primer Walking	-47	Forward	GCTCAGAATCTGCAGAAGTGC	63	
	Primer Walking	-95	Forward	GCGGCGACTCACAGTTCTA	111	
	Primer Walking	-105	Forward	AACGTTGCAGGCGGCGA	121	
	Primer Walking	-133	Forward	CACTGTTGCTGAGTCTAGACACC	149	
	Primer Walking	-158	Forward	AAGAGAGGAGGTCTCTAAGTGG	174	
	Primer Walking	-178	Forward	ATAGCGGGAAGCAGAAAAGC	194	
	Primer Walking	-194	Forward	GCCCAACCCACAGCAAATAG	210	
	Primer Walking	-218	Forward	CCTGACGCATGCGCACCT	234	
	Primer Walking	-239	Forward	TCATCCGGGCAAGGCCAA	255	
	Primer Walking	-260	Forward	ACCACCAACGCGTGCTCT	276	
	Primer Walking	-282	Forward	CCACTTACCCCTCCCCCC	298	
	Primer Walking	-302	Forward	CTCCCCCACACCCTACTACA	318	
	Primer Walking	-319	Forward	CCACGCCCTACTCCCCC	335	
	Detection (qPCR)	+1636	Forward	GGCAAGAAGCATGAATGTGA	103**	
	Detection (qPCR)	+4745	Reverse	TGTCTCCTCCATTTGGCTCT		
	Cloning	0	Reverse	CCATGGACCCTGAGTGGTGTGGATG		NcoI
	Cloning	-983	Forward	GGAAGATCTGCAAGGCCAAGCCCCTG	983	BglII
	Cloning	-793	Forward	GGAAGATCTAGTGGATACACTGTTACTGAATC	793	BglII
	Cloning	-711	Forward	GGAAGATCTCAGAGCGAGACTGTCTCA	711	BglII
	Cloning	-606	Forward	GGAAGATCTCCCAAGCTCCTGCTCCTG	606	BglII
	Cloning	-493	Forward	GGAAGATCTCCTCGCACTTTGATGTGG	493	BglII
	Cloning	-393	Forward	GGAAGATCTCCCAAGTTCCACACTCACG	393	BglII
	Cloning	-293	Forward	GGAAGATCTACCCTACTACACCACTTACC	293	BglII
	Cloning	-193	Forward	GGAAGATCTCCCAACCCACAGCAAATAG	193	BglII
Sequencing						
	pCpGL		Forward	CCACTGATTTTTGTTTATGTGAGCAAAC		
	pCpGL		Reverse	TAGAATGGGGCAGGGCCTTT		
Detection						
	NKAP		Forward	GCATCATTTGAATGCTCAGG		
	NKAP		Reverse	TGCAAGGGCTCTCTTCTCAT		
	ZBTB33		Forward	TTCTTCTTGTTGGGGGACTC		
	ZBTB33		Reverse	GTCCCTTCCCCTCATCTCTC		

*Locations are depicted relative to the ATG; **Based on spliced mRNA

Table S3. Pyrosequencing Primers

Assay	Gene	Location	Primer	Sequence	Amplicon Size (bp)	Number of CpGs Covered
	RHOXF1					
		Region A	Forward	ATATTAAGTATGAGAATAGGAATTAGTTT	199	5
			Reverse*	CCTTCTACCCTTAATACTACCCCCAAC		
			Sequencing	AGGGGAGTAGAGTTTTAGTTATG		
		Region B	Forward	AAGGAGTTTTAGGTTTTATGGGTAAT	304	5
			Reverse*	CTCCTAAAACAAACCACTTACCTTATA		
			Sequencing	GGTGGAAATTAGGAGT		
	RHOXF2					
		Region A	Forward	TTTTTTTTTATTAGGGTAAGGTTAAGT	101	3
			Reverse*	AACCTCCTCTCTTACTTTTCT		
			Sequencing	GGGTAAGGTTAAGTTTTTGA		
		Region B	Forward	GGGGGTTGGAGTTTGGTAA	137	3
			Reverse*	CCAACCAAAACAATAAACTTACC		
			Sequencing	TGTAGTTGTAGGAGTTG		
	MEST					
			Forward	TGTGGTTGGAGGTTTTGGGATTA	78	3
			Reverse*	CCAAAAACAACCCCAACT		
			Sequencing	GGTTTTGGGATTAGGG		
	NKAP					
			Forward	GGTAGGTTTAGGGAATTAAGGTAATAGATA	275	5
			Reverse*	CCACCCAACTAATAAATAAATCCATTC		
			Sequencing	AGAGTTAGAAGTTTAGTAAATTTGT		
	ZBTB33					
			Forward	GAGTGAGGTTTTTTTTTGTTGGGGGATT	117	4
			Reverse*	CTCCTTCCCCCCCTCCCTAATTAA		
			Sequencing	GTTGGGGGATTTTTAG		

*Indicates biotinylated primer; Locations are depicted relative to the ATG