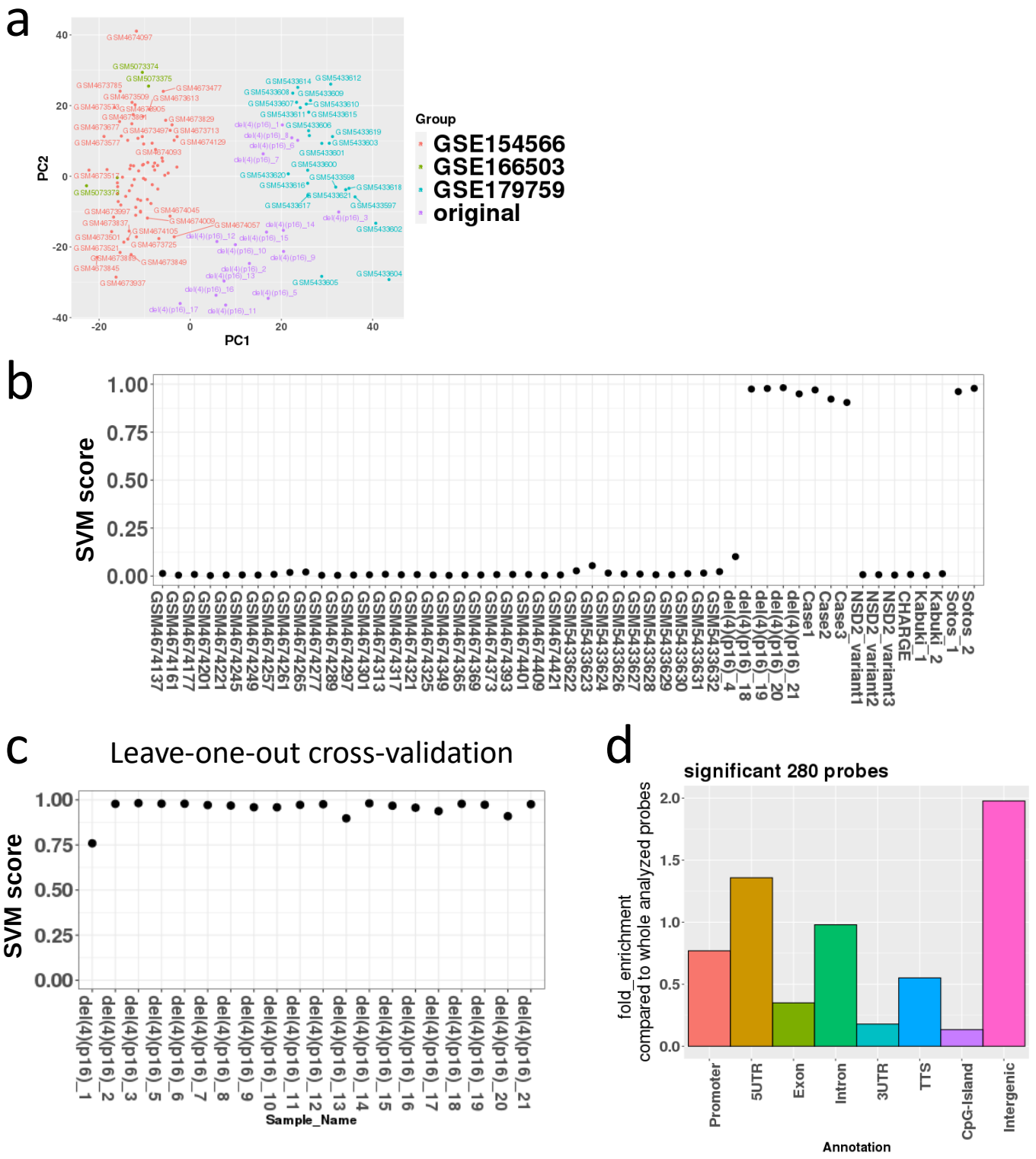
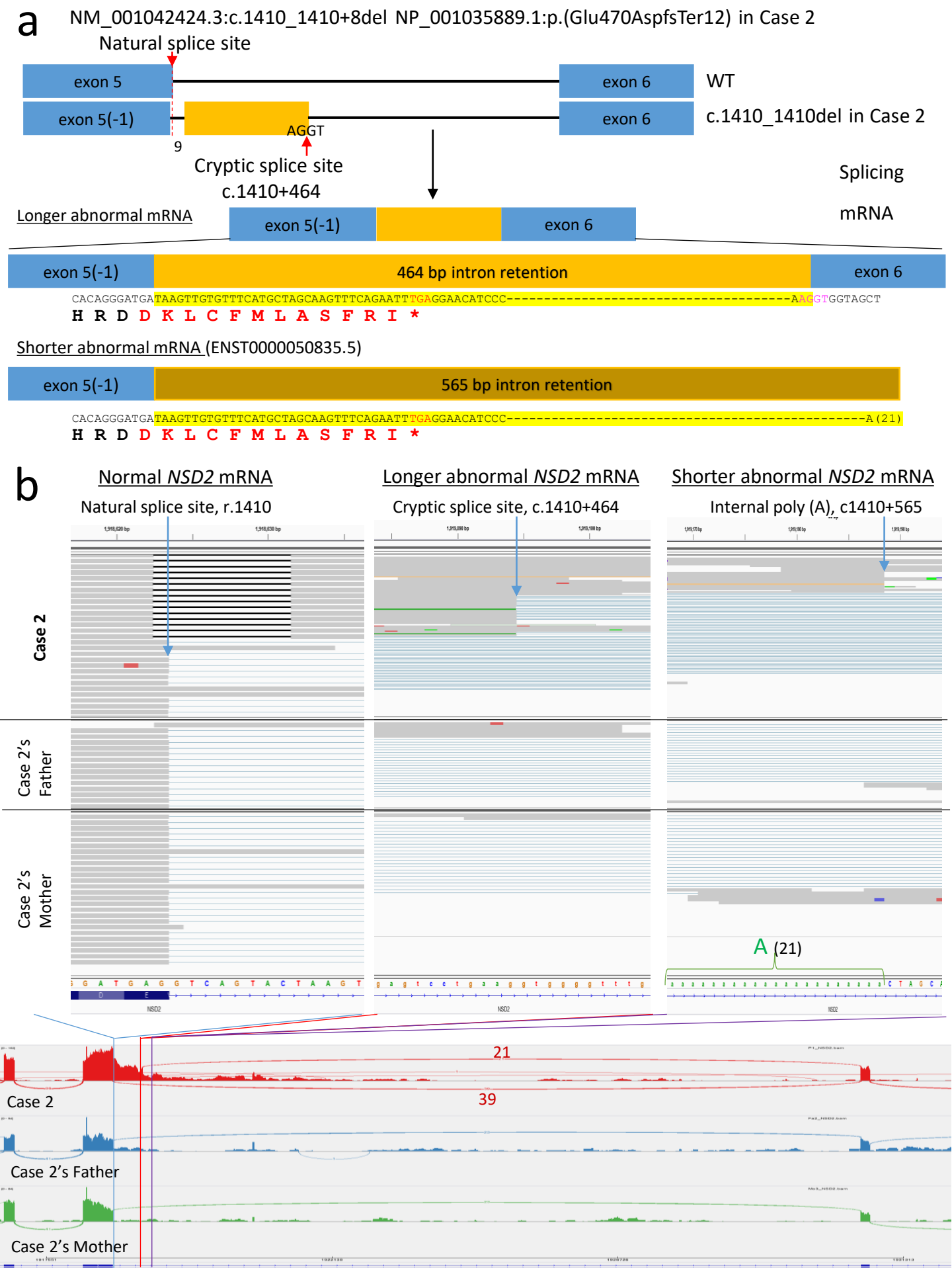
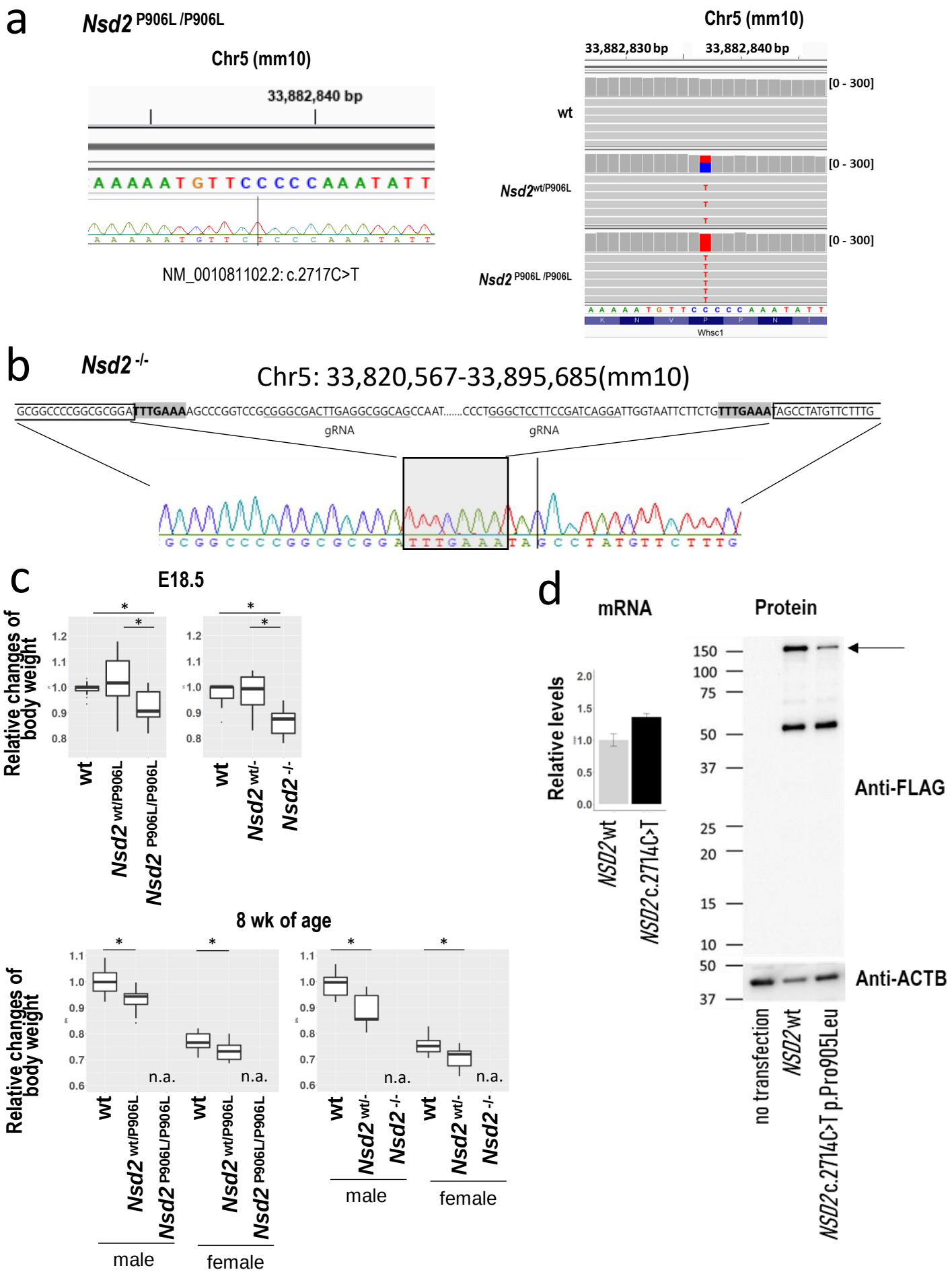


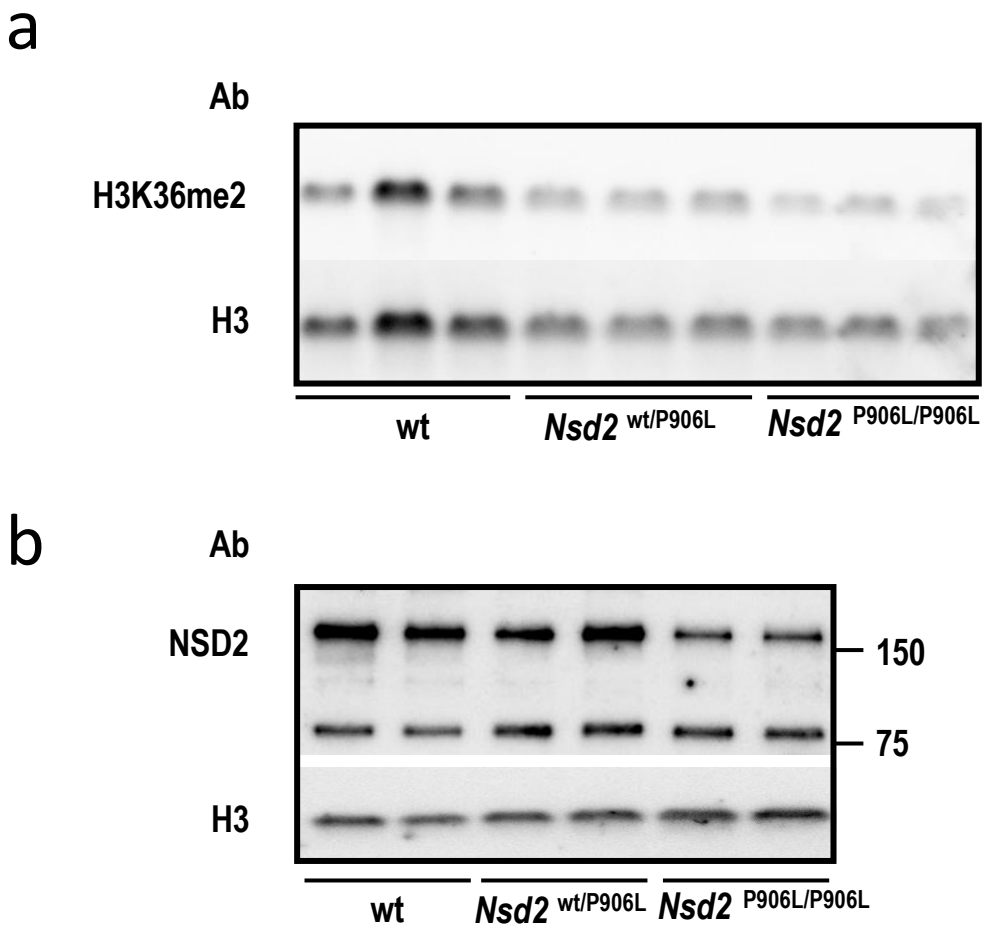
Supplementary Figure 1. The identification of WHS-related DNA methylation signatures in the 280 probes. (a) PCA analysis of individuals in the training set by the beta values of the total probes. (b) An SVM model using beta value of the 280 probes indicated that the 4 WHS cases (del(4)(p16)_18,19,20,21) in the testing set exhibited a DNA methylation pattern similar to the 16 WHS cases in the training set with high SVM scores. An individual with deleted 4p16.2p15.31 that did not include *NSD2* (del(4)(p16)_4) showed low SVM score. The 3 cases with *NSD2* variants (Case 1,2,3) showed high SVM score, while 3 cases with *NSD2* likely benign variants (NSD2_variant 1,2,3) showed low SVM score. (b) Leave-one-out cross-validation showed that the beta value of 280 probes correctly classified WHS from control by SVM. (c) Genomic distribution of the 280 probes indicated an enrichment in 'Intergenic'.


a
NM_001042424.3:c.1410_1410+8del NP_001035889.1:p.(Glu470AspfsTer12) in Case 2
Natural splice site
exon 5
exon 6
WT
exon 5(-1)
exon 6
c.1410_1410del in Case 2
AGGT
9
Cryptic splice site
c.1410+464
Splicing
mRNA
Longer abnormal mRNA
exon 5(-1)
exon 6
exon 5(-1)
464 bp intron retention
exon 6
CACAGGGATGATAAGTTGTGTTTCATGCTAGCAAGTTTCAGAATTTGAGGAACATCCC----------AAGGTGGTAGCT
H R D D K L C F M L A S F R I *
Shorter abnormal mRNA (ENST0000050835.5)
exon 5(-1)
565 bp intron retention
CACAGGGATGATAAGTTGTGTTTCATGCTAGCAAGTTTCAGAATTTGAGGAACATCCC----------A(21)
H R D D K L C F M L A S F R I *
b
Normal NSD2 mRNA
Natural splice site, r.1410
Longer abnormal NSD2 mRNA
Cryptic splice site, c.1410+464
Shorter abnormal NSD2 mRNA
Internal poly (A), c1410+565
Case 2
Case 2's Father
Case 2's Mother
A (21)
NSD2
21
39
Case 2
Case 2's Father
Case 2's Mother

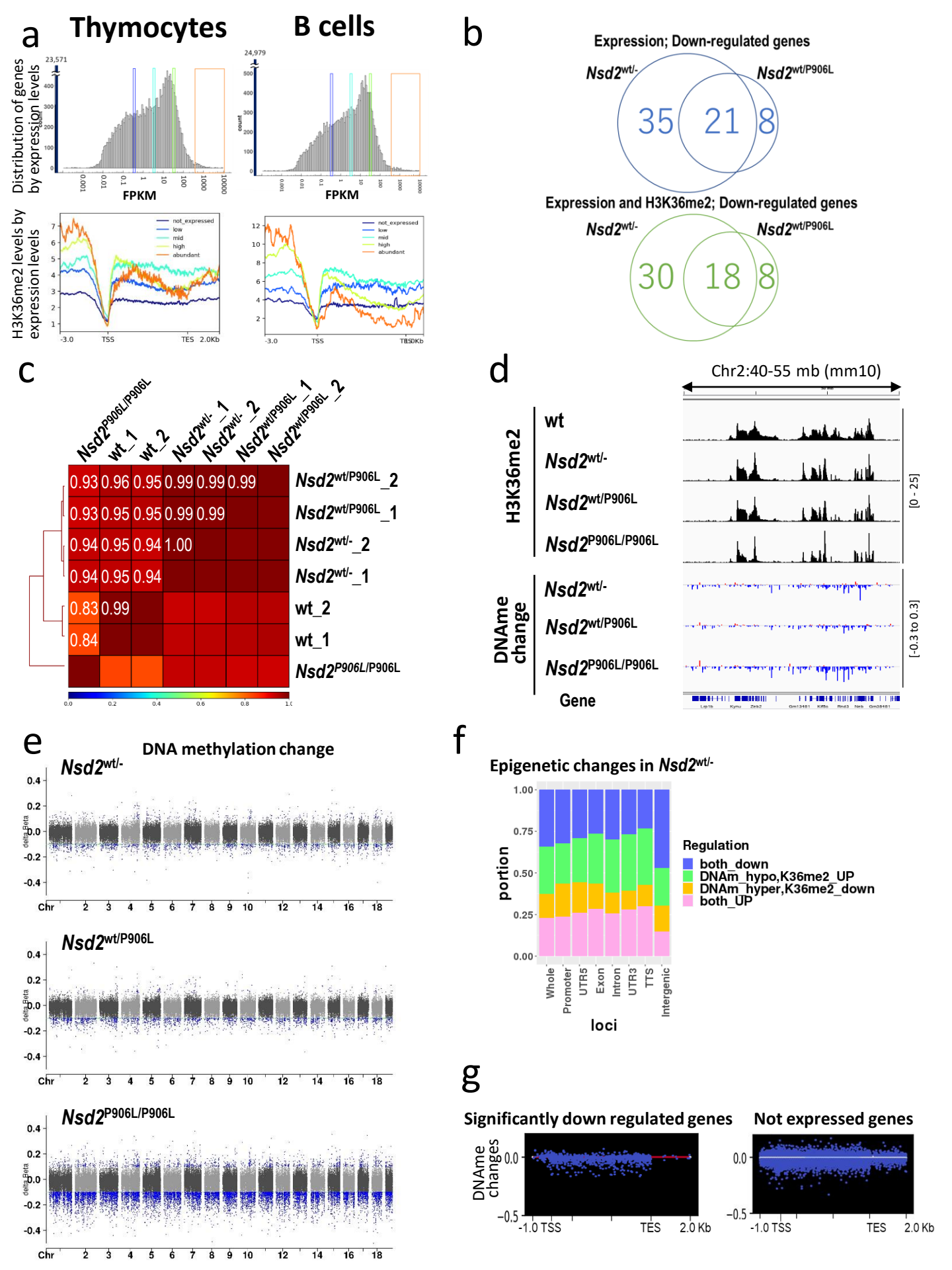
Supplementary Figure 2. Schematic description of *NSD2* transcript from *NSD2*_c.1410_1410+8del. (a) RNA-Seq of peripheral blood cells of Case 2 reveal a cryptic splice site in *NSD2* mRNA. Blue and yellow squares indicate exons and introns, respectively. (b) RNA-Seq of peripheral blood cells of Case 2 revealed existence of *NSD2* transcript with c.1410_1410+8del which indicated escape of nonsense mediated decay. Usage of the cryptic splice site was confirmed by some sequence reads. Case 2-specific transcript ends were confirmed at internal poly(A) site in intron 7 of *NSD2.*


a
Nsd2 P906L /P906L
Chr5 (mm10)
33,882,840 bp
NM_001081102.2: c.2717C>T
Chr5 (mm10)
33,882,830 bp
33,882,840 bp
wt
Nsd2wt/P906L
Nsd2 P906L /P906L
[0 - 300]
Whsc1
b
Nsd2 -/-
Chr5: 33,820,567-33,895,685(mm10)
gRNA
gRNA
c
E18.5
Relative changes of body weight
wt
Nsd2 wt/P906L
Nsd2 P906L/P906L
Nsd2 wt/-
Nsd2 -/-
8 wk of age
n.a.
male
female
d
mRNA
Protein
Relative levels
NSD2 wt
NSD2 c.2714C>T
Anti-FLAG
Anti-ACTB
no transfection
NSD2 wt
NSD2 c.2714C>T p.Pro905Leu

Supplementary Figure 3. Genotype and phenotype of genome-edited mice.
(a) Sequencing analyses for the genome-edited mice carrying c.2717C>T (P906L) substitution in the NM_001081102.2 *Nsd2* transcript. The result of Sanger sequencing of PCR amplicons from tail genomic DNA of *Nsd2*$^{P906L/P906L}$ is shown (left). RNA-sequencing results of *Nsd2*$^{wt/P906L}$ and *Nsd2*$^{P906L/P906L}$ thymocytes are shown (right). Both results confirmed the success of the base substitution. (b) Sequencing analysis for the genome-edited *Nsd2* knock-out mice (*Nsd2*$^{-/-}$). A DNA fragment containing deletion breakpoints was amplified by PCR from tail genomic DNA of *Nsd2*$^{-/-}$, and subjected to Sanger sequencing. The result demonstrated that a 75,080 bp region covering the entire coding region of *Nsd2* gene was deleted. The deletion breakpoints were presumed to be located within two TTTGAAA intervals, one from chr5:33,820,589 to chr5:33,820,593, and other from chr5:33,895,671 to chr5:33,895,675 (mm10). (c) Relative changes in bodyweight in Nsd2-variant mice, compared with the wild-type mice at embryonic day 18.5 and 8 weeks of age. We showed relative changes since the number of littermates differed in mice by the embryonic lethality of the mutant mice. Statistical analysis was performed by ANOVA, with post hoc comparisons by t.test, or t.test. A significant decrease in intra-uterine growth was confirmed in homozygous mice of both variants. Failure to thrive was confirmed in heterozygous mice of both variants at 8 weeks of age. * $p < 0.05$. (d) Overexpression of N-terminal FLAG-tagged NSD2: c.2714C>T in NSD2 (NM_001042424.3) coding p.Pro905Leu revealed a slight increase in mRNA levels but decrease in protein levels compared with the overexpression of N-terminal FLAG-tagged NSD2 wild-type. The arrow indicates approx.150 kDa of O96028-1.

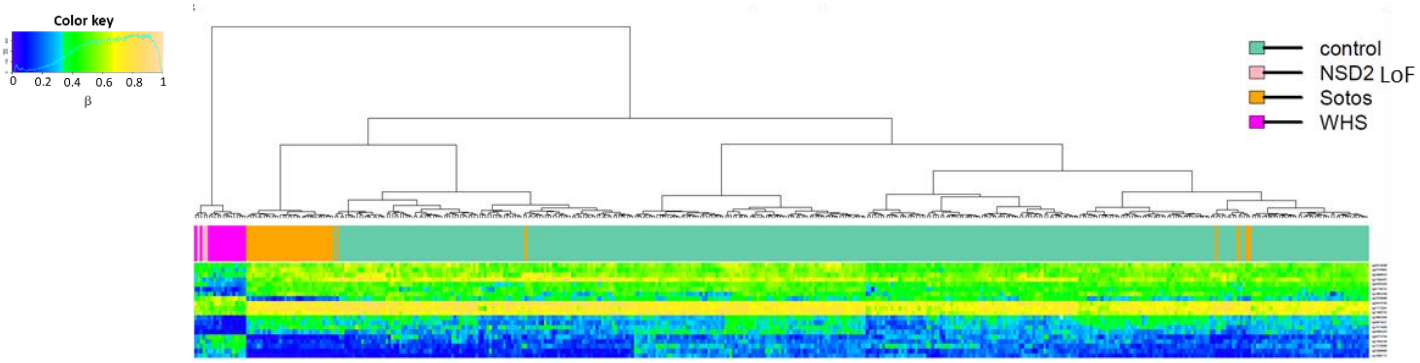


Supplementary Figure 4. Western blotting of biological replicates.
(a) H3K36me2 levels of knock-in mice of Nsd2-p.Pro906Leu. Biological replicates of fetal thymocytes were collected from each genotype. (b) Nsd2 levels of knock-in mice of Nsd2-p.Pro906Leu. Biological replicates of fetal brains were collected from each genotype.


a
Thymocytes
B cells
Distribution of genes by expression levels
23,571
24,979
FPKM
H3K36me2 levels by expression levels
not_expressed
low
mid
high
abundant
TSS
TES
2.0Kb
b
Expression; Down-regulated genes
Nsd2wt/-
Nsd2wt/P906L
35
21
8
Expression and H3K36me2; Down-regulated genes
30
18
8
c
Nsd2P906L/P906L
wt_1
wt_2
Nsd2wt/-_1
Nsd2wt/-_2
Nsd2wt/P906L_1
Nsd2wt/P906L_2
0.93 0.96 0.95 0.99 0.99 0.99
0.93 0.95 0.95 0.99 0.99
0.94 0.95 0.94 1.00
0.94 0.95 0.94
0.83 0.99
0.84
d
Chr2:40-55 mb (mm10)
H3K36me2
wt
Nsd2wt/-
Nsd2wt/P906L
Nsd2P906L/P906L
[0 - 25]
DNAme change
[-0.3 to 0.3]
Gene
e
DNA methylation change
delta_Beta
Chr
f
Epigenetic changes in Nsd2wt/-
portion
loci
Whole
Promoter
UTR5
Exon
Intron
UTR3
TTS
Intergenic
Regulation
both_down
DNAm_hypo,K36me2_UP
DNAm_hyper,K36me2_down
both_UP
g
Significantly down regulated genes
Not expressed genes
DNAme changes
-1.0 TSS
TES
2.0 Kb

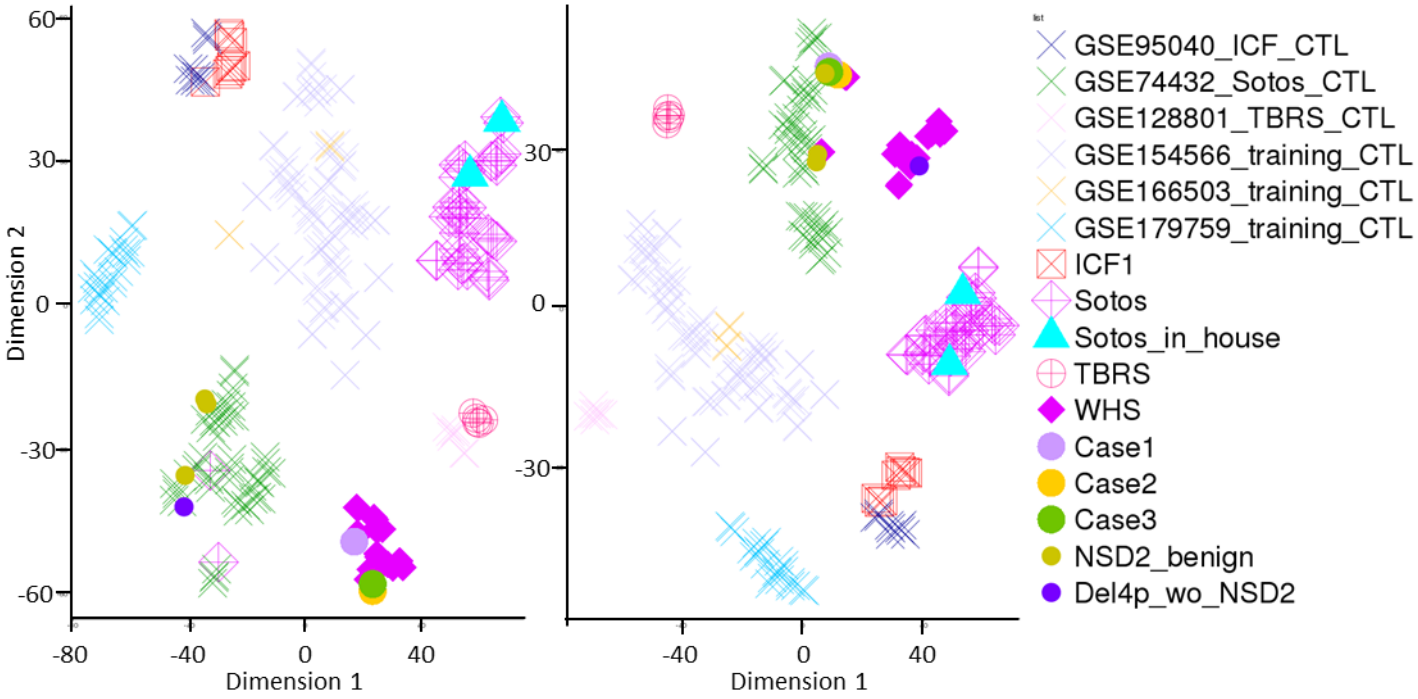
Supplementary Figure 5. Functional assay of *Nsd2* mutant mice. (a) Relevance between H3K36me2 and absolute gene expression levels. (b) Gene expression changes in *Nsd2*-edited mice. Significant expression changes in *Nsd2*$^{wt/-}$ and *Nsd2*$^{wt/P906L}$ thymocytes compared to wild type were indicated by arrows based on FPKM. Blue arrows indicate the genes in which H3K36me2 at promoter altered in the same directions as expression changes. Red arrows indicate the genes that did not accompany H3K36me changes at promoter. (c) H3K36me2 correlation between genotypes. The numbers in heatmap indicate correlation coefficients by pearson correlation test. (d) H3K36me2 decrease and DNA hypo-methylation were more obvious in *Nsd2*$^{P906L/P906L}$ than *Nsd2*$^{wt/-}$ and *Nsd2*$^{wt/P906L}$. (e) DNA hypo-methylation were more than hyper-methylation in *Nsd2* mutant thymocytes. (g) Proportion of combination of DNA methylation changes and H3K36me2 changes in *Nsd2*$^{wt/-}$ thymocytes by genetic regions. Concomitant decrease of DNA methylation and H3K36me2 at intergenic region was most major changes. (g) DNA methylation changes in gene body in *Nsd2*$^{wt/-}$ thymocytes; significantly down-regulated genes in *Nsd2*$^{wt/-}$ (left) and not expressed genes in thymocytes (right).

a **20 probes distinguishing *NSD2* defects, *NSD1* defects, and control**

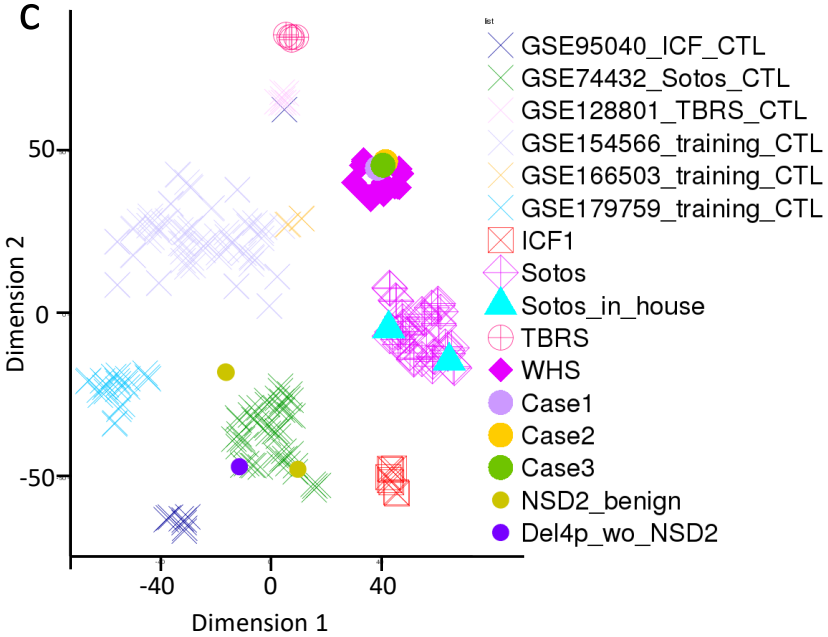


b **Clustering by tSNE with Top 1000 differentially methylated probes**

in *NSD2* defects (WHS) **in *NSD1* defects (Sotos)**



c



Supplementary Figure 6. Distinction between *NSD2* loss of function and other syndromes by DNA methylation signatures. (a) Hierarchical clustering of individuals with *NSD2* loss of function (NSD2 LoF or WHS), *NSD1* loss of function (Sotos syndrome 1), and controls by methylation levels of the 20 probes. Heatmap indicates methylation levels. (b) tSNE clustering of individuals with *NSD2* defects (WHS and Cases), *NSD1* defects (Sotos and Sotos inHouse), *DNMT3A* defects (TBRS), *DNMT3B* defects (ICF1), and controls of each study by methylation levels of the top 1,000 p-values DMP probes. (c) tSNE clustering of individuals with either 4 syndromes; *NSD2* defects (WHS and Cases), *NSD1* defects (Sotos and Sotos inHouse), *DNMT3A* defects (TBRS), *DNMT3B* defects (ICF1) by mixed top 100 p-values DMP probes.