

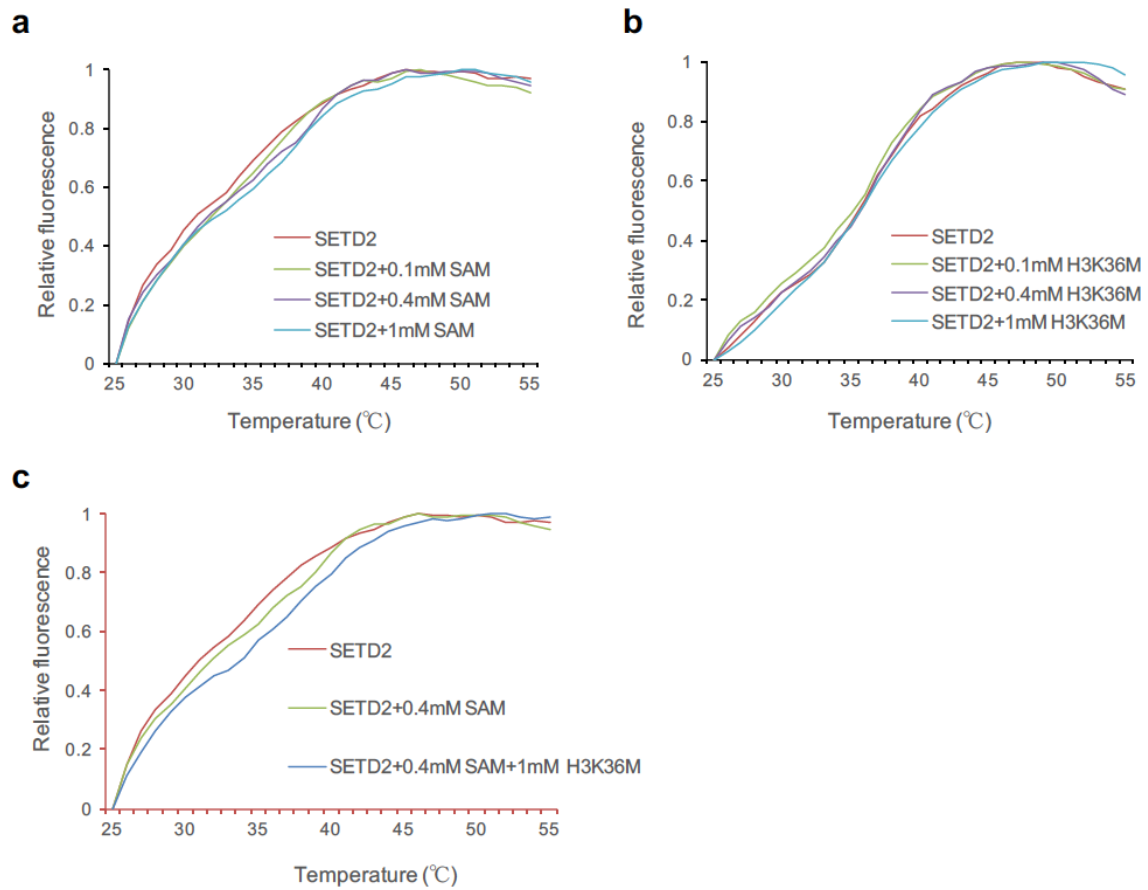
Molecular basis for the role of oncogenic histone mutations in modulating H3K36 methylation

Yinglu Zhang^{1,2}, Chunmin Shan^{1,2}, Jiyong Wang^{1,2}, Kehan Bao¹, Liang Tong^{1,3}, and Songtao Jia^{1,3}

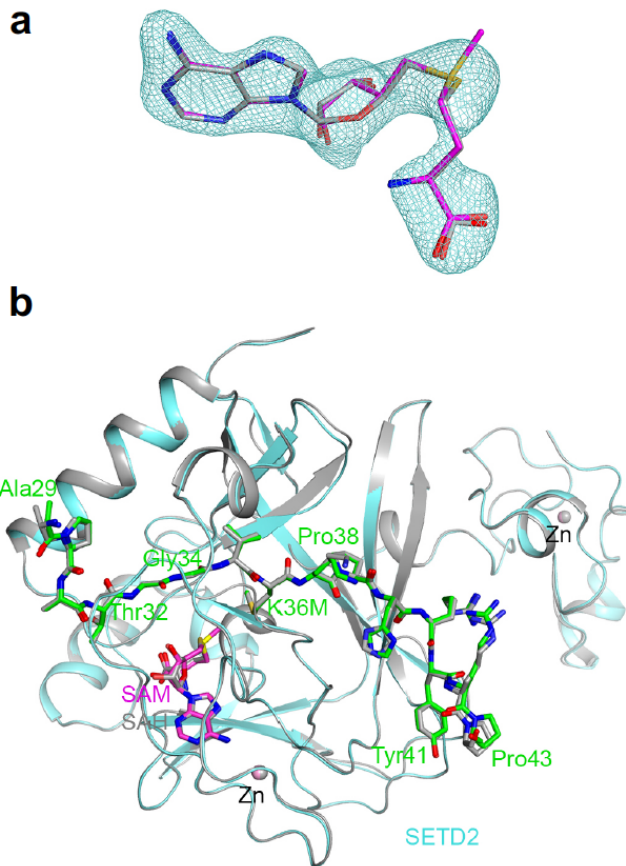
¹Department of Biological Sciences, Columbia University, New York, NY 10027

²These authors contributed equally to this work

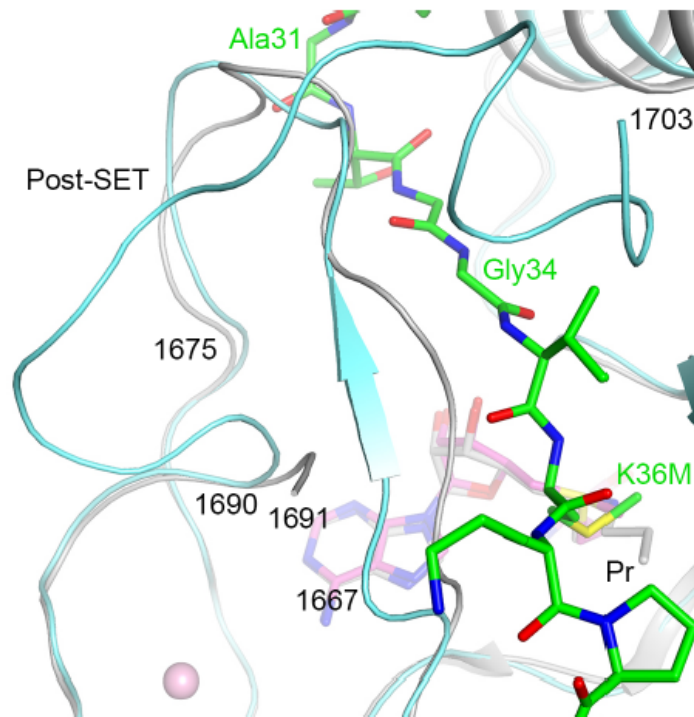
³Corresponding authors: ltong@columbia.edu, jia@biology.columbia.edu



Extended data Figure 1. Thermal shift assay to measure the interaction between the SET domain of SETD2, SAM and an H3K36M peptide.



Extended data Figure 2. The structure of the SET domain of SETD2 in complex with an H3K36M peptide and SAH. (a) Omit $F_o - F_c$ electron density at 2.4 Å resolution for SAH (in gray for carbon atoms), contoured at 3σ . The bound position of SAM is also shown (magenta) and is nearly identical. (b) The overall structure of the SET domain in complex with the H3K36M peptide and SAH (in gray) is essentially the same as that of the SAM complex (SET domain in cyan, H3K36M peptide in green stick models and SAM in magenta).

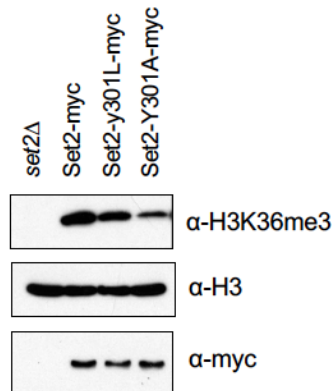


Extended data Figure 3. Overlay of the structure of SETD2 SET domain (cyan) in complex with H3K36M peptide (green) and SAM (magenta) with that of the SET domain in complex with the inhibitor N-propyl sinefungin (Pr-SNF) (gray)²⁵. There are still large conformational differences, although the steric clash between the H3K36M peptide and the post-SET motif is not as severe. The propyl group of Pr-SNF is labeled Pr.

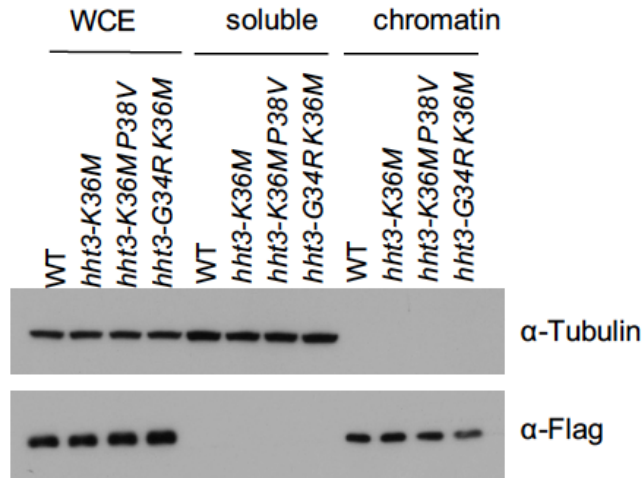
s.p. SET2 191 **G**F**G**L**R**A**D**A**N**L**P****K**D**T**F**V****Y****E****Y****I****G****E****V****I****P****E****Q****K****F****R****K****R****M****R****O****Y****D****S****E****G****I****K****H****F****Y****F****M****M****L****O****K****G****E****Y****I****D****A****T****K****R**
 h.s. SETD2 1561 **G****W****G****L****R****A****A****K****D****L****P****S****N****T****F****V****L****E****Y****C****G****E****V****L****D****H****K****E****F****K****A****R****V****K****E****Y****A****R****N****K****N****I****H****Y****F****M****A****L****K****N****D****E****I****I****D****A****T****O****K**
 h.s. NSD1 1953 **G****W****G****L****R****T****K****T****D****I****K****K****G****E****F****V****N****E****Y****V****G****E****L****I****D****E****E****C****R****A****R****I****R****Y****A****Q****E****H****D****I****T****N****F****Y****M****L****T****L****D****K****D****R****I****I****D****A****G****P****K**
 h.s. NSD2 1074 **G****W****G****L****V****A****K****R****D****I****R****K****G****E****F****V****N****E****Y****V****G****E****L****I****D****E****E****C****M****A****R****I****K****H****A****H****E****N****D****I****T****H****F****Y****M****L****T****I****D****K****D****R****I****I****D****A****G****P****K**
 h.s. NSD3 1156 **G****W****G****L****R****T****K****R****S****I****K****K****G****E****F****V****N****E****Y****V****G****E****L****I****D****E****E****C****R****L****R****I****K****R****A****H****E****N****S****V****T****N****F****Y****M****L****T****V****T****K****D****R****I****I****D****A****G****P****K**

s.p. SET2 251 **G****S****L****A****R****F****C****N****H****S****C****R****P****N****C****Y****V****D****K****W****M****V****G****D****K****L****R****M****G****I****F****C****K****R****D****I****R****G****E****E****L****T****F****D****Y****N****V****D****R****Y****G****A**
 h.s. SETD2 1621 **G****N****C****S****R****F****M****N****H****S****C****E****P****N****C****E****T****Q****K****W****T****V****N****G****L****R****V****G****F****F****T****T****K****L****V****P****S****G****S****E****L****T****F****D****Y****Q****F****O****R****Y****G****K**
 h.s. NSD1 2013 **G****N****Y****A****R****F****M****N****H****C****C****O****P****N****C****E****T****Q****K****W****S****V****N****G****D****T****R****V****G****L****F****A****L****S****D****I****K****A****G****T****E****L****T****F****N****Y****N****L****E****C****L****G****N**
 h.s. NSD2 1134 **G****N****Y****S****R****F****M****N****H****S****C****O****P****N****C****E****T****L****K****W****T****V****N****G****D****T****R****V****G****L****F****A****V****C****D****I****P****A****G****T****E****L****T****F****N****Y****N****L****D****C****L****G****N**
 h.s. NSD3 1216 **G****N****Y****S****R****F****M****N****H****S****C****N****P****N****C****E****T****Q****K****W****T****V****N****G****D****V****R****V****G****L****F****A****L****C****D****I****P****A****G****M****E****L****T****F****N****Y****N****L****D****C****L****G****N**

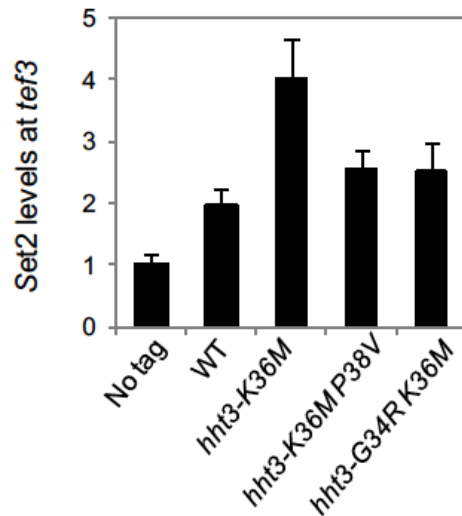
Extended data Figure 4. Sequence alignment of fission yeast (s.p.) and human (h.s.) H3K36 methyltransferases. Amino acids that mediate interaction with H3K36M are indicated by red asterisks, and amino acids that restrict the H3G34 access channel are indicated by blue arrows.



Extended data Figure 5. Western blot analyses of H3K36me3, H3, and Set2-myc levels.



Extended data Figure 6. Chromatin fraction analyses to measure the incorporation of mutant histones into chromatin. Cell lysates (whole cell extract, WCE) were separated into soluble and chromatin-bound fractions. Western blot analyses were performed to measure Tubulin (cytoplasmic marker) or mutant histone levels in different fractions.



Extended data Figure 7. ChIP analyses of Set2-Flag levels at a target gene *tef3*, normalized to the silent mating-type locus. The mutant histone is GFP-tagged. Error bars represent the standard deviation of three replicates.

Extended data Table 1. Yeast strains used in this study

Name	<i>Mat</i>	<i>leu1</i>	<i>his2</i>	<i>ura4</i>	<i>ade6</i>	additional genotype
BR172	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>otr::ura4⁺</i>
SPJ4573	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-Flag::hphMX6</i>
SPJ4729	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-Flag::hphMX6 set2Δ::natMX6</i>
SPJ4647	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-K36M-Flag::hphMX6</i>
SPJ4767	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-K36R-Flag::hphMX6</i>
SPJ4768	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34R-Flag::hphMX6</i>
SPJ4739	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34V-Flag::hphMX6</i>
SPJ4813	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34W-Flag::hphMX6</i>
SPJ4820	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34L-Flag::hphMX6</i>
SPJ4771	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34A-Flag::hphMX6</i>
SPJ4787	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34C-Flag::hphMX6</i>
SPJ4783	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34S-Flag::hphMX6</i>
SPJ4779	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34D-Flag::hphMX6</i>
SPJ4775	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34N-Flag::hphMX6</i>
SPJ4799	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34T-Flag::hphMX6</i>
SPJ4735	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-K36MP38V-Flag::hphMX6</i>
SPJ4751	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34RK36M-Flag::hphMX6</i>
CS567	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>set2-myc::kanMX6</i>
CS568	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-Flag::hphMX6 set2-myc::kanMX6 otr::ura4⁺</i>
CS556	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-K36M-Flag::hphMX6 set2-myc::kanMX6 otr::ura4⁺</i>
CS649	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-K36MP38V-Flag::hphMX6 set2-myc::kanMX6 otr::ura4⁺</i>
CS647	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34RK36M-Flag::hphMX6 set2-myc::kanMX6 otr::ura4⁺</i>
CS112	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-GFP::kanMX6</i>
JW3464	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-GFP::kanMX6 set2-Flag::kanMX6</i>
JW3411	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-K36M-GFP::natMX6 set2-Flag::kanMX6 otr::ura4⁺</i>
JW3415	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-K36MP38V-GFP::natMX6 set2-Flag::kanMX6 otr::ura4⁺</i>
JW3428	<i>mat1Msmt0</i>	<i>leu1</i>	<i>his2</i>	DS/E	M210	<i>hht3-G34RK36M-GFP::natMX6 set2-Flag::kanMX6 otr::ura4⁺</i>

Extended data Table 2. DNA oligos used in this study.

primer name	sequence	annotation
cenH-RT-D	TTCAAGTCTTCTTTATACGTTTGC	for ChIP-qPCR of primer set at silent mating-type region
cenH-RT-C	CCATAGTAGTATGGCTATGAATG	
tef3-RT-R	TGCGACAGATTGCTTAACTG	for ChIP-qPCR of primer set at the <i>tef3</i> locus
tef3-RT-S	GTAATACTAGTTGGTAAAAGGG	