

Table of Contents for Supplemental Material

Supplemental Table_S1 is provided as an .xlsx file.

Supplemental_Table_S2 is provided as an .xlsx file.

<u>File</u>	<u>Page</u>
Supplemental_Methods.pdf	2
Supplemental_Code.pdf	3-5
Supplemental_Table_S3.pdf: Association test p-values	6
Supplemental_Table_S4.pdf: Primer sequences for Sanger sequencing	7-8
Supplemental_Table_S5.pdf: Primer sequences for qRT-PCR	9
Supplemental_Fig_S1.pdf: Lesions in <i>Tet1</i> and <i>Birc6</i> fall into conserved protein domains	10
Supplemental_Fig_S2.pdf: Lesions in <i>Mre11a</i> , <i>Brca1</i> , <i>Rad50</i> , <i>Brca2</i> , <i>Fan1</i> and <i>Spin1</i> fall into conserved protein domains	11
Supplemental_Fig_S3.pdf: Sanger sequence traces	12-17

Supplemental methods

Mouse DNA: To obtain DNA, mouse tails were digested overnight at 55°C in tissue lysis buffer containing proteinase K, followed by purification with phenol chloroform, and precipitation with isopropanol. Resuspended DNA pellets were quantified using both a Nanodrop (Thermo Fisher) and Qubit (Thermo Fisher).

Sequence analysis: Genomic libraries were amplified by PCR using 10 cycles and 750 ng of the PCR-enriched library was hybridized with biotinylated probes that target exonic regions; the captured exome libraries were amplified by an additional 10 cycles of PCR. Exomic libraries were validated on a Bioanalyzer 2100 DNA High Sensitivity chip (Agilent Technologies) for size and by qPCR using the Kapa Library Quantification Illumina/ABI Prism Kit protocol (KAPA Biosystems) for quantities.

Sequence reads were mapped to the mm10 reference genome with *bwa*, and variants were called with *samtools* requiring a minimum SNP quality of 20 and a read depth ≥ 5 . Putative Insertion-Deletion (InDel) calls were realigned and SNP quality was recalibrated with GATK. Custom scripts combined these programs to annotate variants with respect to their predicted impact on protein sequence (Supplemental_Code).

Supplemental_Code.pdf

Custom scripts for calling detrimental variants:

```

==== project <Suppressor mutations screening in Mecp2-null mice>====
==== operating system and method <Linux clust and Linux commands see below>====
==== copyright 2020 Rong Zeng (rong.zeng@uhn.ca)====

### include <mouse reference genome:GRCm38 or mm10>

### include <69 Mouse lines:
520,654,751,4799,918_15L,J_57L,A_27R,A_333N,M_199LL,M_333R,136,137,591,722,933_15N,A_8
7N,A_120N,A_134L,A_83L,620,A_189N,A_173N,M_420L,M_79N,920_39R,S_165N,4633,J_136N,J_
157N,A_230RR,604,A_347N,272,747,752,4604,4730,904_16LL,905_20N,920_36LL,929_11R,934-
2N,J_21N,A_252LL,A_315N,A_322N,A_345L,A_210N,F_227N,F_245L,F_251L,J_3N,J_92N,J_129L,
M_351LL,M_453N,M_466LL,S_23RR,S_36R,S_158R,S_239L,S_301N,S_305L,U_45N,U_372R,U_67
2N,U_734LL,F_99N,F_311L>

### include Wellcome Trust Sanger Institute mouse SNP database <Newsnp.vcf>

### include <NGS tools: Fastx_toolkits,Fastqc,Trimmomatic,BWA,SAMtools,Picard,GATK,Annovar>

### include <script.js> adapted from the script by Pierre Lindenbaum at "L'Insitut du Thorax"

// function accept(ctx)

{
var y,g2;
var sampleList=header.getSampleNamesInOrder());

var g1=ctx.getGenotype("LINE");
/** ignore non-called */
if(g1== null || ! g1.isCalled() ) return false;
/** loop over the other samples */
for(y=0;y< sampleList.size();++y)
{
g2=ctx.getGenotype( sampleList.get(y) );

if(g2.getSampleName().equals(g1.getSampleName())) continue;
/** ignore non-called */
if(! g2.isCalled() ) continue;

/** is g1==g2 ? */
if( g1.sameGenotype( g2 ) ) return false;
}
/* found no other same genotype */
return true;
}
accept(variant);

```

====LINUX commonands for next generation sequencing data analysis, variant calling and filtering in ENU Mecp2/y mice project====

reads alignmnet <Forward reads>

bwa aln genome.fa ./LINE/LINE_R1.fastq.gz > LINE_R1.sai &

reads alignmnet <Reverse reads>

bwa aln genome.fa ./LINE/LINE_R2.fastq.gz > LINE_R2.sai &

SAM format <paired-end reads>

bwa sampe -f LINE.sam -r "@RG\tID:LINE\tLB:LINE\tSM:LINE\tPL:ILLUMINA" genome.fa
LINE_R1.sai LINE_R2.sai ./LINE/LINE_R1.fastq.gz ./LINE/LINE_R2.fastq.gz &

sort the SAM file and convert to BAM file

java -jar SortSam.jar SO=coordinate INPUT=LINE.sam OUTPUT=LINE.bam
VALIDATION_STRINGENCY=LENIENT CREATE_INDEX=true &

Check the head of LINE.bam

samtools view -H LINE.bam

flag possible PCR duplicates

java -Xmx25G -jar MarkDuplicates.jar INPUT=LINE.bam OUTPUT=LINE.marked.bam
METRICS_FILE=metrics CREATE_INDEX=true VALIDATION_STRINGENCY=LENIENT &

creates a table of possible indels

java -jar GenomeAnalysisTK.jar -T RealignerTargetCreator -R genome.fa -o LINE.marked.bam.list -I
LINE.marked.bam &

realigns reads around possible indels

java -jar GenomeAnalysisTK.jar -I LINE.marked.bam -R genome.fa -T IndelRealigner -targetIntervals
LINE.marked.bam.list -o LINE.marked.realigned.bam &

fixing the mate information

java -jar FixMateInformation.jar INPUT=LINE.marked.realigned.bam OUTPUT=LINE.fixed.bam
SO=coordinate VALIDATION_STRINGENCY=LENIENT CREATE_INDEX=true &

base quality Recalibration

java -jar GenomeAnalysisTK.jar -T BaseRecalibrator -I LINE.fixed.bam -R genome.fa -knownSites
SNPMOUSENEW.vcf -o recalibration_reportLINE.grp &

printReads

java -jar GenomeAnalysisTK.jar -T PrintReads -R genome.fa -I LINE.fixed.bam -BQSR
recalibration_reportLINE.grp -o outputLINE.bam &

Unified Genotypes from all lines < outputLINEs.bam has included outputline.bams from all founder
mouse lines in the study>

java -jar GenomeAnalysisTK.jar -glm BOTH -R genome.fa -T UnifiedGenotyper -I outputLINEs.bam --
dbnp SNPMOUSENEW.vcf -o thirtyonesnp.vcf -metrics snps.metrics -stand_call_conf 50.0 -
stand_emit_conf 10.0 -dcov 200 -A AlleleBalance &

```
# change specific LINE name in the script.js and filter variants of specific line comparing with all lines
java -Xmx50g -jar jvarkit/2014.10.15/dist/vcffilterjs.jar -f script.js ALLLINEsnp.vcf > LINE.vcf&
```

```
# splitting the genotypes of specific lines
java -Xmx25g -jar GenomeAnalysisTK.jar -R genome.fa -T SelectVariants --variant LINE.vcf -o
LINE.vcf -sn LINE &
```

```
# filter variants of specific line comparing with variants in the
java -Xmx25g -jar GenomeAnalysisTK.jar -R genome.fa -T SelectVariants --variant LINE.vcf --
discordance Newsnp.vcf -o LINEv.vcf &
```

```
# extract heterozygotes
java -Xmx4g -jar GenomeAnalysisTK.jar -T SelectVariants -R genome.fa --variant LINEv.vcf -select
'vc.getGenotype("LINE").isHet()' -o LINEhet.vcf &
```

```
# hard-Filtering
java -Xmx4g -jar GenomeAnalysisTK.jar -R genome.fa -T VariantFiltration -o LINEf.vcf --variant
LINEh.vcf --clusterWindowSize 10 --filterExpression "MQ0 >= 4 && ((MQ0 / (1.0 * DP)) > 0.1)" --
filterName "HARD_TO_VALIDATE" --filterExpression "DP < 5 " --filterName "LowCoverage" --
filterExpression "QUAL < 30.0 " --filterName "VeryLowQual" --filterExpression "QUAL > 30.0 && QUAL
< 50.0 " --filterName "LowQual" --filterExpression "QD < 1.5 " --filterName "LowQD" &
```

```
# convert the VCF file to the annovar file format:
annovar/convert2annovar.pl --format vcf4old --includeinfo LINEf.vcf > LINEt.annovar &
```

```
# annotate variant
annovar/annotate_variation.pl --buildver mm10 LINEt.annovar mm10db/ &
```

```
# convert annotated file to .txt file <exonic_region>
annovar/table_annovar.pl -otherinfo -buildver mm10 LINEt.annovar.exonic_variant_function mm10db/
-protocol refGene -operation g -nastring NA &
```

```
# convert annotated file to .txt file <exonic_intronic boundary region>
annovar/table_annovar.pl -otherinfo -buildver mm10 LINEt.annovar.variant_function mm10db/ -
protocol refGene -operation g -nastring NA &
```

Supplemental_Table_S3: Association test p-values for marginal gene effects and gene-gene interactions

Line	Gene	Longevity	Health	Limb Claspings	Activity	Muscle Tone	Weight
352	<i>Prdm15</i>	< 0.001	-	-	-	-	-
856	<i>Zzz3</i>	< 0.001	-	-	-	-	-
895	<i>Sqle</i>	0.330	-	-	-	-	-
	<i>Rbbp8</i>	< 0.001	-	-	-	-	-
	<i>Sqle</i> * <i>Rbbp8</i>	< 0.001	-	-	-	-	-
520	<i>Mark4</i>	0.580	0.138	0.191	0.365	0.840	-
	<i>Cd22</i>	0.960	0.037	0.015	0.325	0.646	-
	<i>Clint1</i>	0.270	0.498	0.212	0.663	0.755	-
	<i>Fnip1</i>	0.780	0.970	0.480	0.766	0.956	-
	<i>Kdm4a</i>	0.035	0.024	0.023	0.344	0.866	-
	<i>Mark4</i> * <i>Cd22</i>	nc	nc	nc	nc	nc	-
	<i>Mark4</i> * <i>Clint1</i>	0.129	0.714	0.419	0.048	0.116	-
	<i>Mark4</i> * <i>Fnip1</i>	0.900	0.572	0.180	0.453	0.396	-
	<i>Mark4</i> * <i>Kdm4a</i>	0.338	0.040	0.200	0.219	0.413	-
	<i>Cd22</i> * <i>Clint1</i>	0.042	0.344	0.694	0.048	0.031	-
	<i>Cd22</i> * <i>Fnip1</i>	0.830	0.383	0.725	0.295	0.140	-
	<i>Cd22</i> * <i>Kdm4a</i>	0.448	0.313	0.593	0.985	0.658	-
	<i>Clint1</i> * <i>Fnip1</i>	0.328	0.434	nc	0.902	0.652	-
	<i>Clint1</i> * <i>Kdm4a</i>	0.090	0.471	0.754	0.057	0.134	-
<i>Fnip1</i> * <i>Kdm4a</i>	0.310	0.193	0.927	0.076	0.066	-	
4654	<i>Gtf3c5</i>	0.030	0.974	0.055	0.010	0.014	0.460
	<i>Dennd4a</i>	0.170	0.549	0.273	0.132	0.326	0.313
	<i>Tm7sf2</i>	0.096	0.267	<0.001	<0.001	<0.001	0.251
	<i>Gtf3c5</i> * <i>Dennd4a</i>	0.492	0.105	0.535	0.109	0.858	0.351
	<i>Gtf3c5</i> * <i>Tm7sf2</i>	0.065	0.742	0.844	0.189	0.110	0.338
	<i>Dennd4a</i> * <i>Tm7sf2</i>	0.030	0.086	0.298	0.205	0.472	0.822
4751	<i>Paox</i>	0.240	0.495	0.222	0.134	0.199	-
	<i>Argef15</i>	<0.001	0.840	0.263	<0.001	<0.001	-
	<i>Spin1</i>	0.018	0.322	0.004	0.321	0.079	-
	<i>Paox</i> * <i>Argef15</i>	0.133	0.525	0.993	0.786	0.516	-
	<i>Paox</i> * <i>Spin1</i>	0.921	0.058	0.807	0.669	0.948	-
	<i>Argef15</i> * <i>Spin1</i>	0.909	0.415	0.123	0.278	0.100	-
4799	<i>Atp8a1</i>	0.170	0.004	0.003	0.007	0.005	0.552
	<i>Aacs</i>	0.011	< 0.001	< 0.001	nc	nc	0.556
	<i>Atp8a1</i> * <i>Aacs</i>	0.025	0.006	0.191	nc	nc	0.104
918_15L	<i>Celsr3</i>	0.002	0.002	< 0.001	nc	nc	0.332
J_57L	<i>Col24a1</i>	0.990	0.234	0.135	0.655	0.960	0.424
	<i>Apoa5</i>	0.059	0.769	0.001	0.019	0.002	0.061
	<i>Hcn2</i>	0.022	0.276	0.959	< 0.001	< 0.001	0.041
	<i>Col24a1</i> * <i>Apoa5</i>	0.190	0.706	0.622	0.703	0.202	0.007
	<i>Col24a1</i> * <i>Hcn2</i>	0.180	0.392	0.117	0.779	0.736	0.583
	<i>Apoa5</i> * <i>Hcn2</i>	0.555	0.433	0.149	0.058	0.043	0.169
A_27R	<i>Tet1</i>	< 0.001	0.460	< 0.001	< 0.001	< 0.001	-
	<i>Pelo</i>	0.950	0.337	0.132	0.691	0.968	-
	<i>Tet1</i> * <i>Pelo</i>	0.394	0.553	0.735	0.764	0.668	-
A_333N	<i>Shank1</i>	0.071	0.035	< 0.001	-	-	-
	<i>Fan1</i>	0.004	0.011	< 0.001	-	-	-
	<i>Shank1</i> * <i>Fan1</i>	0.782	0.730	nc	-	-	-
M_199LL	<i>Birc6</i>	0.320	< 0.001	-	-	-	< 0.001
	<i>Dbnl</i>	0.082	0.662	-	-	-	0.928
	<i>Birc6</i> * <i>Dbnl</i>	< 0.001	0.709	-	-	-	0.675
M_333R	<i>Zbtb41</i>	0.017	0.887	-	0.117	0.114	-
	<i>Mtor</i>	0.430	0.340	-	0.644	0.433	-
	<i>Zbtb41</i> * <i>Mtor</i>	0.939	0.106	-	0.861	0.728	-

nc: model failed to converge; - : phenotype not tested for association; The significant associations identified at the 5% significance level are in BOLD.

Supplemental_Table_S4: Primers used for Sanger Sequencing

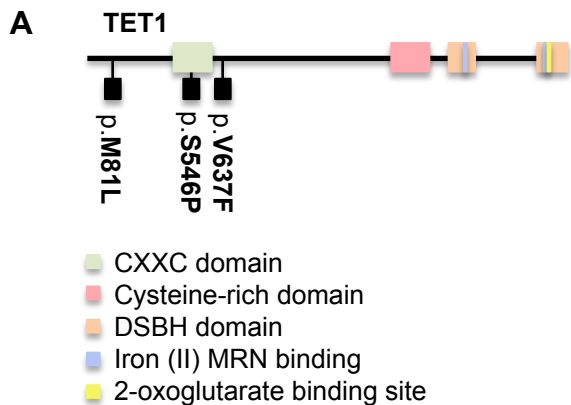
Line (Gene)	Forward Primer (5' to 3')	5' to 3' Reverse Primer
352 (<i>Prdm15</i>)	GGACAGGGCAAATTAAGCAA	CTCCTTCGACCACCATCAGT
856 (<i>Zzz3</i>)	GCCAGAGCTGTCCTCTATCCT	TGCTGACTGGCACTCTCT
895 (<i>Sqle</i>)	CTTGTATAAGGAGGTGCTTAATGGA	GAAATCCACCTGCCTCTGC
895 (<i>Rbbp8</i>)	CATTCTCTCTTGACAAGGTGACAG	TTTCTTTGTGGTGTGTCAGGCAGTTC
520 (<i>Cd22</i>)	TGCTGGGAAAGCTCATGTGTA	CCTCTTACCATAGGCCGCTT
520 (<i>Kdm4a</i>)	TTGAAAGGTAAGTGCACCCCG	CCAAGCACACAGTAGGCACT
520 (<i>Clint1</i>)	TGACAAGTCAGAAAAACCTGAAGTC	CTCGCTCCGTTACAACCTTT
520 (<i>Mark4</i>)	CTGTGGGTGGTGGCTTTTTTG	AAGTTCCCGTGGCTGTACC
520 (<i>Fnip1</i>)	ACTGCAGTGTGTAATGCACC	ACAACAGAAAGCAGTTCCAAGG
4654 (<i>Tm7sf2</i>)	AAGTCAAGGCAGAGTCGCAA	GCAGGCTGTAGGTGAATGGT
4654 (<i>Gft3c5</i>)	GCCAACCCCTTAGACTTCCA	CCCATTAGTGCTGTTGCCCA
4654 (<i>Dennd4a</i>)	AGTGCAGCAGAGACCAAGAGA	CTACAGAGCTGGCACTCACAC
4751 (<i>Arhgef15</i>)	GCACCTCCATCTTCTGGGAG	GAGCTGGTTCTGGGGAGATG
4751 (<i>Spin1</i>)	AGGCCAAGCCCTCTGTCTAT	GTTCCGAGGGACATCTAAGC
4751 (<i>Paox</i>)	CTCTGGAGGACCCTGAGCTA	AAGGAGCGGTTACCGAAGC
4799 (<i>Aacs</i>)	ATCCGTGGATAACCAGGCAAG	CCAGGTCTCCCCCTAGACAT
4799 (<i>Atp8a1</i>)	ACTTGGAGCTACAGCCATTGA	CGAAACCACTTGTCTCCAGC
918-15L (<i>Celsr3</i>)	TGTCATGAACTCTCCGGTGG	GCAGTCCACATACCATGCTG
J_57L (<i>Hcn2</i>)	TTCCTGCAGCGCCAATTC	CCGGTGACTCGAACTGC
J_57L (<i>Apoa5</i>)	AATAAGGCCCTGAGCAGACT	ATGAGTACCAGAAGCTGGGT
J_57L (<i>Col24a1</i>)	ACAAGGGCATTGGATATCTTGA	GGCACTTACCTTTAAGCCTTCT
A_27R (<i>Tet1</i>)	CTGTCAGACATGGGGCATCA	AGTCGCCAGCTAAGAGAGGT
A_27R (<i>Pelo</i>)	GTGAACATCCCCAGGAAACG	GGCTCCTCCCCATTCTAAAC
A_333N (<i>Fan1</i>)	TGCCAGTGTGGCTCACATAC	GCTAGGATGCAGGAAGCACT
A_333N (<i>Shank1</i>)	ATGAGGTCTAACCAGGCCCA	AGTTGGAGTGGAGAGGGCTT
M_199LL (<i>Birc6</i>)	TGCTTGGTGAAGTGCAGTGTA	TGCGAGGTGAGGTTGAATGT
M_199LL (<i>Dbn1</i>)	TCACTGGGGCAAGTGGTTCT	TCAAGGATTTACCAAGTAGCCA
M_333R (<i>Zbtb41</i>)	TGCAATGCGAAGTTTCCGTC	ATGAACGAGCTGCCATTTGC
M_333R (<i>Mtor</i>)	ACTGCTGTGCTCTATGACCTG	TCCGAGTTTCTGCCTCTGAT
J_157N (<i>Tbl1xr1</i>)	TGCAAGCAAACCTCCAGTGC	CCATGGGACATAAAGCACCC
S_165N (<i>NCoR1</i>)	ATGCAGGCCTGGTAGAACAT	AAAGCTAAAGGTCAGGTCTCTC
591 (<i>Tet1</i>)	GCAGCCGAGAATGAAGTCCA	GCGGCGTAGAATTACATGGG
591 (<i>Arhgap15</i>)	CAGAAATGAACCGAACCCCC	CCAGAACAGGCTGCAAAACAA
591 (<i>Anapc5</i>)	TGGCTGTCAGTCTTCCTTCT	TCTTTCCCGCCCATTACCG
591 (<i>Mmadhc</i>)	CATTTCTCCTTTCTCCAGTTTTT	AGTGCCAGCTGCTAGTTTTCC
591 (<i>Hat1</i>)	TGTGTGATACCTGTTGGTTTTGTC	ATGGAGGAGGGTTCCAAAAGG
137 (<i>Atp6v1b1</i>)	CAGAGGTTCTCACGCTCCTT	GGACTGATTTAGAGCGCCGT
137 (<i>Grm8</i>)	GGAGAAAGAGGGGTGCCTTG	GGGTGGGTCTCCATTAGCAC
722 (<i>Bscl2</i>)	CTGGAAGCACACTCAGTGA	CAGAGGCAGACCTAGAACCC
722 (<i>Kif21a</i>)	TCCACGCCATGAACAAAGAG	CTGTTCCAAGACATCCAGCCC

722 (<i>Adcy6</i>)	GACCTGTGCCTCTGTCCTTTT	TGATGTAGACGAAGAACACGGG
722 (<i>Map3k12</i>)	GCAGGATGAGACAGGTACAGAG	TTTTGTCAGTCCAGTCTCCCA
933-15N (<i>Foxn4</i>)	TGGAACAGGTTTCTGCCTCC	TATATGGACATGCCCCCAGAA
933-15N (<i>Parp1</i>)	TCAGCCCCGGTGTCTGTAAT	GGAAATTTGGGAGCTCACGG
933-15N (<i>Cox16</i>)	TGCGCTGAGTCTACAACCTC	CCCTAGTCAGCAGTATCCGA
933-15N (<i>Tdp1</i>)	ACAGTGCGAATCTGTCCAAG	GGATTGGAATCCTCACCACG
A87N (<i>Psmb2</i>)	AGCGATTTGGGTGAAGACTGT	GCTTTCTGCTGTTGAAGGCTG
A87N (<i>Emcl</i>)	CTTCCCAGAGTCTGAACCCC	CAGTTCAAGGACCTACAGGCA
A87N (<i>CD37</i>)	TTGTGAGTCTACTCATGGCTACC	AGCCCTACCTGGGAAGATAGTA
A87N (<i>Pcf11</i>)	ACGCCTGTCTGGGATTATCAG	CCAAAGACCAAGCCCCGAGT
A87N (<i>Cep72</i>)	CTAGCTCAGGCCAAGCATGT	ACTTGTCTTCCCCTTGTCCC
A120N (2310022B05Rik)	AGGTAAATTGAGGTCGGGGC	ACGGGGAAAACCTGGCAAC
A120N (<i>Unc79</i>)	TCCAATCTGTCAATTTCTGAGGC	CACCACTCTCCTGTCTCCTGA
A120N (<i>Eif2b5</i>)	GGTCTTTCTGGTCTTAGGGCA	GGTACTGGACTCTGGACAACCT
A120N (<i>Fanca</i>)	GAGCATGCTGGGTAAGAGACT	TAACCGGAACGTCCGTCAC
A120N (<i>Sidt2</i>)	TTGTCAAGGTGACCTCCAAGAA	TTTCTCAGGCAGCACAAGGT
A134L (<i>Scn11a</i>)	ACTTGCTGGGGAAATCCACT	GCCAGCGTGCATACTTATCTT
A134L (<i>Gm14496</i>)	TGGCTTCTTGTGCAGGGTTG	TAGGCTGGACTCGAACTCAGAA

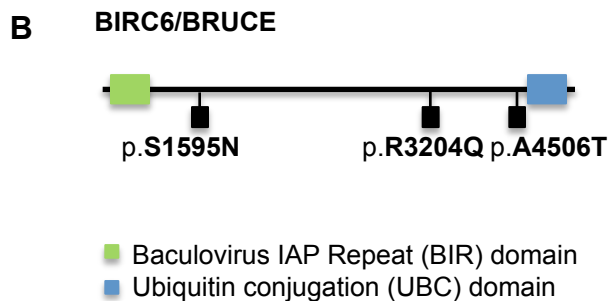
Supplemental_Table_S5: Primers Used for qRT-PCR

Gene	Forward Primer (5' to 3')	Reverse Primer (5' to 3')
<i>Tbp</i>	CCTTGTACCCTTCACCAATGAC	ACAGCCAAGATTCACGGTAGA
<i>Rbbp8</i>	AATGGTCAACAGGATCAAGTAGC	GTAGCCGGTTAATGCCAGAAAA
<i>Mre11a</i>	CCGAGAAAGCAGACAGAGAAA	AAAGGCTGAGGTGGAGGTCT
<i>Rad50</i>	GGTGAACAAGGACCTCGACA	CCGCCAAAGATCACGAATAA
<i>Nbn</i>	CCCATGTGGAAGCTGCTC	AGTTTTTCCTCCCAACAACG
<i>Atr</i>	GAATGGGTGAACAATACTGCTGG	TTTGGTAGCATACTGGCGA
<i>Cdk1</i>	GACAGAGAGGGTCCGTCGTA	TGGCCAGTGACTCTGTGTCT
<i>Rpa2</i>	ATGGATGTTTCGTCAGTGGGT	AAAGGCCACCAAGCTCTTTT
<i>Xrcc5</i>	AGATTGCCAGCTTTGAGGAA	GCTTTGATGCAGTCCATACTTTT
<i>Trp53bp1</i>	TGGAGTCTTCTCATTTGGGT	GCAGCTTCTACTGTCACTGCC
<i>Prkdc</i>	AAGGATGTATGCTGCCTTGG	ACCACCATTCCCAAAGCTC

Supplemental_Fig_S1.pdf



	p.M81L	p.S546P	p.V637F
RAT	-TRAGAARLNQDRNQV-	-EGGISHV S M T SSADD-	-EKKPK V LK T DFNN-
MOUSE	-TRAGAAR M NRDRNQV-	-EGEISRV S ITGSA-D-	-EKKPK V LK T DFNN-
HORSE	-TRAGAAR M NLDRTEV-	-EEESSQ I SK T SMVD--	-EKKPK V LK A DFDN-
COW	-TRAGAAR M NLDRTEV-	-EGGSSQ I S A TRLVD--	-EKKPK V LK A DFDN-
CAT	-TRAGAAR M NLDRTEV-	-EGGSSQ I S V TSMVD--	-EKKPK V LK A NFDN-
DOG	-TRAGAAR M NLDRTEV-	-EGESSQ V S M T S VVD--	-EKKPK V LK A DFDN-
HUMAN	-TRAGAAR M NLDRTEV-	-DRGSSQ V S V T S T V HV-	-EKKPK V LK A DFDN-
CHIMP	-TRAGAAR M NLDRTEV-	-DRGSSQ V S V T S T V HV-	-EKKPK V LK A DFEN-

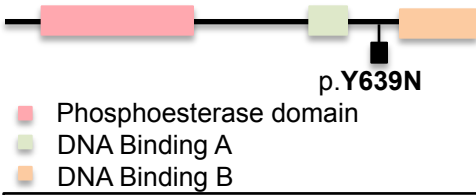


	p.S1595N	p.R3204Q	p.A4506T
XENOPUS	-DDASTFTV S TFGVTP-	-LQATPPH R RARSAAW---	-TAEIVS A ATTSLRQ-
RAT	-DDASTFTV S SFGVTP-	-LQATPPH R RARSAAW---	-TAEIVH A ATANLRQ-
MOUSE	-DDASTFTV S SFGVPP-	-LQATPPH R RARSAAW---	-TAEIVH A ATTSLRQ-
HORSE	-DDA---- M S S FGVTP-	-LQATPPH R RARSAAW---	-TAEIVY A ATTSLRQ-
COW	-DDA---- M S S FGVTP-	-LQATPPH R RARSAAW---	-TAEIVY A ATTSLRQ-
CAT	-DDASTFTV S SFGVTP-	-LQATPPH R RARSAAW---	-TAEIVH A ATTSLRQ-
DOG	-DDA---- M S S FGVTP-	-LQATPPH R RARSAAW---	-TAEIVH A ATTSLRQ-
HUMAN	-DDA---- M S S FGVTP-	-LQATPPH R RARSAAW---	-TAEIVY A ATTSLRQ-
CHIMP	-DDA---- M S S FGVTP-	-LQATPPH R RARSAAW---	-TAEIVY A ATTSLRQ-

Supplemental_Fig_S1_Legend: Amino acid alignments generated using Clustal Omega show that the individual amino acids as well as their surrounding sequence is well conserved among organisms. (A): Protein structure of TET1: the missense mutations cluster around the CpG binding domain. (B) Protein structure of BIRC6: The mutations occur throughout the protein, but not in known functional elements.

Supplemental_Fig_S2.pdf

A MRE11



p.Y639N

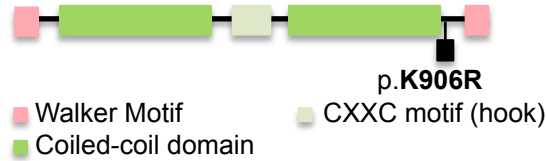
XENOPUS	-RNMPKKT Y SEDIEDD-
RAT	-RNVATKN Y SETIEVD-
MOUSE	-RNVAPKN Y SETIEVD-
HORSE	-QDVATKN Y SEVMEVD-
COW	-RNVATKN Y TEVIEVD-
CAT	-RNVATKN Y TEMIEVD-
DOG	-RNVATKN Y SEVIEVD-
HUMAN	-RNVTTKN Y SEVIEVD-
CHIMP	-RNVTTKN Y SEVIEVD-

D BRCA1



p.K50E		p.F227I
XENOPUS	-FCKFCML Q LLSKKKKE-	-SD Q AEECFSE F SETAERT-
RAT	-FCKFCML K LLN Q KK-G-	-AKKAACD F SE-GIRN-
MOUSE	-FCKFCML K LLN Q KK-G-	-AEEAACE F SE-GIRN-
HORSE	-FCKFCML K LLN Q KK-G-	-TKKAACE F SEKAITN-
COW	-FCKFCML K LLN Q KK-G-	-FFPAACE F SEKDITN-
CAT	-FCKFCML K LLN Q RK-G-	-AKKASCE F SE-DITN-
DOG	-FCKFCML K LLN Q RK-G-	-AKKAACE F SG-DITN-
HUMAN	-FCKFCML K LLN Q KK-G-	-AKKAACE F SETDVTN-
CHIMP	-FCKFCML K LLN Q KK-G-	-AKKAACE F SETDVTN-

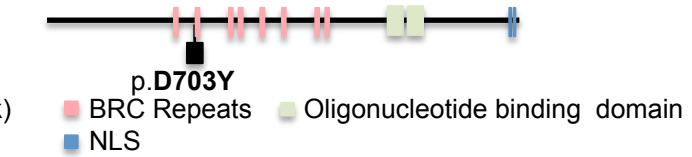
B RAD50



p.K906R

XENOPUS	-QCLSREI K EAKEQVF-
RAT	-QSLNREI K DAKEQIN-
MOUSE	-QSLNREI K DAKEQIS-
HORSE	-QSLYREI K DAKEQVS-
COW	-QSLHREI K DAKEQLS-
CAT	-QSLFREI K DAKEQLS-
DOG	-QSLSREI K DAKEQLS-
HUMAN	-QSLYREI K DAKEQVS-
CHIMP	--SLYREI K DAKEQVS-

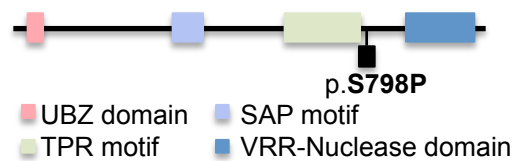
E BRCA2



p.D703Y

XENOPUS	-MLNS-----NVSNDM-
RAT	-PERSCEND Q KSPKVS-
MOUSE	-PERTCEND Q KSPKVS-
HORSE	-QEKHCED D SKGQRV-
COW	-QDKHCEDNPKSQGV-
CAT	-QEKHCED D TKSORVA-
DOG	-QEKHWED D AKKPRVS-
HUMAN	-QEGQCEN D PKSKKVS-
CHIMP	-QEGQCEN D PKSKKVS-

C FAN1



p.S798P

XENOPUS	-PQTGMGK S VFIMEEL-
RAT	-PQHGMGK S VFVIEG-
MOUSE	-PQHGMGK S VFVMES-
HORSE	-PQSGMGK S VFVMEA-
COW	-PQRGSGK S VFVLEA-
CAT	-PQRGMGK S VFVMEA-
DOG	-PQHGMGK S VFVMEA-
HUMAN	-PQRGMCK S VFVMEA-
CHIMP	-PQRGMCK S VFVMEA-

F SPIN1



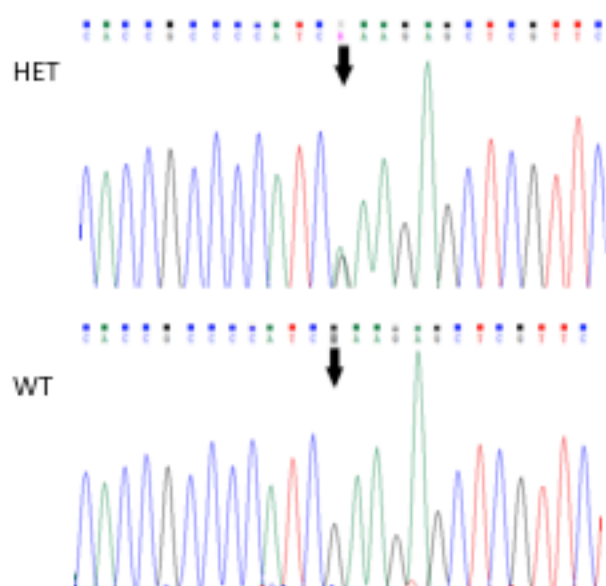
p.Y232X

XENOPUS	-FDDDFHI Y VYDLVKT-
RAT	-FDDDFHI Y VYDLVKT-
MOUSE	-FDDDFHI Y VYDLVKT-
HORSE	-FDDDFHI Y VYDLVKT-
COW	-FDDDFHI Y VYDLVKT-
CAT	-FDDDFHI Y VYDLVKT-
DOG	-FDDDFHI Y VYDLVKT-
HUMAN	-FDDDFHI Y VYDLVKT-
CHIMP	-FDDDFHI Y VYDLVKT-

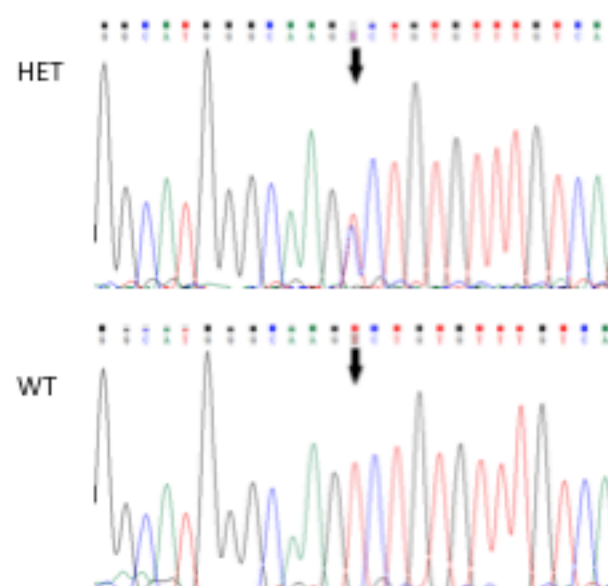
Supplemental Fig_S2 Legend: All amino acids as well as their surrounding sequence are highly conserved. Protein structures of (A) MRE11, (B) RAD50, (C) FAN1, (D) BRCA1, (E) BRCA2, and (F) SPIN1 with the location of the chemically-induced mutations are shown.

Supplemental_Fig_S3.pdf: Sanger sequence traces

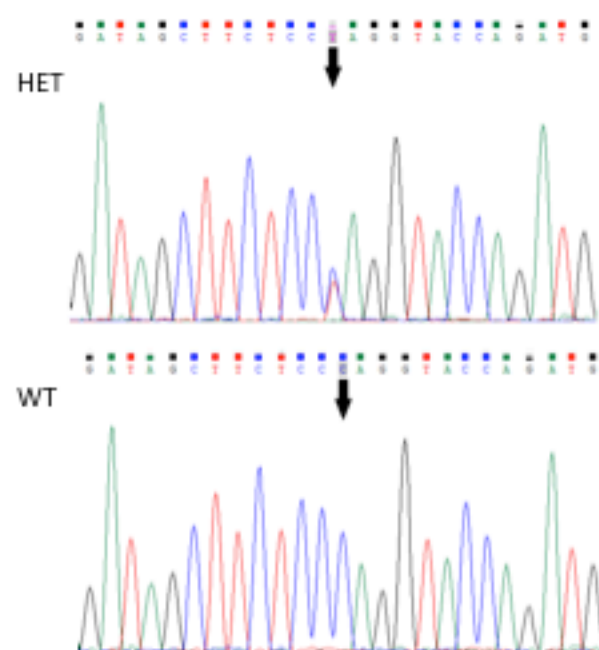
Birc6 exon49:c.9611G>A (M_199LL)



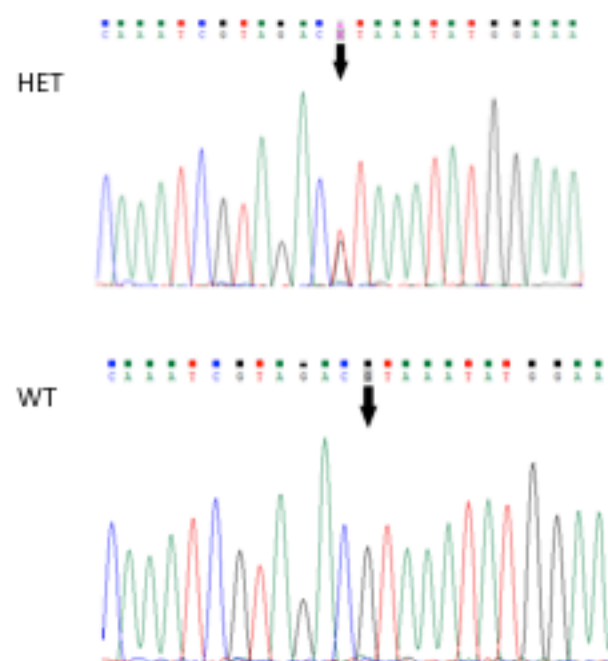
Fan1 exon10:c.2392T>C (A_333)



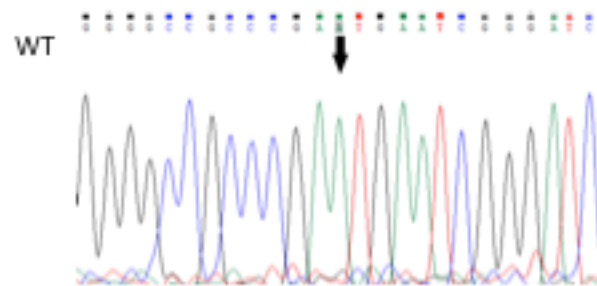
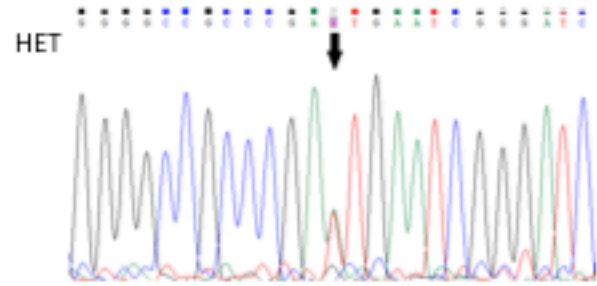
Rbbp8 exon15:c.2227C>T (895)



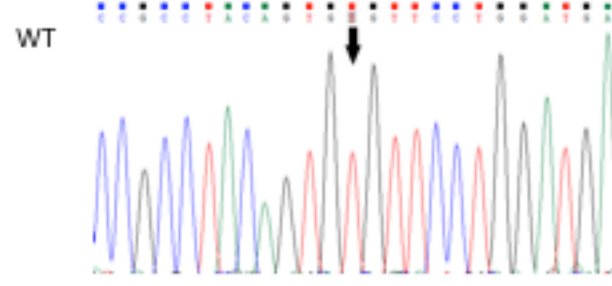
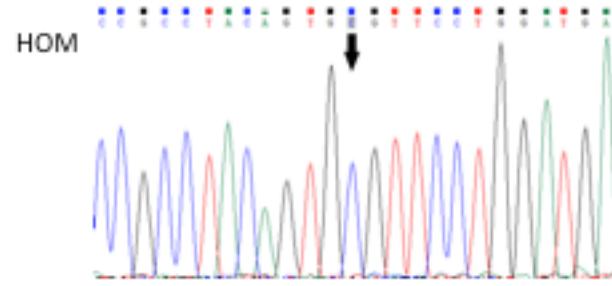
Spin1 exon4:c.696C>A:(4751)*



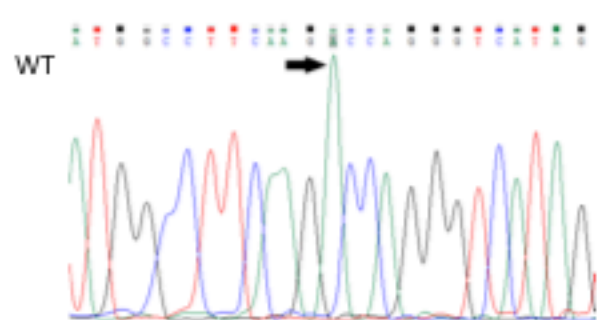
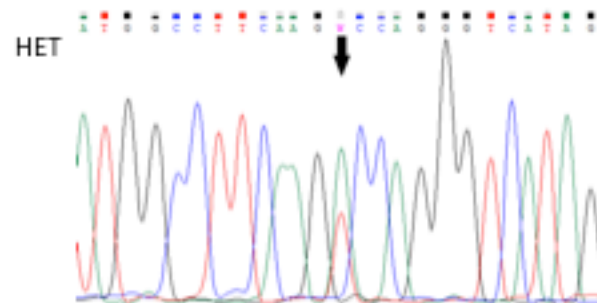
Tet1 exon1:c.241A>T (A_27R)



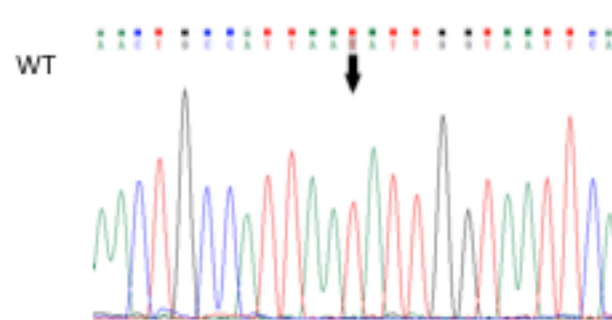
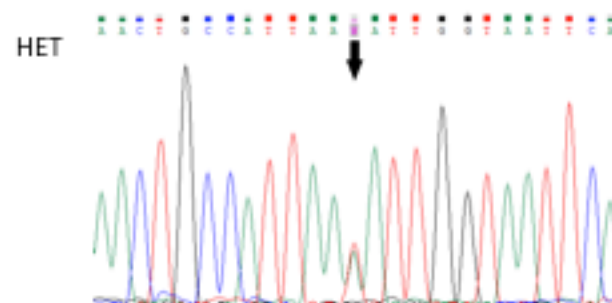
Aacs exon8:c.770T>C (4799)



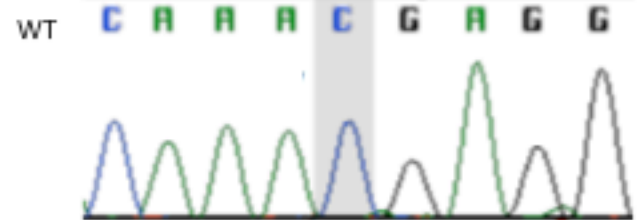
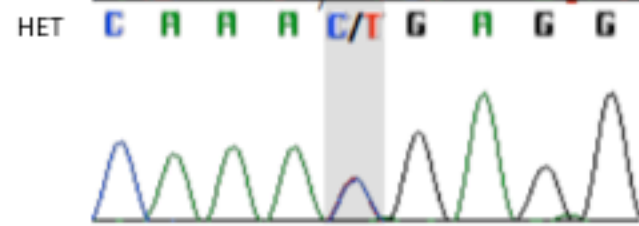
Apoa5 exon4:c.1061A>T (J_57L)



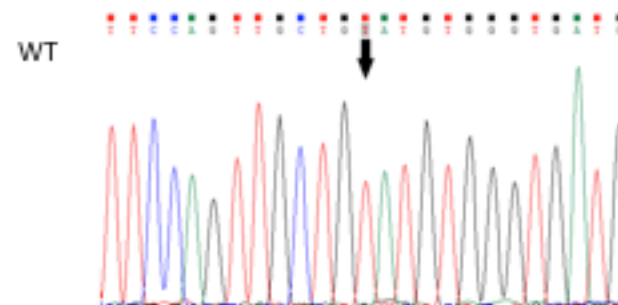
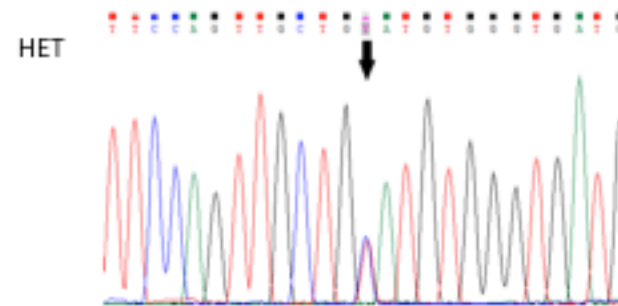
Atp8a1 exon22:c.2037T>A (4799)



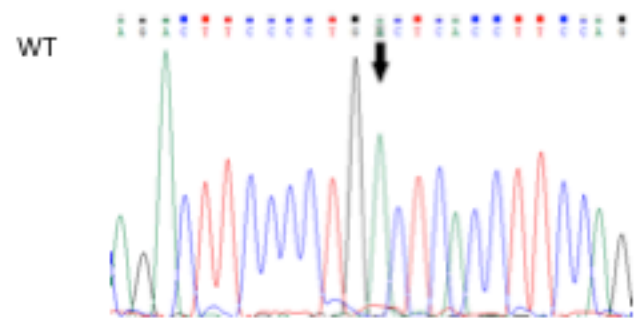
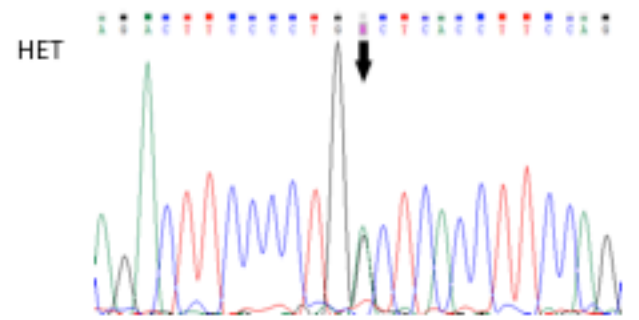
Sqle exon6:c.1195C>T (895)



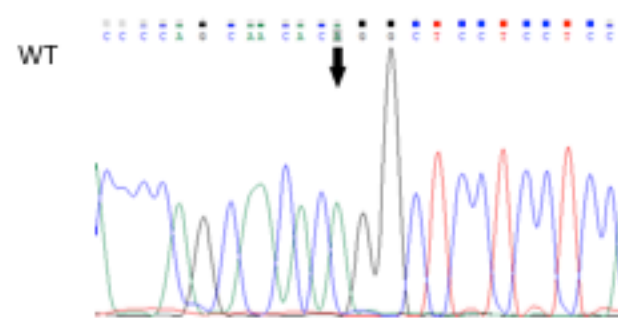
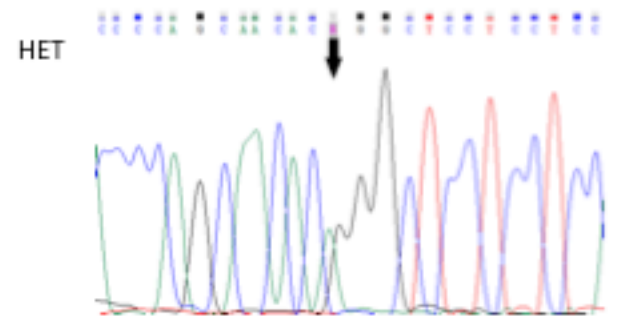
Tm7sf2 exon6:c.697T>C (4654)



Arhgef15 exon2:c.275A>G (4751)

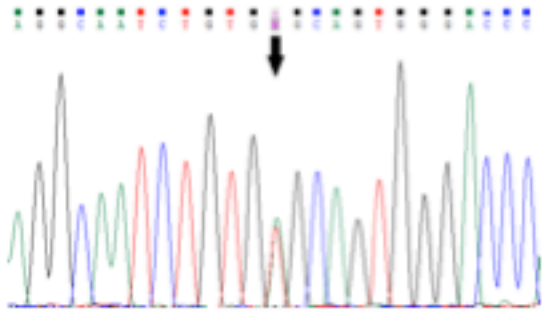


Cd22 exon13:c.2414A>G (520)

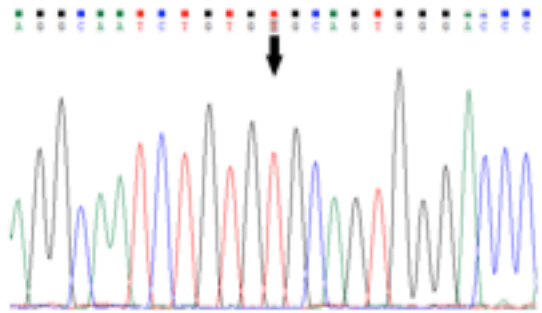


Celsr3 exon22:c.7415T>A (918-15L)

HET

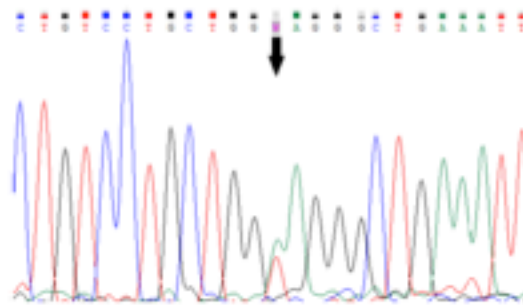


WT

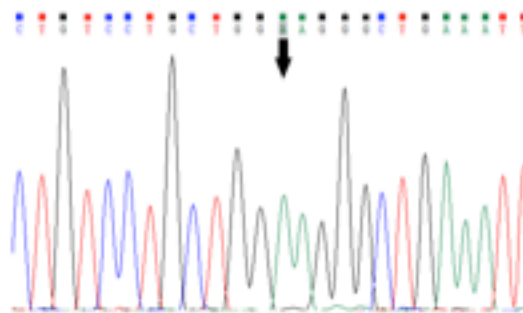


Dbn1 exon8:c.709T>A (M_199LL)*

HET

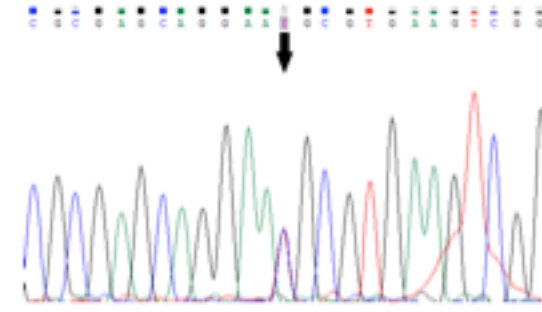


WT

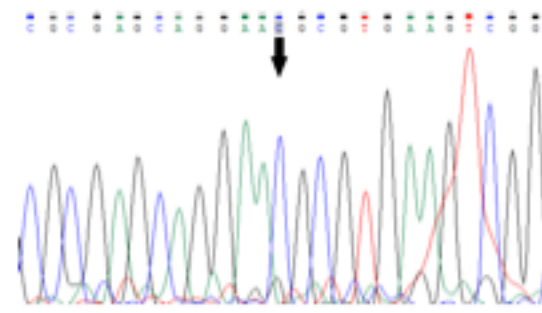


Hcn2 exon1:c.502C>T (J_57L)

HET

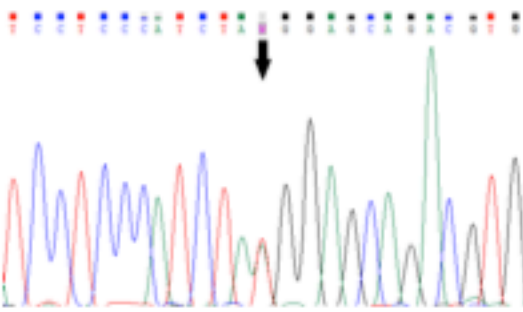


WT

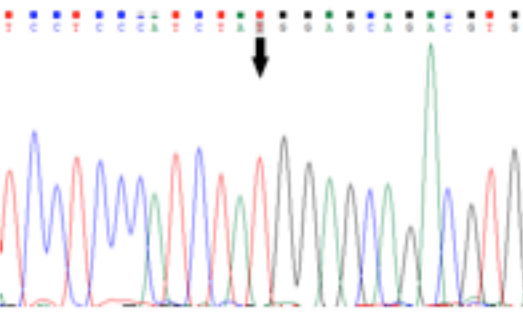


Kdm4a exon4:c.396T>A (520)

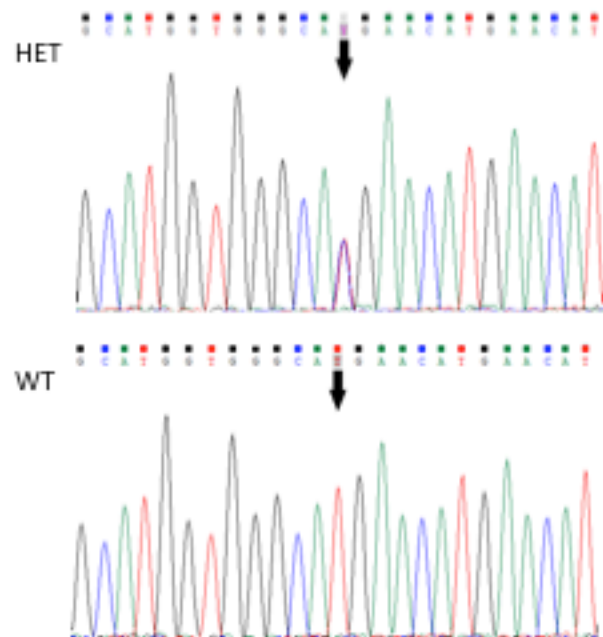
HET



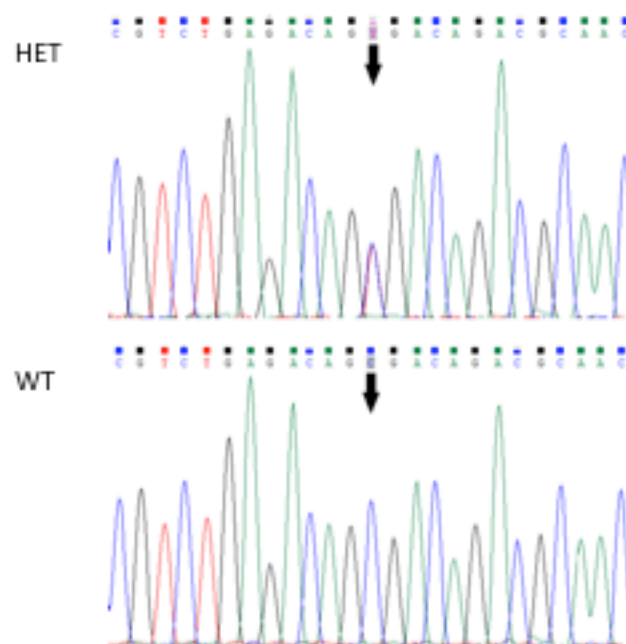
WT



Clint1 exon12:c.1730T>C (520)



Dennd4a exon15:c.1778C>T (4654)



*Reverse complement used for sequencing