nature chemical biology

ARTICLES

https://doi.org/10.1038/s41589-021-00776-9

Check for updates

Genome-wide screens uncover KDM2B as a modifier of protein binding to heparan sulfate

Ryan J. Weiss[1,8], Philipp N. Spahn[2,8], Austin W. T. Chiang[2], Qing Liu[1], Jing Li[1], Kristina M. Hamill[3], Sandra Rother[1], Thomas M. Clausen[1,4], Marten A. Hoeksema[1], Bryce M. Timm[3], Kamil Godula[3,5], Christopher K. Glass[1,6], Yitzhak Tor[3,5], Philip L. S. M. Gordts[5,6], Nathan E. Lewis[5,7] ✉ and Jeffrey D. Esko[1,5] ✉

Heparan sulfate (HS) proteoglycans bind extracellular proteins that participate in cell signaling, attachment and endocytosis. These interactions depend on the arrangement of sulfated sugars in the HS chains generated by well-characterized biosynthetic enzymes; however, the regulation of these enzymes is largely unknown. We conducted genome-wide CRISPR–Cas9 screens with a small-molecule ligand that binds to HS. Screening of A375 melanoma cells uncovered additional genes and pathways impacting HS formation. The top hit was the epigenetic factor *KDM2B*, a histone demethylase. *KDM2B* inactivation suppressed multiple HS sulfotransferases and upregulated the sulfatase SULF1. These changes differentially affected the interaction of HS-binding proteins. *KDM2B*-deficient cells displayed decreased growth rates, which was rescued by *SULF1* inactivation. In addition, *KDM2B* deficiency altered the expression of many extracellular matrix genes. Thus, *KDM2B* controls proliferation of A375 cells through the regulation of HS structure and serves as a master regulator of the extracellular matrix.

All animal cells express one or more HS proteoglycans (HSPGs), which are comprised of a small set of proteins with the common feature of containing one or more covalently linked HS chains. HS is a linear polysaccharide assembled in the Golgi by the copolymerization of alternating residues of *N*-acetyl-D-glucosamine (GlcNAc) and D-glucuronic acid (GlcA) on a tetrasaccharide primer linked to the proteoglycan core protein. As the chains assemble, they undergo a series of modification reactions that include *N*-deacetylation and *N*-sulfation of a subset of GlcNAc residues, epimerization of adjacent GlcA residues to L-iduronic acid (IdoA), 2-*O*-sulfation of IdoA and GlcA residues and 6-*O*-sulfation and occasional 3-*O*-sulfation of glucosamine units (Fig. 1a). There is great structural heterogeneity in chain length, size and spacing of the modified tracts, and in the extent of sulfation and epimerization within the modified segments. The arrangement of sulfate groups and uronic acid epimers varies temporally and spatially during development and in different tissues in the adult. The modified regions of the chains make up binding sites for various protein ligands, which are known collectively as the HS interactome[1]. Binding can result in tethering and presentation of protein ligands at the cell surface, protein oligomerization and allosteric regulation. In this way, HSPGs regulate important biological processes including cell proliferation, motility, development and various physiological phenomena[2].

The enzymes involved in HS biosynthesis have been identified, cloned, biochemically characterized and mutated in both cells and model organisms[3]. Although much is known about the biosynthetic enzymes and HS composition, little is known about the mechanisms that dictate tissue-specific patterning of HS chains. The organization of sulfated residues along the chains, in part, reflects the substrate specificity of the enzymes, their sequential mode of action and possibly their physical association into complexes[3,4]. However, other nonenzymatic factors can regulate both patterning of the chains and expression of proteoglycan core proteins. Such factors include transcriptional regulatory elements, 5′UTR and 3′UTR sequences present in messenger RNAs, microRNAs and molecular chaperones (refs. [5,6] and references therein). Thus, the composition of the chains, their binding properties and ultimately their biological activity depends on diverse factors in addition to the catalytic properties of the biosynthetic enzymes.

The development of CRISPR–Cas9 as a genetic tool has enabled genome-wide loss-of-function screens. Here, a genome-wide CRISPR–Cas9-mediated screen was developed to uncover and characterize novel regulators of HS assembly. High-throughput screening assays were adapted using a low-molecular-weight, HS-dependent ligand, guanidinylated neomycin (GNeo), to identify lentiviral-encoded single-guide RNAs that induce resistance to a GNeo–toxin conjugate or reduce binding of a fluorescent GNeo derivative to HS on the cell surface. Enrichment analysis of the sequencing data recovered known HS biosynthetic enzymes, as well as many genes with previously unrecognized links to HS assembly. One of the top-scoring genes, *KDM2B*, is a histone lysine demethylase. Subsequent gene inactivation, RNA sequencing, chromatin immunoprecipitation sequencing (ChIP–seq) and biochemical analysis revealed *KDM2B* to be a regulator of HS structure, HS binding to protein ligands and growth of cells in monolayer culture

[1]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA, USA. [2]Department of Pediatrics, University of California, San Diego, La Jolla, CA, USA. [3]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, USA. [4]Department for Immunology and Microbiology, Faculty of Health and Medical Sciences, University of Copenhagen and Department of Infectious Disease, Copenhagen University Hospital, Copenhagen, Denmark. [5]Glycobiology Research and Training Center, University of California, San Diego, La Jolla, CA, USA. [6]Department of Medicine, University of California, San Diego, La Jolla, CA, USA. [7]Department of Pediatrics, The Novo Nordisk Foundation Center for Biosustainability and Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA. [8]These authors contributed equally: Ryan J. Weiss, Philipp N. Spahn ✉e-mail: n4lewis@eng.ucsd.edu; jesko@health.ucsd.edu

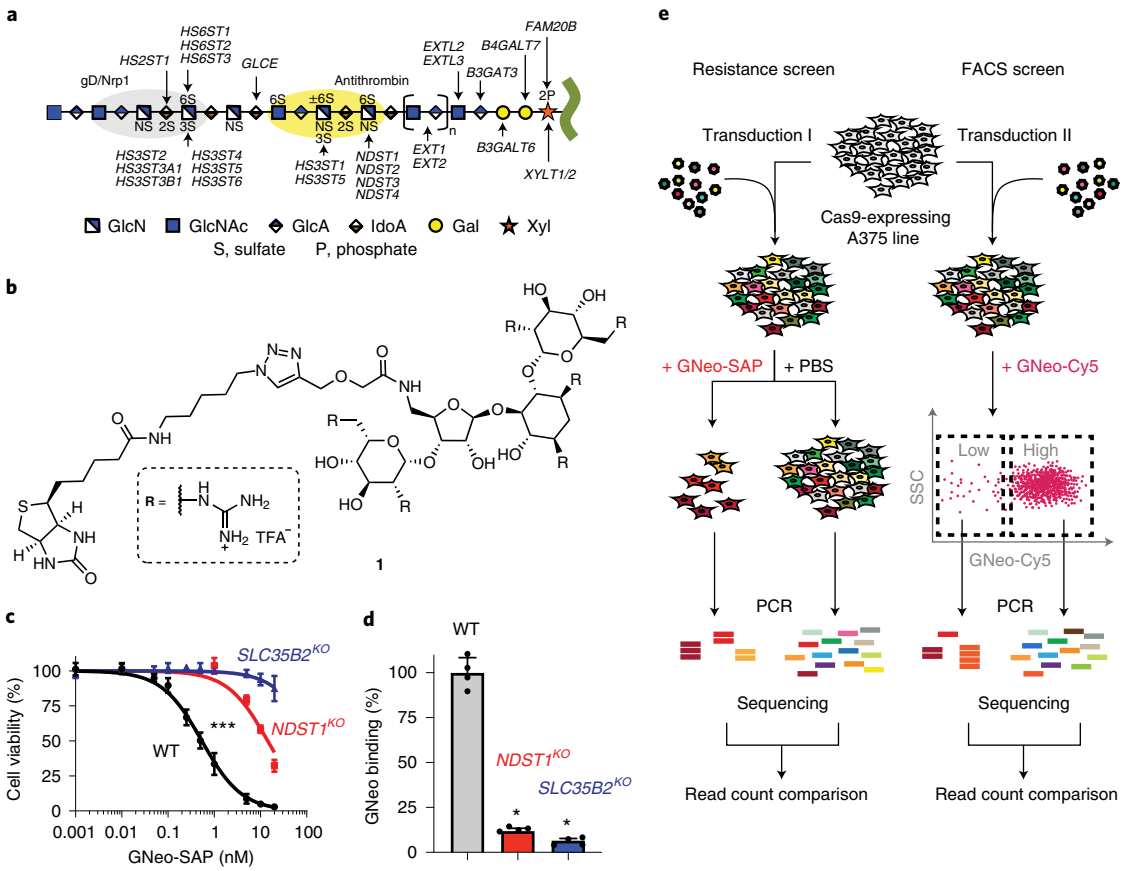


Fig. 1 | Finding novel regulators of HS biosynthesis through genome-wide screens. a, HS assembles while attached via a tetrasaccharide to a proteoglycan core protein. NDSTs, HS2ST, HS3STs and HS6STs install sulfate groups at specific sites along the HS chain, and an epimerase (GLCE) converts GlcA to IdoA. The chain is rendered according to the Symbol Nomenclature for Glycans. **b**, GNeo is a highly selective ligand for HS. **c**, Treatment of A375 wild-type (WT), *NDST1* mutant and *SLC35B2* mutant cell lines with increasing concentrations of GNeo-SAP. NDST1 *N*-deacetylates subsets of *N*-acetylglucosamine units and adds *N*-sulfate to the glucosamine residues, which is required for GNeo binding[9]. SLC35B2 is a Golgi transporter for 3′-phosphoadenosine-5′-phosphosulfate, which is the high-energy sulfate donor for sulfation of HS and other glycosaminoglycans[3] (Kruskal–Wallis test on combined dataset, $n=4$). **d**, Flow cytometry analysis of WT, *NDST1* mutant and *SLC35B2* mutant cell lines after incubation with GNeo-biotin conjugated to streptavidin-Cy5 (Mann–Whitney test, $n=4$). **e**, A375 cells, stably expressing Cas9, were transduced with the genome-wide GeCKO sgRNA library and subjected to a screen for resistance against GNeo-SAP (left) and a screen for low GNeo-Cy5 binding (right) (Methods). Gal, galactose; Xyl, xylose; SSC, side scatter. Data are presented as mean ± s.d., $^{***}P<0.001$, $^{*}P<0.05$.

and soft agar. Moreover, our findings suggest that *KDM2B* regulates several components of the matrisome[7], a set of gene products and interactions that make up the extracellular matrix, including proteoglycans.

Results

CRISPR screens identify regulators of HS biosynthesis. Because A375 human melanoma cells express cell surface HS and have been used for genome-wide CRISPR screens[8], these cells were chosen as a model to conduct pooled, genome-wide CRISPR–Cas9 knockout screens to find genes that alter HS on the cell surface. In previous work we described GNeo (Fig. 1b), a highly specific ligand for HS with minimal affinity to other types of cell surface glycans[9]. To interrogate GNeo–HS interactions, biotinylated GNeo with a modified linker (GNeo-biotin; Supplementary Information) was synthesized and conjugated to streptavidin-saporin. Saporin, a type-1 ribosome-inactivating protein from the plant *Saponaria officinalis*, does not penetrate the cell membrane on its own but, when conjugated to GNeo, it becomes a potent HS-dependent cytotoxic compound (GNeo-SAP; Fig. 1c)[9]. GNeo-biotin also was conjugated to streptavidin-Cy5 to generate a fluorescent conjugate (GNeo-Cy5) that binds to HS on the cell surface and can be quantitated using flow cytometry (Fig. 1d). These agents allow parallel, partially orthogonal, screens for sgRNAs that either render cells resistant to GNeo-SAP or that bind less GNeo-Cy5. The agents were tested in A375 cells mutated by CRISPR–Cas9 gene targeting of *SLC35B2* or *NDST1*, factors involved in transport of the active sulfate donor, 3′-phospho-adenyl-5′-phosphosulfate, and *N*-sulfation of glucosamine residues in HS, respectively[3]. Mutations in these genes rendered cells resistant to GNeo-SAP and reduced binding of GNeo-Cy5 (Fig. 1c,d).

To identify novel genes that regulate HS assembly, a genome-wide library of single-gene mutants of A375 cells was generated. A large population (1×10^8) of cells stably expressing the Cas9 nuclease was transduced at a low multiplicity of infection (MOI) (~0.3) with a pool of lentiviruses representing the GeCKO v.2 sgRNA library[10]. We then employed a dual-screening strategy to select for mutants deficient in HS assembly. In the toxin-based resistance screen, a population of GeCKO-transduced cells was treated with GNeo-SAP to enrich for resistant mutants (Fig. 1e). In a parallel screen, cells were sorted using flow cytometry by incubation of cells with GNeo-Cy5, with collection of cells showing decreased binding to the fluorescent probe. In the resistance screen, cells were cultured in GNeo-SAP for 4 days at the dosage required to kill 99%, after

which the toxin was removed and surviving cells were expanded for another 8 days to generate sufficient material for genomic DNA extraction. PBS-treated control populations were cultivated in parallel over the same time span and passaged every other day while maintaining ~100× coverage of the sgRNA library. For the fluorescent activated cell sorter (FACS) screen, GeCKO-transduced A375 cells were harvested after 2 days, incubated with 50 nM GNeo-Cy5 on ice and subjected to cell sorting to separate the GNeo-negative (low) subpopulation from the GNeo-Cy5-positive (high) cells (Fig. 1e). In both screens, cells were harvested, genomic DNA was extracted and sgRNA loci were amplified using nested PCR[8] to find sgRNAs with high fold enrichment over the controls. Sequencing data were processed using the open-source analysis platform PinAPL-Py[11]. Briefly, sgRNA counts were compared between the GNeo-SAP- and PBS-treated populations, and the GNeo-Cy5-low and -high populations, respectively. The fold enrichment data of individual sgRNAs (Supplementary Datasets 1 and 2) targeting the same gene were combined into a single enrichment score for each gene (Methods), thus yielding a ranking of genes based on their degree of positive selection under the screening conditions (Supplementary Datasets 3 and 4). We validated the library construction and screening protocol by conducting a resistance screen using diphtheria toxin (DTX) (Extended Data Fig. 1)[12]. Primary hits in the DTX screen included the membrane-bound precursor of heparin-binding epidermal growth factor receptor, part of the receptor complex for DTX required to elicit its cytotoxic function[13], and genes involved in diphthamide formation (DPH1 and DPH2), the intracellular target of DTX[14].

Analysis of recovered sgRNA sequences from the GNeo-SAP and GNeo-Cy5 screens showed distinct shifts in read count distribution, with read count abundance concentrated on a few sgRNAs and genes (Fig. 2a,b), indicating successful enrichment of positively selected mutants. Enrichment analysis of the screens recovered several known key enzymes of HS biosynthesis, in concordance with the requirement of HS for GNeo binding (Fig. 2c,d). To prioritize genes for further study, we identified 187 overlapping high-ranking genes from both screens that, by gene association analysis as implemented in STRING[15], revealed clusters related to HS biosynthesis, ubiquitination, vesicle transport, microtubules and signaling pathways (Supplementary Dataset 5 and Fig. 2e). We were particularly interested in transcriptional regulators of HS assembly[6], and focused on one cluster comprised of transcription factors and genes involved in chromatin modification (Fig. 2e). This cluster included *KDM2B*, a histone demethylase that represses cell differentiation pathways and can promote oncogenic proliferation[16] with no known ties to HS biosynthesis. Interestingly, analysis of the sequencing data with more restrictive significance thresholds (Supplementary Datasets 6 and 7) resulted in *KDM2B* being shown as the only gene shared between the top-scoring hits from the GNeo-SAP and GNeo-Cy5 screens. Thus, *KDM2B* was chosen for further study.

***KDM2B* inactivation reduces binding of GNeo.** KDM2B, a Jumonji (JmjC) domain histone demethylase, is a chromatin-remodeling protein with many roles in the regulation of gene expression[17]. It recruits the polycomb repressor complex 1 (PRC1)[18] and demethylates histone marks H3K36me2 (ref. [19]) and H3K4me3 (ref. [20]), leading to altered gene expression. To validate and investigate *KDM2B* as a potential regulator of HS assembly, we generated a *KDM2B* mutant cell line in A375 cells (hereafter referred to as *KDM2B*^{C5}) using CRISPR–Cas9 (Fig. 3a, Extended Data Fig. 2a and the source data for Fig. 3). *KDM2B*^{C5} cells showed an increase of more than threefold in resistance to GNeo-SAP killing (half-maximal inhibitory concentration (IC_{50} = 1.3 nM) compared to A375-Cas9 wild-type cells (IC_{50} = 0.4 nM) (Fig. 3b). Additionally, *KDM2B*^{C5} cells showed a decrease of ~two-fold in GNeo-Cy5 binding to HS on the cell surface (Fig. 3c).

***KDM2B* inactivation results in alterations in HS structure.** The reduced binding and uptake of GNeo conjugates to cells suggested that *KDM2B* might alter the amount or structure of HS. To determine this possibility, we isolated cell surface HS from wild-type and *KDM2B*^{C5} cells by trypsin digestion and anion-exchange chromatography. The samples were subsequently depolymerized into disaccharides using mixed heparin lyases. Depolymerization products were then aniline tagged by reductive amination and analyzed by ion-pairing, reverse-phase chromatography and quantitative high-resolution mass spectrometry (MS) with mass standards. Analysis of the various disaccharides did not reveal a significant change in disaccharide composition, although there was a trend in reduction of the trisulfated disaccharide D2S6 (Fig. 3d and Supplementary Table 1). The overall degree of sulfation also did not differ between mutant and wild type (Supplementary Table 2) and, contrary to expectations based on GNeo binding, the overall amount of cell surface HS increased somewhat in *KDM2B*^{C5} cells (Fig. 3d, inset).

To examine HS structure in greater detail, wild-type and *KDM2B*^{C5} cells were radiolabeled with $^{35}SO_4$ for 24 h at 37 °C, and [^{35}S]HS was purified by anion-exchange chromatography. Size-exclusion chromatography of [^{35}S]HS from wild-type and *KDM2B*^{C5} cells showed no significant difference in overall molecular mass, with an average value of ~36 kDa (Fig. 3e). Fractionation of [^{35}S]HS from wild-type or *KMD2B*^{C5} cells over a GNeo-Sepharose affinity column (Methods) using stepwise elution with buffers containing 0.1–1.5 M NaCl showed that *KDM2B*^{C5} HS exhibited an altered profile, with a significant decrease in binding at the highest (1–1.5 M) and lowest (0.1–0.4 M) salt concentrations compared to wild-type HS (Fig. 3f). [^{35}S]HS was also treated with heparin lyase III, which cleaves the chains in regions devoid of sulfation and IdoA. Fractionation of the liberated sulfated domains over the GNeo affinity column revealed a significant increase in recovery of [^{35}S]HS fragments from *KDM2B*^{C5} cells in the flowthrough fraction (0.1 M), and a significant decrease in domains requiring high salt for elution (0.8–1 M) (Fig. 3g). These findings indicate that inactivation of *KDM2B* results in changes in the overall domain structure of HS.

***KDM2B* regulates growth factor binding to HS.** To investigate whether the structural changes in HS derived from *KDM2B*^{C5} cells could affect HS–protein interactions, we quantified binding of a library of known HS-interacting proteins to wild-type and *KDM2B*^{C5} cells using flow cytometry. Dramatically altered binding of a subset of proteins was noted. For example, binding of human fibroblast growth factor 1 (FGF1), FGF7, FGF10, S100 calcium-binding protein A12 (S100A12)[21] and antithrombin was reduced to differing extents, whereas binding of FGF2, interleukin-8 and neuropilin-1 (NRP1) did not exhibit any differences in binding to the wild type and mutant (Fig. 4a). No difference was detected in binding of an antibody recognizing *N*-sulfated regions of HS (10E4), consistent with the lack of any change in *N*-sulfated disaccharides (Fig. 4a). Binding of an antibody that detects an HS neoepitope generated by heparin lyase digestion (3G10) was unaltered (Fig. 4a), indicating no change in the number of HS chains expressed on the cell surface. To corroborate these findings we identified a second clone bearing different mutant alleles, which gave a similar binding profile (Extended Data Fig. 2a and Extended Data Fig. 3a).

To investigate more directly the interaction of HS with FGF1 and antithrombin, we mixed [^{35}S]HS from *KDM2B*^{C5} mutant and wild-type cells with these proteins and rapidly filtered the mixture through a nitrocellulose membrane. Unbound [^{35}S]HS flowed through the filter whereas [^{35}S]HS bound to FGF1 or antithrombin was trapped. Quantitation of bound [^{35}S]HS showed that material derived from the mutant bound to a lesser extent (Fig. 4b).

Finally, to rule out off-target effects of gene targeting in *KDM2B*^{C5} cells, we transfected full-length human *KDM2B* complementary DNA back into the mutant clone and showed restoration of cell

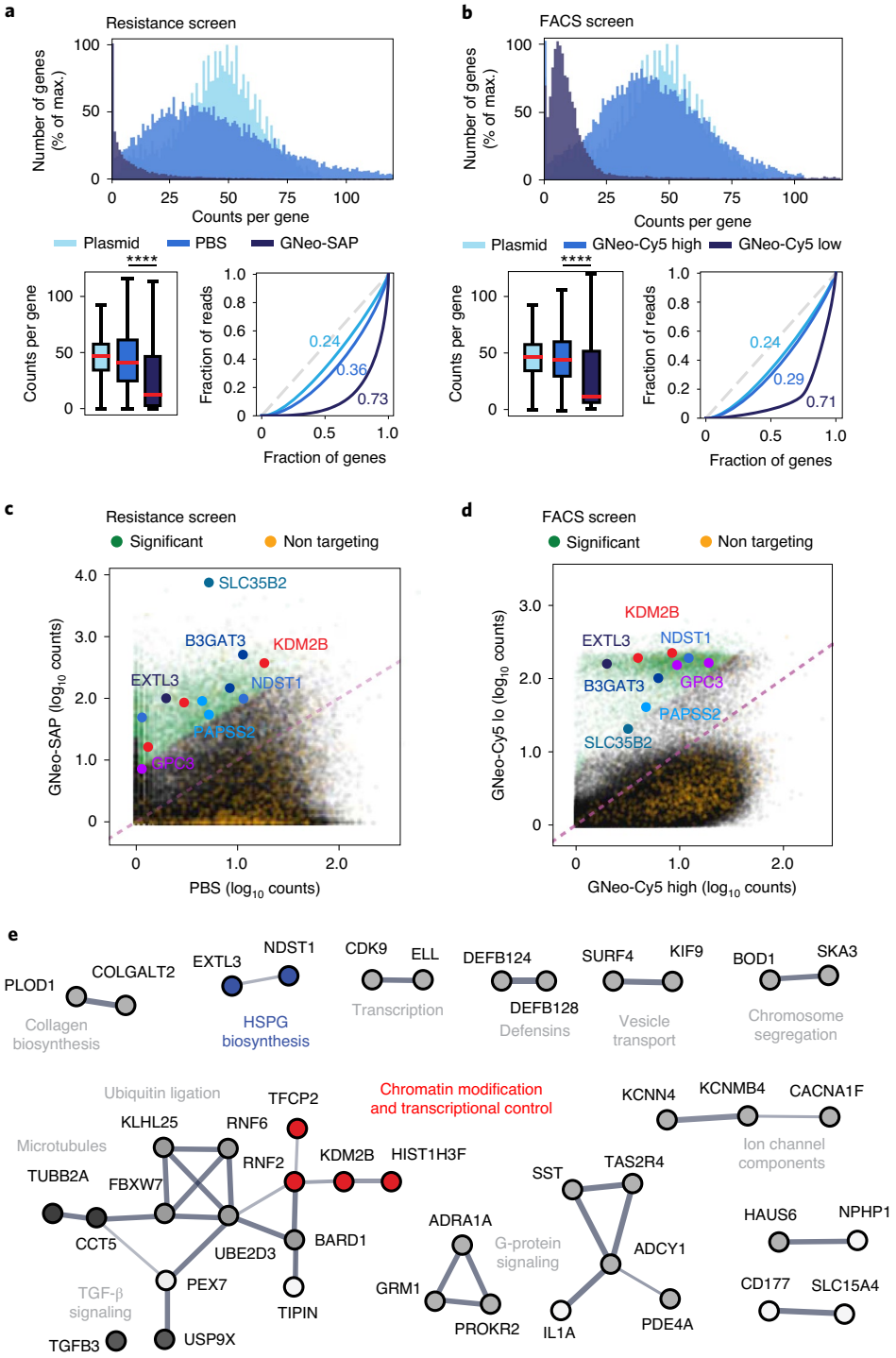


Fig. 2 | Sequencing data analysis of CRISPR–Cas9 screens selecting for regulators of HS biosynthesis. a, Top: frequency distribution of gene counts after treatment with either 10 nM GNeo-SAP or PBS (plasmid, GeCKO v2 plasmid library). Bottom left: boxplots showing the gene count distribution for each sample (Kolmogorov–Smirnov test; $n = 21{,}915$; boxplot whiskers showing 5 and 95% percentiles, outliers omitted for clarity). Bottom right: Lorenz curves showing the distribution of sequencing reads over the gene library. Numbers represent Gini coefficients (0: reads cover gene library evenly; 1: reads cover only a single gene). **b**, Top: frequency distribution of gene counts after incubation with GNeo-Cy5 and collection of both low- and high-binding cell fractions (plasmid, GeCKO v2 plasmid library). Bottom left: boxplots showing the gene count distribution for each sample (Kolmogorov–Smirnov test; $n = 21{,}915$; boxplot whiskers showing 5 and 95% percentiles, outliers omitted for clarity). Bottom right: Lorenz curves showing the distribution of sequencing reads over the gene library. Numbers represent Gini coefficients. **c**, Scatterplot of sgRNA counts ($\log_{10}$, normalized) in samples recovered after GNeo-SAP versus PBS treatment. **d**, Scatterplot of sgRNA counts ($\log_{10}$, normalized) in the low- versus the high-binding cell fraction after incubation with GNeo-Cy5. **c**,**d**, The sgRNA fraction with significant fold change is shown in green, and the sgRNA fraction representing nontargeting controls is shown in orange. The most significantly enriched sgRNAs targeting HS enzymes or solute carriers required for HS biosynthesis are captioned. **e**, Protein–protein interaction network of top enriched targets in the GNeo-SAP and GNeo-Cy5 screens. Line width indicates the confidence level of the interaction. Proteins without direct interactions with other proteins from the list were omitted for clarity (Supplementary Dataset 5 provides a full list). Data are presented as mean ± s.d., $^{****}P < 0.0001$.

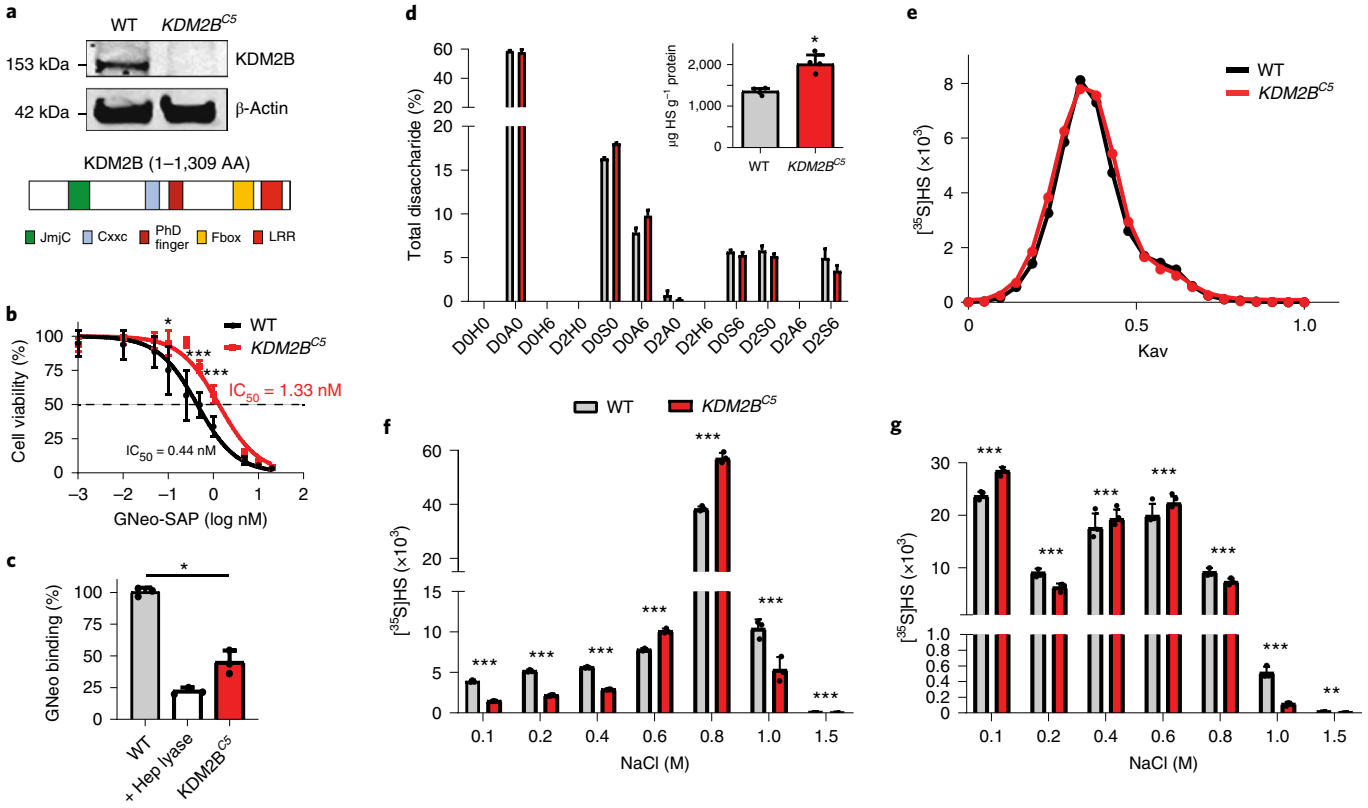


Fig. 3 | Structural analysis of cell surface HS in *KDM2B* mutant cells. a, Immunoblot of CRISPR-mediated inactivation of *KDM2B* in A375 cells and protein domains of human KDM2B. **b**, CellTiter-Blue assay showing increased resistance to GNeo-SAP conjugate over a wide concentration range (0–20 nM) (Mann–Whitney test, $n = 8$). **c**, Decreased cell surface binding of GNeo-biotin-streptavidin-Cy5 to *KDM2B*^{C5} cells compared to wild-type cells. A375 cells pretreated with heparin (hep) lyases (I–III) were used as a positive control (*t*-test on $\log_{10}$ fluorescence data, $n = 3$). **d**, Liquid chromatography–mass spectrometry (LC–MS) quantification of disaccharides from HS in wild-type and *KDM2B*^{C5} cells (*z*-test with absolute quantities in pmol; $n = 3$). Absolute values for disaccharides and the different classes of disaccharides are shown in Supplementary Table 1. The disaccharide structure code is described in Supplementary Table 1 and ref. [49]. Inset: LC–MS quantification of total HS in wild-type and *KDM2B*^{C5} cells (Mann–Whitney test, $n = 3$). **e**, Gel filtration chromatography (Sepharose CL-6B) of [^{35}S]HS from wild-type and *KDM2B*^{C5} cells. **f**, [^{35}S]HS was fractionated over a GNeo affinity column (0.1–1.5 M NaCl) (Poisson *C*-test, $n = 3$). **g**, [^{35}S]HS was digested with heparin lyase III (10 mU) overnight at 37 °C before fractionation over a GNeo affinity column (Poisson *C*-test, $n = 3$). AA, amino acids. Data are presented as mean ± s.d., $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$. Kav is a function of the elution volume of a molecule and indicates the ratio between the elution volume of a given molecule and the total available volume of the column.

surface binding of FGF1 (Fig. 4c and Extended Data Fig. 3b). Similarly, CRISPR-mediated inactivation of *KDM2B* in HeLa cells showed a significant decrease in FGF1 binding and restoration of binding following transduction of full-length human *KDM2B* (Extended Data Figs. 2b and 3c). We generated a mutant KDM2B construct with the demethylase domain inactivated (H242A) and we transfected A375 *KDM2B*^{C5} cells. The mutant cDNA was expressed (Extended Data Fig. 3d) but did not rescue FGF1 binding (Fig. 4c). In summary, KDM2B plays an important role in controlling the binding affinities of HS-dependent ligands, which is probably conserved across human cell lines.

***KDM2B* controls HS enzyme expression.** Considering the notable changes in HS–protein binding and alteration in HS structure, the expression level of various enzymes was analyzed by RNA-seq. While two HS 6-*O*-sulfotransferases showed an increase in expression (*HS6ST1* and *HS6ST3*), four sulfotransferases showed reduced expression following depletion of *KDM2B* (*HS3ST1*, *HS3ST3A1*, *HS3ST3B1* and *HS6ST2*). Additionally, the expression of HS endosulfatases *SULF1* and *SULF2* was increased (Fig. 5a). Thus, most transcriptional changes among HS enzymes pointed to a potential decrease in 3-*O*- and 6-*O*-sulfation levels in *KDM2B*^{C5} cells. *SULF1* expression, in particular, was dramatically elevated at both the mRNA and protein level (Extended Data Fig. 4a,b and Source Extended Data Fig. 4).

To investigate whether KDM2B regulates the methylation of these genes, we performed ChIP–seq on mutant and wild-type cells using antibodies to two known *KDM2B* histone methylation targets, H3K4me3 and H3K36me2 (refs. [19,20]). Although the level of overall H3K4me3 was increased at KDM2B-occupied sites in the mutant, that of H3K36me2 was not affected by loss of KDM2B (Fig. 5b). Additionally, we found that H3K4me3 methylation at the promoter of *MXN1*, a known target of *KDM2B*[22], was significantly elevated in *KDM2B*^{C5} cells compared to wild type (Extended Data Fig. 4c). Interestingly, ChIP–seq for H3K4me3 showed an increase in methylation extending from the promoter in *SULF1* into the gene body, but not in the *B4GALT7* gene (Fig. 5c). ChIP–seq analysis of H3K36me2 did not reveal alterations in methylation in the *SULF1* gene. These findings indicate that KDM2B normally represses *SULF1* expression through the removal of activating H3K4me3 methylation marks.

In addition to transcript level changes of HS biosynthesis enzymes, we found differential expression of several additional gene groups in the transcriptome of *KDM2B* mutant cells. Gene set enrichment analysis identified ontology classes related to the extracellular matrix (ontology classes *Core Matrisome*, *Matrisome-associated*

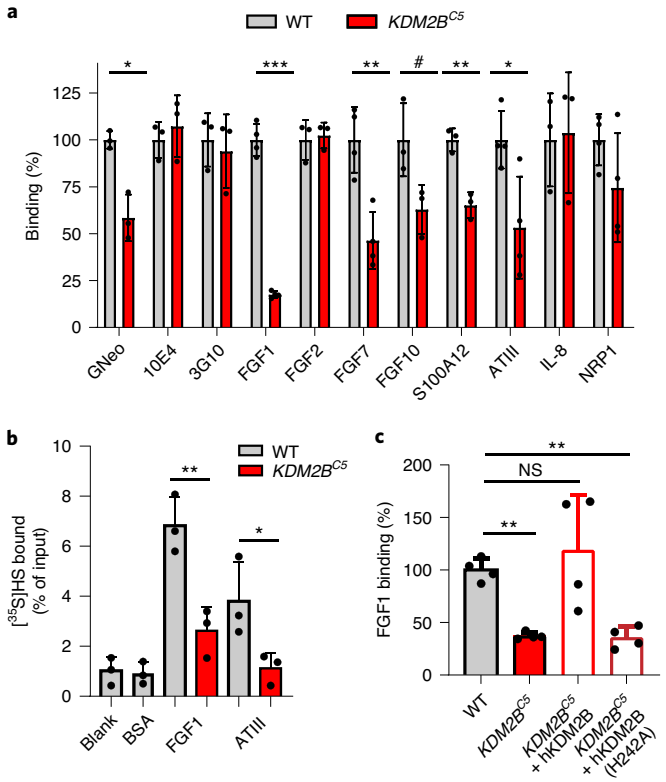


Fig. 4 | Binding of growth factors and other HS-binding proteins to *KDM2B* mutant cells. a, *KDM2B*^{C5} cells show a significant decrease in binding of a subset of protein ligands (*t*-test on $\log_{10}$ fluorescence data, $n=3$). **b**, Filter binding assays show decreased binding of FGF1 and antithrombin to [^{35}S]HS from *KDM2B*^{C5} cells (*t*-test, $n=3$). **c**, Reintroduction of wild-type *KDM2B* cDNA, but not a *KDM2B* catalytic mutant (H242A), restored binding of FGF1 in A375 (*t*-test on $\log_{10}$ fluorescence data, $n=4$). NS, not significant. Data are presented as mean ± s.d., $^{***}P<0.001$, $^{**}P<0.01$, $^{*}P<0.05$, $^{\#}P=0.05$.

and *Extracellular Matrix Organization*) (Extended Data Fig. 4d,e and Supplementary Dataset 8). For example, we detected elevated protein levels of MMP-9, a secreted metalloproteinase that degrades multiple extracellular matrix components, as well as decreased levels of TIMP-3, a matrix metalloprotease inhibitor in the growth medium of *KDM2B*^{C5} cells (Extended Data Fig. 4f). These results indicate that, beyond its novel role as a transcriptional regulator of HS assembly, KDM2B may represent a principal factor controlling many components of the extracellular matrix.

Rescue of *SULF1* restores ligand binding in *KDM2B*^{C5} cells. To determine if transcriptional changes in these HS biosynthetic genes are responsible for the differential changes in binding of multiple proteins to cell surface HS, we investigated whether targeted rescue experiments in *KDM2B*^{C5} cells would restore binding of FGF1, FGF7, FGF10, S100A12 and antithrombin. We focused specifically on *SULF1* and *HS6ST2*, due to previous reports that binding of FGF1, FGF7 and FGF10 to HS requires 6-*O* sulfation[23–25], and on *HS3ST1* and *HS3ST3A1*, because antithrombin has been shown to bind 3-*O*-sulfated sequences of HS[26]. SULF1 is an endo-6-*O*-sulfatase that modulates HS–protein interactions in the extracellular space by removing a subset of 6-*O*-sulfate groups from HS[27], and HS6ST2 is a member of a family of sulfotransferases that catalyze 6-*O*-sulfation of *N*-acetylated and *N*-sulfoglucosamine residues of HS[28]. HS3ST1 and HS3ST3A1 are two of seven isozymes that catalyze 3-*O*-sulfation of glucosamine residues of HS and create binding sites for antithrombin and a small subset of proteins[29]. Indeed, small interfering RNA knockdown of *SULF1* in *KDM2B*^{C5} cells restored binding of FGF1, FGF7 and FGF10 (Fig. 5d and Extended Data Fig. 4a,g), corroborating that KDM2B affects binding of these ligands through its regulation of 6-*O*-sulfation of HS. Overexpression of *HS6ST2*, however, did not rescue binding of any ligands (Fig. 5d and Extended Data Fig. 4h), probably because elevated expression levels of SULF1 in *KDM2B*^{C5} cells antagonize its sulfotransferase action, prohibiting a net increase in 6-*O* sulfation of HS at the cell surface. Interestingly, *HS3ST3A1* and *HS3ST1* overexpression rescued binding of S100A12, a calcium-binding protein that binds strongly to highly sulfated HS chains[21]. Additionally, antithrombin binding in *KDM2B*^{C5} cells was also rescued by overexpression of 3-*O*-sulfotransferases (Fig. 5d and Extended Data Fig. 4i), in accordance with its requirement for specific 3-*O*-sulfated binding sequences[26].

***KDM2B* inactivation impedes SULF1-dependent growth.** Heparan sulfate plays a key role as a coreceptor for several growth factors that activate mitogenic signaling pathways, including members of the FGF signaling pathway that exhibit altered binding in *KDM2B* mutant cells (Fig. 4a). To test whether alteration in HS fine structure might impair cellular proliferation, we examined the growth properties of *KDM2B*^{C5} cells. In monolayer culture, *KDM2B*^{C5} cells divided more slowly than wild-type cells in growth medium containing 10% fetal bovine serum (FBS), with a doubling time of 30 versus 18 h, respectively (Fig. 6a). Under starvation conditions (2% FBS), *KDM2B*^{C5} cells essentially ceased to grow whereas wild-type A375 cells continued to divide, albeit with a reduced doubling time (~27 h) (Fig. 6a). Guided by the observation that silencing of *SULF1* restored the binding of several growth factors in *KDM2B*^{C5} cells (Fig. 5d), we inactivated *SULF1* in *KDM2B*^{C5} cells through CRISPR–Cas9-mediated gene targeting (Extended Data Fig. 5a). Interestingly, inactivation of SULF1 restored growth in 2% serum and enhanced growth under standard conditions (doubling time, 21.8 h) (Fig. 6a). We also saw reduced colony formation in monolayer culture (Extended Data Fig. 5b) and anchorage-independent growth of *KDM2B*^{C5} cells in soft agar (Fig. 6b). In concordance with the observed restoration of growth rates, inactivation of *SULF1* in *KDM2B*^{C5} cells increased colony number and colony size (Fig. 6b). These findings indicate that KDM2B regulates proliferation of A375 cells and that this regulation is partially mediated through repression of *SULF1*. In conclusion, we identified KDM2B as a novel transcriptional regulator of HS assembly and showed that *KDM2B* constitutes a critical genetic component in the maintenance of proliferation of melanoma cells, by controlling HS fine structure.

Discussion

CRISPR–Cas9 screens have become an invaluable discovery tool to unravel metabolic and signaling pathways. Most screens involve direct selection of mutants resistant to a toxin[30] or viral infection[31], or through enrichment by cell sorting[32]. An innovative feature of this study is the application of two orthogonal screens to identify factors involved in HS composition: flow cytometry with GNeo-Cy5 as a ligand and resistance to GNeo-SAP. Previous studies have shown that GNeo binds selectively to HS on the cell surface[9], and thus each screen was expected to enrich for mutations in genes affecting the presentation and composition of HS on the cell surface. Both screens recovered both unique and shared mutations. Thus both screening methods probably select for mutants in screen-specific pathways (for example, mutants conferring metabolic resistance to saporin), but also shared pathways (for example, HS biosynthesis) that we enriched for by matching the targets obtained from the two screens. To date, very little is known about genetic regulation of the enzymes responsible for biosynthesis of HS[6]. The dual-screening method recovered expected hits, such as HS biosynthetic enzymes (*NDST1*

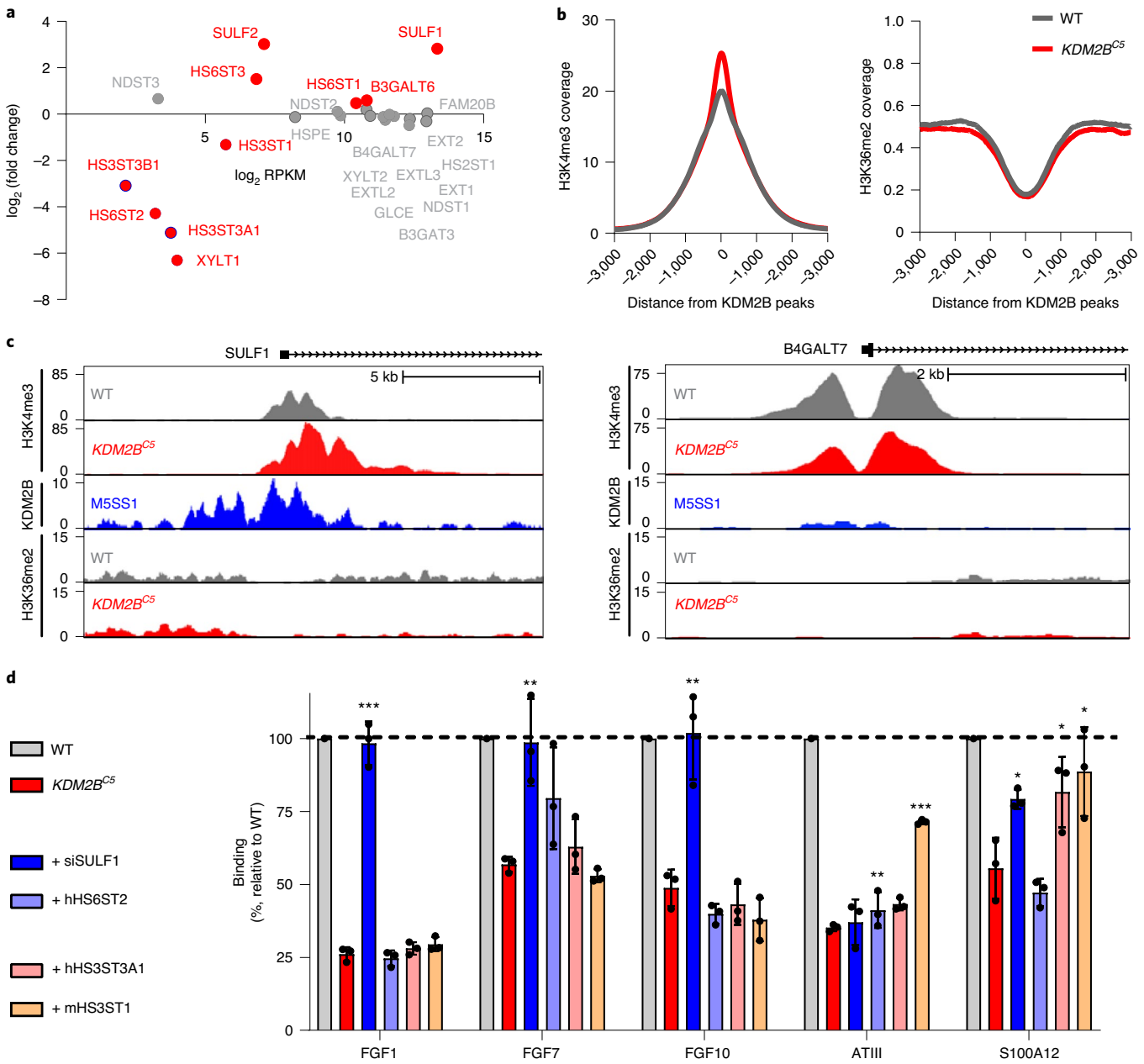


Fig. 5 | Transcriptome and ChIP analysis of *KDM2B* mutant cells. a, mRNA expression of HS biosynthetic enzymes and HS-modifying enzymes from RNA-seq (differentially expressed genes shown in red). **b**, H3K4me3 and H3K36me2 ChIP-seq signals normalized by input sample in WT and *KDM2B*^{C5} cells at KDM2B-occupied genomic regions (GSE108929). **c**, Gene track for H3K4me3, KDM2B and H3K36me2 ChIP-seq at the *SULF1* (left) and *B4GALT7* (right) loci. **d**, Ligand binding of *KDM2B*^{C5} cells transfected with indicated expression plasmids (hHS6ST2/hHS3ST3A1/mHS3ST1) or a siRNA-targeting SULF1 relative to mock-transfected wild-type (WT) cells (*t*-test, n = 3). RPKM, reads per kilobase of transcript per million mapped reads. Data are presented as mean ± s.d., $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$.

and *EXTL3*), but also several unexpected pathways indicating that the presence of HS on the cell surface depends on the regulation of other cellular processes (for example, signaling and chromatin remodeling). Importantly, our gene-ranking method is restrictive in that it grants higher scores to genes for which more sgRNAs were enriched. This can lead to genes receiving lower scores if they had only one or two sgRNAs enriched due to imperfect targeting by the CRISPR library, which could explain why we did not recover all HS biosynthetic enzymes in the top tier of the target list.

The discovery of KDM2B as a regulator of HS assembly is novel. KDM2B binds DNA through a zinc-finger domain and executes its gene-regulatory function in two ways: it can act as a JmjC-demethylase, targeting H3 methyl groups, and it can act as a recruiter of PRC1 to nonmethylated CpG islands in promoter regions[18]. Both processes typically lead to gene silencing[16], although JmjC-domain-containing demethylases also target inactivating methylation marks leading to gene activation[33]. The inability of a transfected catalytic mutant (H242A) to rescue FGF1 binding in *KDM2B*^{C5} cells, and the increase in H3K4me3 methylation at the *SULF1* promoter, suggest that KDM2B regulates HS biosynthesis through its demethylase activity. Interestingly, RNF2, another key component of PRC1, was highly enriched in both screens (Supplementary Dataset 5), strongly suggesting that HS biosynthetic genes could be regulated through PRC1.

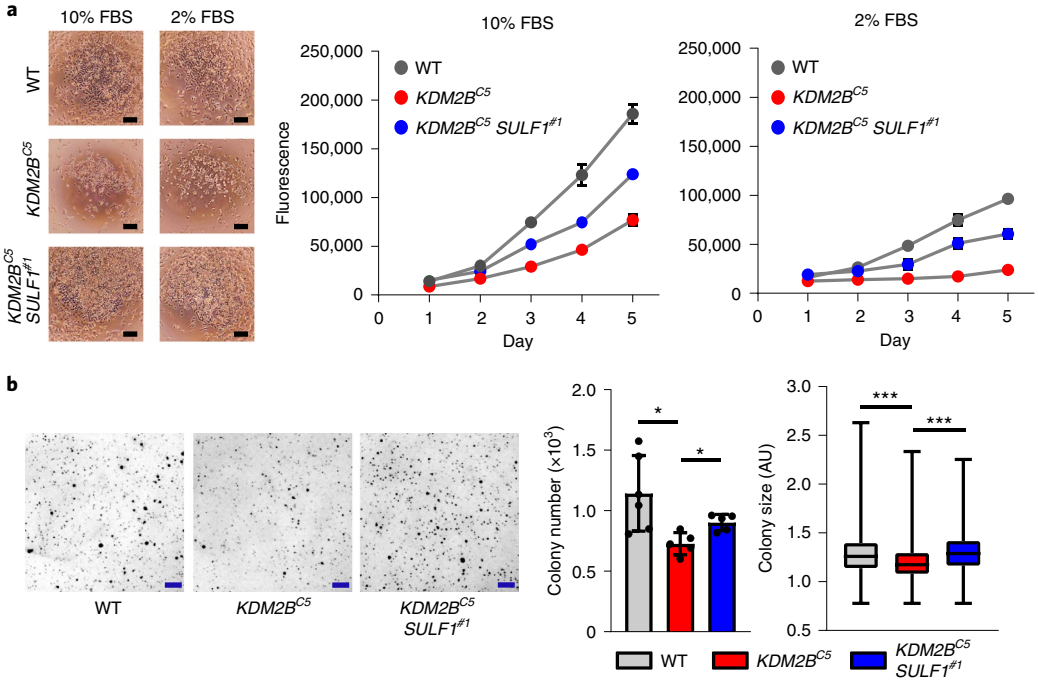


Fig. 6 | *KDM2B* inactivation results in a SULF1-dependent decrease in cell growth. a, Viable cell density was monitored by CellTiter-Blue over 5 days under normal (10% FBS) and low-serum (2% FBS) conditions ($P < 0.001$. Two-way analysis of variance; $n = 3$). Images taken 48 h after seeding. Scale bars, 100 µm. **b**, Anchorage-independent growth was analyzed by soft agar assay. Colony number and size were measured after 11 days of incubation (colony number: *t*-test, $n = 5$; colony size: *t*-test on pooled dataset, $n = 3{,}000$). Scale bars, 1 mm. AU, arbitrary units. Data are presented as mean ± s.d., $^{***}P < 0.001$, $^{*}P < 0.05$.

The formation of binding sites in HS depends on the activity of multiple enzymes that determine the arrangement of sulfate groups and IdoA residues along the chain. KDM2B modulates the expression of several of these enzymes. A common feature of certain growth factors whose binding was altered (FGF1, FGF7 and FGF10) is a requirement for 6-*O*-sulfate groups on specific glucosamine residues in HS. Interestingly, bottom-up compositional studies did not reveal an overall change in composition, most likely because SULF1 removes a minor, but critical, subset of 6-*O*-sulfate groups. Previous studies have shown that SULF1 causes differential impact on the extent of 6-*O*-sulfation depending on the tissue[34,35]. Other methods are needed to examine the domain structure and sequence[36–38].

Loss of function of *KDM2B* has been noted in other cell lines and correlated with reduced cell growth and viability in glioma cells[39], triple-negative breast cancer cells[40] and others[40]. The cause of reduced cell growth has been attributed to alterations in the cell cycle[41], ribosome biogenesis[42], expression of other epigenetic factors[43], autophagy[44] and growth factor signaling[45]. Our findings indicate that, in melanoma cells, the regulation of cell growth by *KDM2B* is at least partially mediated by its regulation of HS structure. Our transcriptomic analysis also indicates that additional effects on growth could be mediated by changes in extracellular matrix composition. Other phenotypes ascribed to KDM2B based on inactivation of the gene in mice, such as stem cell proliferation and lineage commitment, should be re-evaluated in regard to the fact that they be related to changes in HS proteoglycan or extracellular matrix composition[46,47].

Intriguingly, KDM2B regulates not only HS biosynthesis but also a notable number of genes encoding components of the extracellular matrix (ECM) or ECM remodeling, including two key regulators of ECM turnover, MMP-9 and TIMP-3. In accordance with these findings, a recent study reported that *KDM2B* inhibition results in increased expression of several ECM proteins in synovial sarcoma cells[22]. Interestingly, several genes in the matrisome have previously been recognized as PRC1 targets[48]. We cannot exclude the possibility that some of the changes in expression of matrisome proteins could be indirectly linked to *KDM2B* through changes in HS. Nevertheless, our observations suggest that the genes involved in determining the fine structure of HS are part of a larger genetic network that determines the composition and architecture of the ECM, and that changes in ECM composition may underlie the alterations in cell division caused by *KDM2B* inactivation.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-021-00776-9.

Received: 28 May 2020; Accepted: 18 February 2021;
Published online: 12 April 2021

References

1. Xu, D. & Esko, J. D. Demystifying heparan sulfate-protein interactions. *Annu. Rev. Biochem.* **83**, 129–157 (2014).
2. Bishop, J. R., Schuksz, M. & Esko, J. D. Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature* **446**, 1030–1037 (2007).
3. Esko, J. D. & Selleck, S. B. Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem.* **71**, 435–471 (2002).
4. Presto, J. et al. Heparan sulfate biosynthesis enzymes EXT1 and EXT2 affect NDST1 expression and heparan sulfate sulfation. *Proc. Natl Acad. Sci. USA* **105**, 4751–4756 (2008).
5. Kreuger, J. & Kjellen, L. Heparan sulfate biosynthesis: regulation and variability. *J. Histochem. Cytochem.* **60**, 898–907 (2012).
6. Weiss, R. J. et al. ZNF263 is a transcriptional regulator of heparin and heparan sulfate biosynthesis. *Proc. Natl Acad. Sci. USA* **117**, 9311–9317 (2020).

7. Naba, A. et al. The extracellular matrix: tools and insights for the "omics" era. *Matrix Biol.* **49**, 10–24 (2016).
8. Shalem, O. et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
9. Elson-Schwab, L. et al. Guanidinylated neomycin delivers large, bioactive cargo into cells through a heparan sulfate-dependent pathway. *J. Biol. Chem.* **282**, 13585–13591 (2007).
10. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
11. Spahn, P. N. et al. PinAPL-Py: a comprehensive web-application for the analysis of CRISPR/Cas9 screens. *Sci. Rep.* **7**, 15854 (2017).
12. Zhou, Y. et al. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487–491 (2014).
13. Shishido, Y., Sharma, K. D., Higashiyama, S., Klagsbrun, M. & Mekada, E. Heparin-like molecules on the cell surface potentiate binding of diphtheria toxin to the diphtheria toxin receptor membrane-anchored heparin-binding epidermal growth factor-like growth factor. *J. Biol. Chem.* **270**, 29578–29585 (1995).
14. Su, X., Lin, Z. & Lin, H. The biosynthesis and biological function of diphthamide. *Crit. Rev. Biochem. Mol. Biol.* **48**, 515–521 (2013).
15. Szklarczyk, D. et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
16. Yan, M., Yang, X., Wang, H. & Shao, Q. The critical role of histone lysine demethylase KDM2B in cancer. *Am. J. Transl. Res.* **10**, 2222–2233 (2018).
17. Vacik, T., Ladinovic, D. & Raska, I. KDM2A/B lysine demethylases and their alternative isoforms in development and disease. *Nucleus* **9**, 431–441 (2018).
18. Farcas, A. M. et al. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *eLife* **1**, e00205 (2012).
19. He, J. et al. Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nat. Cell Biol.* **15**, 373–384 (2013).
20. Frescas, D., Guardavaccaro, D., Bassermann, F., Koyama-Nasu, R. & Pagano, M. JHDM1B/FBXL10 is a nucleolar protein that represses transcription of ribosomal RNA genes. *Nature* **450**, 309–313 (2007).
21. Zhang, X., Ong, C., Su, G., Liu, J. & Xu, D. Characterization and engineering of S100A12-heparan sulfate interactions. *Glycobiology* **30**, 463–473 (2020).
22. Banito, A. et al. The SS18-SSX oncoprotein hijacks KDM2B-PRC1.1 to drive synovial sarcoma. *Cancer Cell* **34**, 346–348 (2018).
23. Ashikari-Hada, S. et al. Characterization of growth factor-binding structures in heparin/heparan sulfate using an octasaccharide library. *J. Biol. Chem.* **279**, 12346–12354 (2004).
24. Kreuger, J., Salmivirta, M., Sturiale, L., Giménez-Gallego, G. & Lindahl, U. Sequence analysis of heparan sulfate epitopes with graded affinities for fibroblast growth factors 1 and 2. *J. Biol. Chem.* **276**, 30744–30752 (2001).
25. Sugaya, N., Habuchi, H., Nagai, N., Ashikari-Hada, S. & Kimata, K. 6-O-sulfation of heparan sulfate differentially regulates various fibroblast growth factor-dependent signalings in culture. *J. Biol. Chem.* **283**, 10366–10376 (2008).
26. Thacker, B. E., Xu, D., Lawrence, R. & Esko, J. D. Heparan sulfate 3-*O*-sulfation: a rare modification in search of a function. *Matrix Biol.* **35**, 60–72 (2013).
27. Lamanna, W. C., Frese, M. A., Balleininger, M. & Dierks, T. Sulf loss influences N-, 2-O-, and 6-O-sulfation of multiple heparan sulfate proteoglycans and modulates fibroblast growth factor signaling. *J. Biol. Chem.* **283**, 27724–27735 (2008).
28. Habuchi, H. et al. The occurrence of three isoforms of heparan sulfate 6-*O*-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted *N*-sulfoglucosamine. *J. Biol. Chem.* **275**, 2859–2868 (2000).
29. Shworak, N. W. et al. Molecular cloning and expression of mouse and human cDNAs encoding heparan sulfate D-glucosaminyl 3-O-sulfotransferase. *J. Biol. Chem.* **272**, 28008–28019 (1997).
30. Tian, S. et al. Genome-wide CRISPR screens for Shiga toxins and ricin reveal Golgi proteins critical for glycosylation. *PLoS Biol.* **16**, e2006951 (2018).
31. Labeau, A. et al. A genome-wide CRISPR-Cas9 screen identifies the dolichol-phosphate mannose synthase complex as a host dependency factor for dengue virus infection. *J. Virol.* **94**, e01751–19 (2020).
32. Bassaganyas, L. et al. New factors for protein transport identified by a genome-wide CRISPRi screen in mammalian cells. *J. Cell Biol.* **218**, 3861–3879 (2019).
33. Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* **7**, 715–727 (2006).
34. Nagamine, S. et al. Organ-specific sulfation patterns of heparan sulfate generated by extracellular sulfatases Sulf1 and Sulf2 in mice. *J. Biol. Chem.* **287**, 9579–9590 (2012).
35. Lamanna, W. C. et al. Heparan sulfate 6-O-endosulfatases: discrete in vivo activities and functional co-operativity. *Biochem. J.* **400**, 63–73 (2006).
36. Miller, R. L. et al. Shotgun ion mobility mass spectrometry sequencing of heparan sulfate saccharides. *Nat. Commun.* **11**, 1481 (2020).
37. Wu, J. et al. Sequencing heparan sulfate using HILIC LC-NETD-MS/MS. *Anal. Chem.* **91**, 11738–11746 (2019).
38. Wei, J. et al. Characterization and quantification of highly sulfated glycosaminoglycan isomers by gated-trapped ion mobility spectrometry negative electron transfer dissociation MS/MS. *Anal. Chem.* **91**, 2994–3001 (2019).
39. Wang, Y. et al. KDM2B overexpression correlates with poor prognosis and regulates glioma cell growth. *Onco Targets Ther.* **11**, 201–209 (2018).
40. Zheng, Q. et al. Histone demethylase KDM2B promotes triple negative breast cancer proliferation by suppressing p15INK4B, p16INK4A, and p57KIP2 transcription. *Acta Biochim. Biophys. Sin. (Shanghai)* **50**, 897–904 (2018).
41. Tzatsos, A. et al. Lysine-specific demethylase 2B (KDM2B)-let-7-enhancer of zester homolog 2 (EZH2) pathway regulates cell cycle progression and senescence in primary cells. *J. Biol. Chem.* **286**, 33061–33069 (2011).
42. Galbiati, A. et al. Epigenetic up-regulation of ribosome biogenesis and more aggressive phenotype triggered by the lack of the histone demethylase JHDM1B in mammary epithelial cells. *Oncotarget* **8**, 37091–37103 (2017).
43. Kuang, Y. et al. Histone demethylase KDM2B upregulates histone methyltransferase EZH2 expression and contributes to the progression of ovarian cancer in vitro and in vivo. *Onco Targets Ther.* **10**, 3131–3144 (2017).
44. Zhao, E. et al. Inhibition of cell proliferation and induction of autophagy by KDM2B/FBXL10 knockdown in gastric cancer cells. *Cell Signal.* **36**, 222–229 (2017).
45. Kottakis, F. et al. FGF-2 regulates cell proliferation, migration, and angiogenesis through an NDY1/KDM2B-miR-101-EZH2 pathway. *Mol. Cell* **43**, 285–298 (2011).
46. Andricovich, J., Kai, Y., Peng, W., Foudi, A. & Tzatsos, A. Histone demethylase KDM2B regulates lineage commitment in normal and malignant hematopoiesis. *J. Clin. Invest.* **126**, 905–920 (2016).
47. Ye, B. et al. LncKdm2b controls self-renewal of embryonic stem cells via activating expression of transcription factor Zbtb3. *EMBO J.* **37**, e97174 (2018).
48. Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H. & Helin, K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.* **20**, 1123–1136 (2006).
49. Lawrence, R., Lu, H., Rosenberg, R. D., Esko, J. D. & Zhang, L. Disaccharide structure code for the easy representation of constituent oligosaccharides from glycosaminoglycans. *Nat. Methods* **5**, 291–292 (2008).

Methods

Cell culture. HEK293T cells (ATCC CRL-3216) and A375 cells (ATCC CRL-1619) were grown in DMEM (Gibco) supplemented with 10% (v/v) FBS (Gibco) and 1% (v/v) penicillin/streptomycin (Gibco) at 37 °C under an atmosphere of 5% CO_2/95% air. HeLa cells (ATCC CCL-2) were grown in minimal essential medium supplemented with 10% (v/v) FBS (Gibco) and 1% (v/v) penicillin/streptomycin (Gibco). Cells were passaged every 3–4 days and revived from liquid nitrogen after ~10 passages. All transfected cell lines were cloned and stored under liquid nitrogen.

Lentivirus library generation. The pooled human GeCKO v2 plasmid collection (sublibraries A + B) was obtained from Addgene (no. 1000000049), transferred into Lucigen Endura *Escherichia coli* cells (catalog. no. 60242) using electroporation (Electroporator 2510, Eppendorf), and amplified using maxi plasmid preparation (Zymo Research). Deep sequencing confirmed that the plasmid collection covered >93% of sgRNAs and genes in the GeCKO library. A total of 12 flasks (175 cm^2, Corning) with 1.75×10^6 HEK293T cells each were grown to 70% confluence. The medium was replaced with 10 ml of OptiMEM (Gibco) 1 h before transfection. For each flask, a total of 14 µg of GeCKO v.2 plasmid collection (7 µg per sublibrary), 7 µg of lentiviral envelope plasmid (pMD2.g; Addgene, no. 12259) and 11 µg of lentiviral packaging plasmid (psPAX2; Addgene, no. 12260) were diluted in 2.8 ml of OptiMEM and supplemented with 142 µl of Plus reagent (Invitrogen). In parallel, 71 µl of Lipofectamine LTX (Invitrogen) was diluted in 2.8 ml of OptiMEM and incubated at room temperature for 5 min, following which plasmids and Lipofectamine were added and the mixture was incubated for another 20 min before being added to the flasks. After 6 h, the transfection mix was replaced with DMEM (+10% FBS and 1% bovine serum albumin (BSA)). After 2 days, the virus-containing medium was collected and centrifuged for 10 min at 3,000*g* (4 °C) to sediment cell debris. Supernates were filtered through 0.45-µm protein-binding membranes (Millipore) and centrifuged for 2 h and 24,000*g* at 4 °C. Concentrated virus was resuspended in DMEM (+10% FBS and 1% BSA) at 4 °C overnight, then stored in aliquots at −80 °C. The MOI was determined via spinfection with various viral dilutions on A375 cells in 12-well plates (3×10^6 per well), followed by transfer of one half of each well into two wells for subsequent treatment with 1 µg ml^{-1} puromycin (Invitrogen). After 3 days, cells were counted and the MOI (μ) was computed from the fraction of surviving cells (q) using the formula: $\mu = -\log(1-q)$.

Library transduction. To maintain ~100-fold coverage of the genome, a total of 1×10^8 A375 cells were distributed across three 12-well plates (Corning) in DMEM/FBS medium containing 8 µg ml^{-1} polybrene (Sigma); the viral suspension was added to achieve a MOI of ~0.3. Plates were centrifuged at 2,000*g* for 2 h (37 °C) and incubated overnight. After transfer of cells into 175-cm^2 plates (Corning), they were incubated in medium containing puromycin (1 µg ml^{-1}) for 7 days. Aliquots of the cell library were stored in liquid nitrogen.

Screening assays. For the resistance screen, two replicates of 1×10^7 GeCKO-library-transduced A375 cells were seeded in 175-cm^2 plates (Corning) and treated with 10 nM guanidinoneomycin-saporin (GNeo-SAP) or 5 nM diphtheria toxin (List Labs) (Supplementary Information). Cells were maintained in GNeo-saporin for 4 days, after which the medium was replaced with toxin-free DMEM/FBS medium and cells were grown for a further 8 days. Two control plates were grown in parallel, treated with PBS (Gibco) and passaged every 2 days before harvesting after 8 days. For the FACS screen, two replicates of 1×10^7 GeCKO-library-transduced A375 cells were grown on 175-cm^2 plates for 2 days, harvested, washed in PBS, stained with GNeo-Cy5 and subjected to FACS (BD Influx sorter).

Sample preparation and sequencing. Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen), and CRISPR RNA barcodes were amplified through nested PCR. Genomic DNA extracts (2 µg) were diluted with double-distilled H_2O (ddH_2O) to a volume of 67 µl, and mixed with 33 µl of PCR1 reaction master mix (20 µl of Q5-reaction buffer (NEB), 2 µl of 10 mM deoxynucleotide triphosphate (dNTP; NEB), 5 µl of 10 µM primer F1, 5 µl of 10 µM primer R1 and 1 µl of Q5 (NEB)). PCR products were generated using the recommended conditions (NEB) at 70 °C annealing temperature for 18 cycles. After completion, all reactions were pooled and 5 µl of the PCR products was diluted with 62 µl of ddH_2O and mixed with 33 µl of PCR reaction 2 master mix (20 µl of Q5-reaction buffer, 2 µl of 10 mM dNTP, 5 µl of 10 µM primer F0x, 5 µl of 10 µM primer R0y and 1 µl of Q5). PCR products were generated under the recommended conditions using an annealing temperature of 70 °C for 24 cycles. PCR products were gel purified (QIAquick Gel Extraction kit, Qiagen) and submitted for sequencing on a HiSeq 4000 (Illumina).

Amplicon sequencing analysis. Analysis of the raw amplicon sequencing data, including quality control, sequence alignment, read counting, sgRNA enrichment analysis and gene ranking, was carried out using PinAPL-Py v.2.9 (ref. [11]). Briefly, a ranking score (Π_g) for each gene g was computed based on the read counts in the treatment condition $n_x(i)$ and in the control condition $n_0(i)$ of each sgRNA i in the GeCKO library, according to:

$$\Pi_g = q(I) \sum_{i \in I} \frac{\log_{10} n_x(i)}{\log_{10} n_0(i)}$$

where I is the set of sgRNAs i targeting gene g that yield significant enrichment $n_x(i)/n_0(i)$ according to the negative binomial model implemented in PinAPL-Py. q is a scaling function to grant additional weight to genes that receive support from multiple crRNAs. $q(I) = |I|$ was used in this study. Statistical significance of high Π_g scores was estimated by calculating Π_g for 10,000 permutations in which sgRNAs were assigned to genes randomly. In a restrictive setting, a P value threshold of 0.01 was used after Sidak correction for multiple tests to assess sgRNA enrichment and gene scores. In a permissive setting, a false discovery rate (FDR) of 0.1 was used instead, and a minimal read count of five counts per million reads was applied to decrease noise (Supplementary Datasets 1–7 provide a full list of parameters used). Additional analytical runs were carried out with different parameter settings, but KDM2B was consistently recovered among the top-ranked genes. Genes reaching significant enrichment in both screens were analyzed with STRING to find association clusters (www.string-db.org).

RNA-seq quantification and differential expression analysis. RNA-seq quality was assessed using FastQC. Adapter sequences and low-quality bases were trimmed using Trimmomatic[50]. Sequence alignment was performed using *STAR*[51] against the human genome (GRCH.38; GCF_000001405.38) with default parameters. The expression of each gene was quantified using HTSeq[52]. Differential gene expression analysis was accomplished using DESeq2. After Benjamini–Hochberg FDR correction, genes with adjusted $P < 0.05$ and fold change >1.5 were considered differentially expressed. Raw data were deposited at the Gene Expression Omnibus (GEO) and Short Read Archive (accession no. GSE145789). Functional annotation and gene set enrichment analysis of the top 200 differentially expressed genes were carried out using Metascape (http://metascape.org/).

ChIP–seq. A375 wild-type or *KDM2B*^{C5} cells (3×10^6) were fixed with 1% formaldehyde for 10 min, and the crosslinking reaction was quenched by the addition of 1.3 M glycine (1:10). Snap-frozen cell pellets were resuspended in 130 µl of RIPA lysis buffer (20 mM Tris/HCl pH 7.5, 1 mM EDTA, 0.5 mM EGTA, 0.1% SDS, 0.4% sodium deoxycholate, 1% NP-40 alternative, 0.5 mM DTT) supplemented with protease inhibitors. Sonicated lysates were centrifuged and incubated overnight at 4 °C with 2 µg of antibodies directed against H3K4me3 (no. 04-745, Millipore) or H3K36me2 (no. ab9049, Abcam) and 20 µl of protein A magnetic beads (Thermo Fisher). The beads were collected using a magnet and washed. DNA was amplified by PCR for 14 cycles in a 25-µl reaction volume using NEBNext Ultra II PCR master mix and 0.5 µM each of primers Solexa 1GA and Solexa 1GB. Sample concentrations were quantified by a Qubit dsDNA HS Assay Kit (Thermo Fisher) and 75-base pair (bp), single-end sequenced on a HiSeq 4000 (Illumina). ChIP–seq data were mapped to the hg38 genome using bowtie2 v.2.2.9 (ref. [53]), and tag directories were made using HOMER v.4.11 (ref. [54]). KDM2B ChIP–seq data were downloaded from GEO GSE108929. The binding sites of KDM2B were identified using HOMER 'findPeaks -style factor' and overlayed with H3K4me3 and H3K36me2 signal in A375 wild-type or *KDM2B*^{C5} cells using HOMER annotatePeaks.pl, with the parameters '-norm 1e7 -size 6000 -hist 25' to help compute the histograms of 25-bp bins within ±3,000 bp of KDM2B peaks. All sequencing data have been made available by deposition in the GEO database GSE163162. The University of California, Santa Cruz genome browser was used to visualize sequencing data. For KDM2B, three available antibodies were tested but none of them enriched KDM2B-bound targets. Since Banito et al. reported successful KDM2B ChIP, these data were downloaded from GEO GSE108929 (M5SS1) and are shown in Fig. 5c for *SULF1* along with a negative control gene, *B4GALT7* (ref. [22]).

ELISA experiments. The concentrations of MMP-9 and TIMP-3 in the supernatants after cell cultivation for 96 h were determined by sandwich enzyme-linked immunosorbent assay (ELISA) according to the manufacturer's protocols (R&D Systems). In brief, high-binding 96-well ELISA plates were coated overnight with capture antibodies for MMP-9 (1 µg ml^{-1}) or TIMP-3 (2 µg ml^{-1}) diluted in PBS. The plates were then incubated for 2 h after the addition of supernatants (100 µl of medium per well). MMP-9 and TIMP-3 were quantified using detection antibodies for MMP-9 (100 ng ml^{-1}; R&D Systems, no. DY911) and TIMP-3 (2 µg ml^{-1}; R&D Systems, no. DY973) diluted in 1% BSA in PBS, in comparison to calibration curves ranging 0–4 ng ml^{-1} MMP-9 or TIMP-3 in 1% BSA/PBS.

Immunoblotting. Nuclear proteins from A375 cell lines were extracted using an NE-PER nuclear and cytoplasmic protein extraction reagent kit (Thermo Fisher). Protein was quantified by BCA assay (Thermo Scientific-Pierce). Samples (11 µg of protein) were subjected to SDS–polyacrylamide gel electrophoresis (PAGE) (4–12% Bis-Tris, NuPAGE, Invitrogen), blotted on Immobilon-FL polyvinyl polypyrrolidone (PVDF) membranes (Millipore) and probed for KDM2B (rabbit

anti-JHDM1B, Millipore, no. 09-864, 1:1,000) and β-actin (mouse anti-Actin, Sigma, no. AC-74, 1:1,000). Membranes were blocked with 5% nonfat milk in Tris-buffered saline and 0.1% Tween 20 for 1 h at room temperature, then incubated overnight at 4 °C with respective primary antibodies. Mouse and rabbit primary antibodies were incubated with secondary Odyssey IR dye antibodies (LI-COR, 1:14,000) and visualized with an Odyssey IR imaging system (LI-COR Biosciences).

For detection of secreted SULF1 protein, conditioned medium from A375 wild-type and *KDM2B*^{C5} cells was collected after 3 days, centrifuged (1,000*g*, 5 min) and the supernatant was passed through a 0.2-µm filter. Filtered medium was transferred to a 30-kDa PES spin filter (Thermo Scientific-Pierce), concentrated by centrifugation (4,200*g*, 30 min) and exchanged three times by the addition of 50 mM Tris + buffer (50 mM Tris, 5 mM $MgCl_2$, 5 mM $CaCl_2$, 100 mM NaCl, 1× protease inhibitors, pH 7.5) to the concentrated sample. The resulting solution (~tenfold concentrated) was supplemented with protease inhibitors (Roche) and quantified for total protein content using BCA Assay (Thermo Scientific-Pierce). Aliquots of concentrated and unconcentrated conditioned medium were prepared with matching total protein levels (5 µg µl^{-1}), analyzed by SDS–PAGE (4–15% TGX Mini Protean Gel, Bio-Rad) and blotted onto PVDF membranes. Total protein content in the SDS–PAGE gel was evaluated by washing in InstantBlue Coomassie Protein Stain (Expedeon) overnight at 4 °C. SULF1 protein was detected with an anti-SULF1 primary antibody (rabbit polyclonal, Sigma, no. SAB1410410, 1:1,000) and a secondary anti-rabbit HRP antibody (Cell Signaling Technologies, no. 7074 S, 1:2,000). The membrane was visualized by the addition of Immobilon Crescendo Western HRP substrate (Millipore), followed by exposure in a Bio-Rad ChemiDoc XRS imaging system (Bio-Rad).

Cell line generation. Cas9-expressing A375 cells were generated by transfection with Fugene 6 (Promega), envelope plasmid (pMD2.g, Addgene, no. 12259), packaging plasmid (psPAX2, Addgene, no. 12260) and Cas9 expression plasmid (lentiCas9-Blast, Addgene, no. 52962) (4 µg each), followed by selection with 2 µg ml^{-1} blasticidin for 4 days. A375 mutant cell lines were generated by ligation of sgRNAs targeting *KDM2B* (5′-CTTGGTCAAGCGTCCGACTG-3′), *NDST1* (5′-ATGACGCACCTGTCCAACTA-3′) or *SLC35B2* (5′-TCCGCCTGAAGTACTGCACC-3′) into lentiGuide-Puro (Addgene, no. 52963) and cotransfection of each plasmid with Fugene 6 into HEK293T cells, along with viral plasmids pMD2.g (Addgene, no. 12259) and psPAX2 (Addgene, no. 12260), to make lentiviral particles, which were subsequently used to transduce Cas9-expressing A375 cells. After treatment of the transduced cell pool with puromycin (1 µg ml^{-1}) for 3 days, surviving cells were seeded onto a 96-well plate at low density and clonal populations were established. The A375 KDM2B SULF1 double-mutant cell line (*KDM2B*^{C5} *SULF1*$^{\#1}$) was generated by assembling a *SULF1*-targeting Cas9/crRNA/tracrRNA ribonucleoprotein particle in vitro (crRNA: 5′-GTGATTGTGCACATACTTCC-3′) using standard protocols (Integrated DNA Technologies) and transfecting it into *KDM2B*^{C5} cells by electroporation (Neon; Thermo Fisher). After 7 days, the cells were seeded onto a 96-well plate at low density and clonal populations were established. HeLa-KDM2B mutant cells were isolated by single-cell dilution from a HeLa-KDM2B mutant cell pool generated by Synthego, Inc.

Growth experiments and soft agar assays. For growth curves, A375 wild-type, *KDM2B*^{C5} or *KDM2B*^{C5}*SULF1*$^{\#1}$ cells were plated in 48-well plates (6,000 cells per well). Cell viability was measured daily with CellTiter-Blue (Promega). Clonogenic colony formation was assessed by seeding cells at 1,000 cells per well in six-well plates and incubation for 14 days under normal growth conditions. Colonies were visualized with methylene blue (12.5 mM in methanol) for 10 min at room temperature, washed with deionized water and imaged. For quantification, methylene blue was dissolved in 0.7 M sodium citrate in 50% ethanol and absorbance was measured at 650 nm.

For soft agar assays, 8,000 cells per well were plated in six-well plates in complete DMEM containing 0.2% agarose, with a 0.4% agarose underlay. The colonies were fed four drops of DMEM every 4 days. The colonies were stained with 0.01% crystal violet (w/v in 10% ethanol) and scored using the 'Analyze Particles' procedure in ImageJ software. For plotting, colony sizes (in pixels) were scaled by 10^6, log transformed and pooled from all biological replicates.

Plasmid preparations, rescue experiments and quantitative PCR. Lentiviral particles carrying the human *KDM2B* gene were produced by cotransfection of HEK293T cells with a plasmid encoding full-length *KDM2B* cDNA (Vector Builder, Inc.), a psPAX2 packaging plasmid (Addgene, plasmid no. 12260) and the VSV-G-encoding plasmid pMD2.g (Addgene, plasmid no. 12259). Medium containing the lentivirus particles was collected and used to infect HeLa-*KDM2B* mutant cells. After infection, the cells were cultured with 2 µg ml^{-1} puromycin to select for stably transduced cells.

For A375 rescue experiments, the PAM sequence in the pLV-KDM2B plasmid (VectorBuilder, Inc.) was mutated and a catalytic mutant construct was generated by H242A point mutation using the Q5 Mutagenesis kit (NEB) before transfection. The identity of the plasmid was confirmed by Sanger sequencing (Eton Bioscience, Inc.). *KDM2B*^{C5} cells (2×10^5 cells per well) were transfected with Lipofectamine LTX with Plus reagent (Invitrogen) and either the PAM-mutated pLV-KDM2B plasmid, the PAM-mutated H242A mutant KDM2B plasmid, a human HS3ST3A1 expression plasmid (VectorBuilder), a human HS6ST2 expression plasmid (Applied Biological Materials, Inc.), a mouse HS3ST1 expression plasmid or a siRNA-targeting human *SULF1* (SASI_Hs02_00330796, Sigma). Cells were incubated with this mixture for 4 h, after which the medium was replaced with DMEM + 10% FBS. FACS binding experiments were performed 48 h post transfection.

For quantitative PCR (qPCR) experiments, RNA was extracted from A375 cells, reverse transcribed and quantitated by qPCR with reverse transcription using primers for human *KDM2B*, *SULF1*, *HS6ST2*, *HS3ST3A1* and *YWHAZ*. qPCR was performed using SYBR Green Master Mix (Applied Biosystems) following the manufacturer's instructions. The expression of *YWHAZ* was used to normalize the expression of the target genes between samples. The primers used for qPCR are provided in Supplementary Table 3.

Protein biotinylation. A heparin-Sepharose column (100 µl, GE Healthcare) was pre-equilibrated with PBS (Gibco) and loaded with human FGF1 (ProSpec), FGF2 (Peprotech), FGF7 (ProSpec) or FGF10 (ProSpec) dissolved in PBS. The flowthrough was reloaded onto the column twice to ensure complete binding. After washing twice with PBS, a 0.6 mg ml^{-1} solution of Sulfo-NHS-LC-biotin (Thermo Fisher) in PBS was loaded onto the column and incubated for 1 h at room temperature. Each column was washed three times with PBS, then bound biotinylated protein was eluted with 0.4 ml of PBS buffer containing an additional 2 M NaCl. Biotinylated S100A12 was prepared as previously reported[21]. All biotinylated proteins were stored at −80 °C before use.

Flow cytometry. A375 and HeLa cells grown in monolayer culture were washed with PBS, lifted using Cell Dissociation Buffer (Gibco) or 10 mM EDTA and incubated in suspension for 30 min at 4 °C with 0.5 µg ml^{-1} mAb 10E4 (AMSBio, no. 370255-1, Clone F58-10E4, 1:1,000), 5 µg ml^{-1} mAb 3G10 (AMSBio, no. 370260-S, clone F69-3G10, 1:200), 80 nM biotin-FGF1, 2.5 nM biotin-FGF2, 180 nM biotin-FGF7, 35 nM biotin-FGF10 or 1 µg ml^{-1} biotin-S100A12. Alternatively, cells were incubated for 1 h at 4 °C with 500 nM human antithrombin (Aniara) or 10 nM human NRP1-His-biotin-streptavidin-Cy5 conjugate. Bound antithrombin was detected with 2 µg ml^{-1} anti-AT pAb (R&D Systems, no. AF1267, 1:100) followed by 2.5 µg ml^{-1} donkey anti-goat conjugated to AlexaFluor488 (Invitrogen, no. A-11055, 1:1,000). Binding of biotinylated proteins was detected by 1:1,000 dilution of streptavidin-Cy5 (Molecular Probes). For heparin lyase pretreatment, adherent cells were incubated with 5 mU ml^{-1} each of heparin lyases I, II and III for 30 min at 37 °C in growth medium. Flow cytometry was performed on a FACSCalibur (BD Biosciences) instrument (≥10,000 events per sample) and raw data were analyzed using FlowJo Analytical Software (Tree Star, Inc.). Cells were gated according to forward and side scattering. The extent of protein binding was quantified using the geometric mean of the fluorescence intensity. These values were plotted and further analyzed using GraphPad Prism v.8.0.

HS purification and liquid chromatography–MS analysis. A375 wild-type or *KDM2B*^{C5} cells (0.5×10^6) were seeded in a 10-cm-diameter plates and harvested when confluent. Cell surface HS was isolated and disaccharide composition was analyzed as described previously[55].

Radiolabeling and HS structural analysis. A375-Cas9 and *KDM2B* mutant cells (1×10^6) were radiolabeled with [^{35}S]sulfate (50 microcuries ml^{-1}) in F12 medium containing 10% dialyzed FBS for 24 h at 37 °C. HS was purified using the methods described above and analyzed by size-exclusion chromatography (Sepharose 6B-L column 1 × 70 cm; 50 mM sodium acetate, 0.2 M NaCl, pH 6.0). For filter binding assays, recombinant FGF1 (Prospec), antithrombin (Anaira) or BSA (Sigma) was incubated with 20,000 counts of [^{35}S]HS for 1 h at room temperature. Binding was measured by filtration on prewashed nitrocellulose membranes as previously described[56].

A GNeo affinity column was prepared by dissolving 1 mg of GNeo-biotin in 50 mM sodium acetate buffer containing 0.1 M NaCl, pH 6.0. The solution was added to a column (0.5 ml) of streptavidin-Sepharose (GE Healthcare) and incubated for 20 min at room temperature. Next, the column was washed thoroughly to remove unbound GNeo-biotin. Subsequently, [^{35}S]HS was applied to the column and eluted with a stepwise salt gradient from 0.1 to 1.5 M NaCl.

Statistics and reproducibility. Statistical tests and sample sizes are indicated in the figure legends. $^{****}P < 0.0001$, $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$. All tests were two-sided. Tests were performed in either Prism (GraphPad) or R (R-Project). Measurements were taken from distinct samples, and the number of biological replicates is indicated in the figure legends. Error bars represent mean ± standard deviation. Boxplot center refers to median, boxplot bounds refer to 25 and 75% percentiles and whiskers refer to minimum and maximum, unless noted otherwise. Immunoblots were performed twice independently, and representative images are shown in the figures.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Any data generated or analyzed during this study, associated protocols, materials within the manuscript and public databases (GSE163162, GSE145789) are included in the article and related Supplementary information, or are available from the corresponding author. Source data are provided with this paper.

References

50. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
51. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
52. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
53. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
54. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
55. Lawrence, R. et al. Evolutionary differences in glycosaminoglycan fine structure detected by quantitative glycan reductive isotope labeling. *J. Biol. Chem.* **283**, 33674–33684 (2008).
56. Kreuger, J., Lindahl, U. & Jemth, P. Nitrocellulose filter binding to assess binding of glycosaminoglycans to proteins. *Methods Enzymol.* **363**, 327–339 (2003).

Acknowledgements

We thank C. van der Kooi, University of Kentucky, for providing the b1b2 domain of NRP1, and D. Xu, University of Buffalo, for providing the biotinylated S100A12 protein. We also thank the GlycoAnalytics Core Facility at University of California, San Diego for help with analytical experiments. We thank C. Kuo for processing of RNA sequencing data. RNA sequencing and CRISPR amplicon sequencing were conducted at the IGM Genomics Center, University of California, San Diego, La Jolla, CA (MCC grant no. P30CA023100). This work was supported by grant nos. R21 CA199292 (to J.D.E. and N.E.L.), GM33063 (to J.D.E.), GM119850 (to N.E.L.), NSF CHE 200424 (to K.G.) and T32 GM008326 (fellowship support for B.M.T.); and DFG research fellowship no. 420160411 from the German Research Foundation (to S.R.) and no. K12HL141956 (fellowship support for R.J.W.).

Author contributions

R.J.W., P.N.S., P.L.S.M.G., N.E.L. and J.D.E. designed the research. Unless otherwise noted, R.J.W. and P.N.S. performed the experimental work and analyzed the data. A.W.T.C. processed and analyzed the RNA-seq data. Q.L. and J.L. characterized the KDM2B mutant clones and performed immunoblotting and qPCR experiments. K.M.H. synthesized the GNeo-biotin derivative. S.R. performed ELISA experiments. T.M.C. performed soft agar assays. M.A.H. processed and analyzed ChIP–seq experiments. B.M.T. performed immunoblotting experiments. K.G., C.K.G. and Y.T. contributed new reagents. R.J.W., P.N.S., N.E.L. and J.D.E. wrote the paper.

Competing interests

The University of California San Diego and J.D.E. have a financial interest in TEGA Therapeutics, Inc. The terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict-of-interest policies.

Additional information

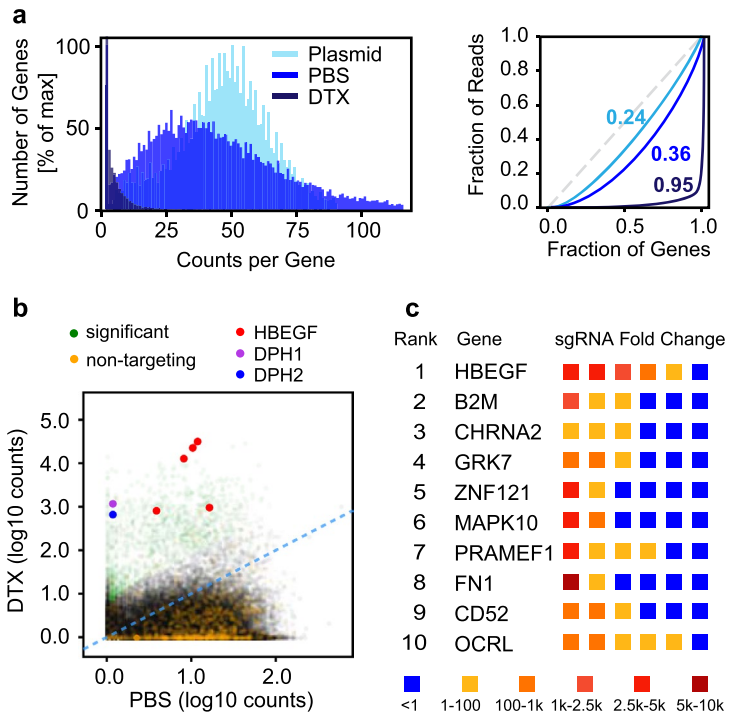
Extended data is available for this paper at https://doi.org/10.1038/s41589-021-00776-9.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-021-00776-9.

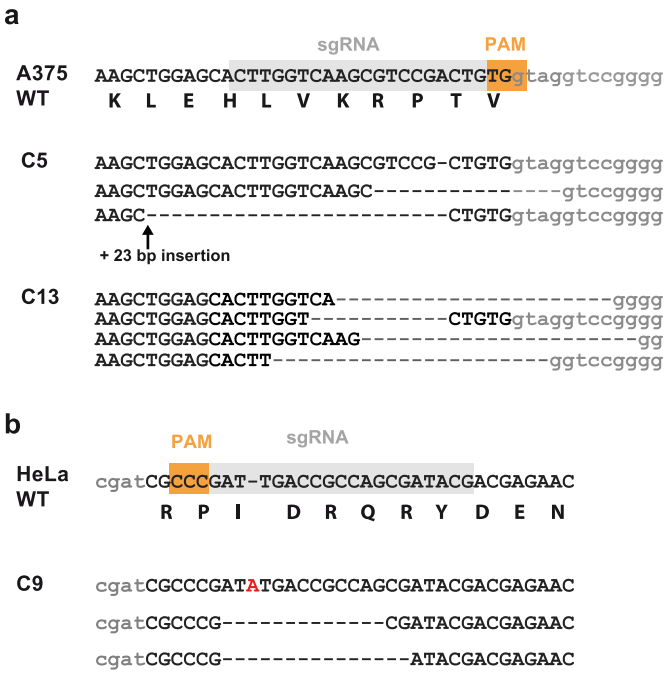
Correspondence and requests for materials should be addressed to N.E.L. or J.D.E.

Peer review information *Nature Chemical Biology* thanks Ulf Lindahl, Linda Troeberg and the other, anonymous, reviewer for their contribution to the peer review of this work.

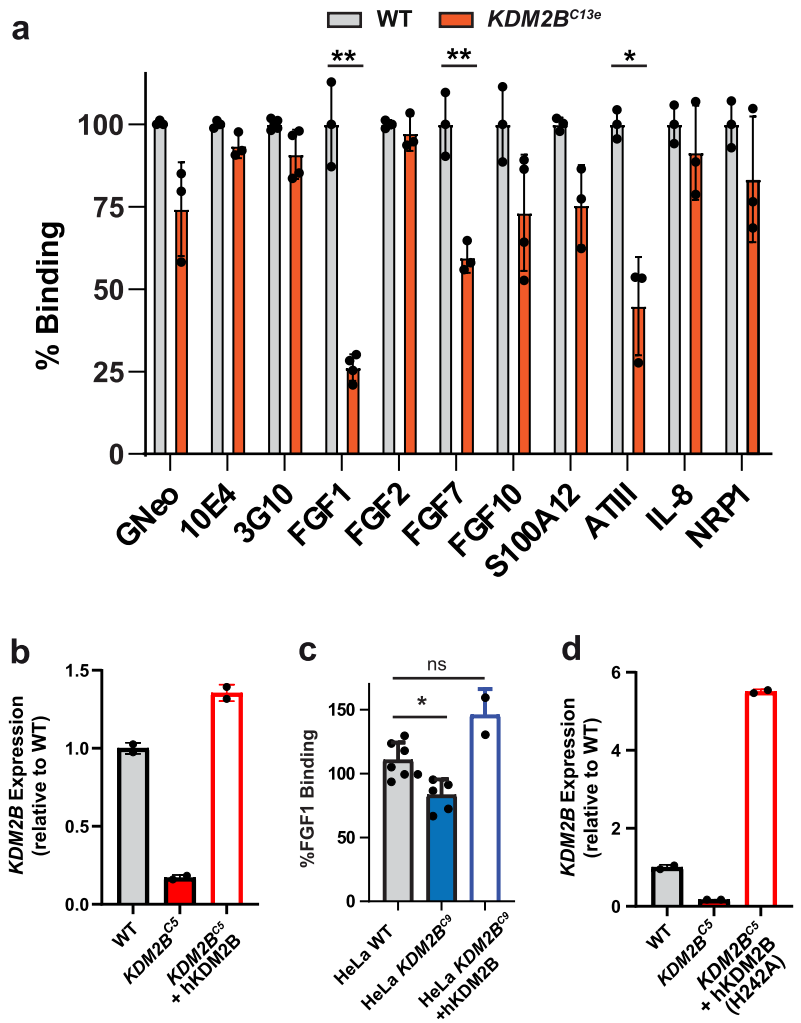
Reprints and permissions information is available at www.nature.com/reprints.



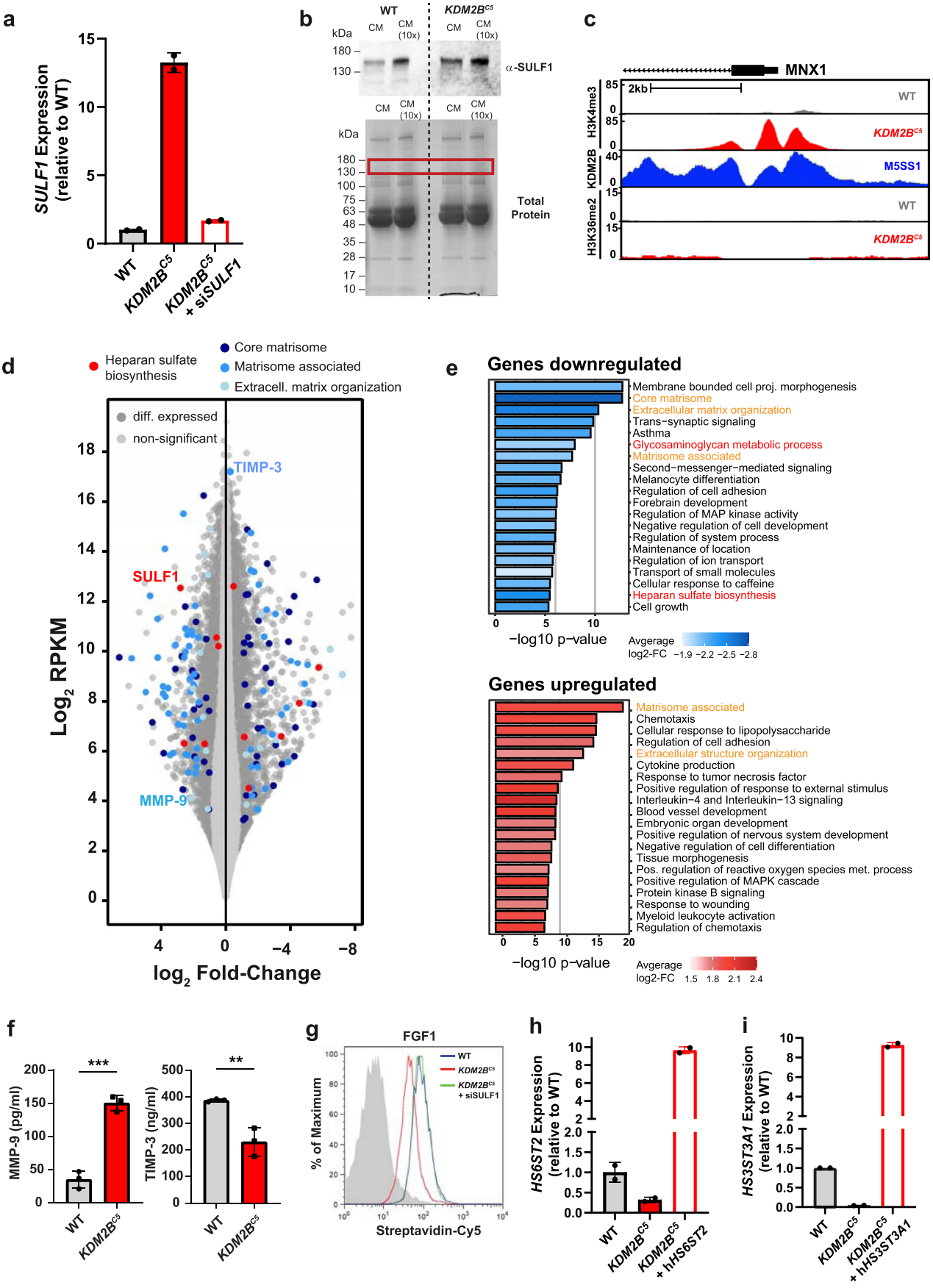
Extended Data Fig. 1 | A genome-wide screen for resistance to diphtheria toxin. **a**, *Left:* Frequency distribution of gene counts after treatment with either diphtheria toxin (DTX) or PBS (Plasmid = GeCKO plasmid library). *Right*: Lorenz curves showing the distribution of sequencing reads over the gene library. Numbers represent Gini coefficients. **b**, Scatterplot of sgRNA counts (log10, normalized) in samples after DTX treatment versus PBS treatment. The sgRNA fraction with significant fold-change is shown in green. sgRNA fraction representing non-targeting controls shown in orange. HBEGF-targeting sgRNAs shown in red, and DPH genes are shown in purple and blue. **c**, Table of the top 10 ranked genes showing enrichment after DTX treatment. The enrichment level of the six gene-targeting sgRNAs is indicated by color.



Extended Data Fig. 2 | *KDM2B*-deficient cell lines. a, Sanger sequencing of two A375 KDM2B knockout clones (C5 and C13) and (**b**) Sanger sequencing of one HeLa knockout clone (C9) after targeting with the indicated sgRNAs. Intron sequence denoted in lower case.

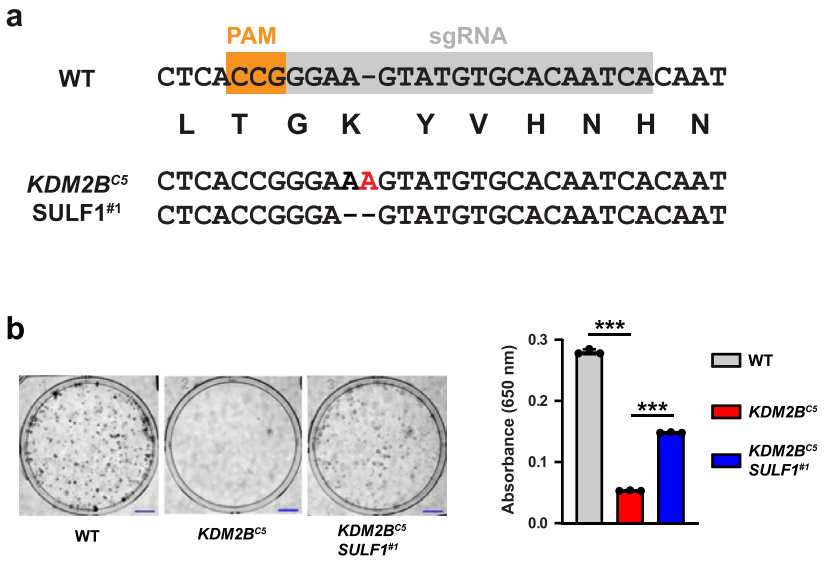


Extended Data Fig. 3 | KDM2B affects protein binding. a, Protein binding in A375 KDM2B knockout clone C13 (t-tests; n = 3). **b**, Fold change in *KDM2B* mRNA expression in A375 *KDM2B*^{C5} cells, and in A375 *KDM2B*^{C5} cells expressing a *KDM2B* cDNA from a lentiviral construct (hKDM2B) (n = 2). **c**, FGF1 binding in HeLa *KDM2B*^{C9} cells, and in HeLa *KDM2B*^{C9} cells expressing *KDM2B* cDNA (hKDM2B) (t-test on $\log_{10}$ fluorescence data, n = 5). **d**, Fold change in *KDM2B* mRNA expression in A375 *KDM2B*^{C5} cells expressing a *KDM2B* cDNA with a point mutation (H242A) in the demethlase domain (n = 2).



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Transcriptomic changes in *KDM2B*-deficient cells. a, Fold change in *SULF1* mRNA expression in A375 *KDM2B*^{C5} cells, and in A375 *KDM2B*^{C5} cells expressing a siRNA against *SULF1* (n = 2). **b**, top: Western blot of SULF1 protein levels in conditioned media (CM) collected from A375 wild-type and *KDM2B*^{C5} cells (CM 10x equals 10-fold concentrated solution) bottom: Total protein (coomassie-stained gel), red frame indicates position of SULF1. **c**, Gene track for H3K4me3, KDM2B, and H3K36me2 ChIP-Seq at the *MNX1* locus, a known target of KDM2B. **d**, Transcriptome-wide expression data displaying differentially expressed genes in *KDM2B*^{C5} cells compared to wild-type. HS biosynthetic genes and extracellular matrix genes are highlighted. **e**, Gene set enrichment analysis of significantly downregulated ($\log_2 \leq -0.5$, $p \leq 0.05$; FDR-corrected) and upregulated ($\log_2 \geq 0.5$, $p \leq 0.05$; FDR-corrected) genes in A375 wild-type and *KDM2B*^{C5} RNA-Seq datasets (n = 3). **f**, MMP-9 and TIMP-3 levels in the supernatant from cultured A375 wild-type and *KDM2B*^{C5} cells (t-test, n = 3). **g**, Histograms showing FGF1 binding in A375 wild-type, *KDM2B*^{C5} cells, and in *KDM2B*^{C5} cells upon treatment with an siRNA targeting *SULF1*. **h**, Fold change in *HS6ST2* mRNA expression in A375 *KDM2B*^{C5} cells, and in A375 *KDM2B*^{C5} cells expressing a *HS6ST2* cDNA from a lentiviral construct (hHS6ST2) (n = 2). **i**, Fold change in *HS3ST3A1* mRNA expression in A375 *KDM2B*^{C5} cells, and in A375 *KDM2B*^{C5} cells expressing a *HS3ST3A1* cDNA from a lentiviral construct (hHS3ST3A1) (n = 2).



Extended Data Fig. 5 | KDM2B affects cell growth through SULF1. a, Sanger sequencing of a A375 *KDM2B SULF1* double knock-out clone (KDM2B^{C5} *SULF1*$^{\#1}$) after targeting with the indicated sgRNA. Both sequence changes result in frameshift mutations. **b**, Clonogenic assay under normal growth conditions. After 14 days, colony growth was quantified by methylene blue staining and absorption readings at 650 nm (t-test, n = 3). Scale bar = 5 mm.

nature research

Corresponding author(s): Jeffrey D. Esko

Last updated by author(s): Jan 27, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All instrument-generated data were collected using the manufacturer's associated software as described in the Online Methods. Plate reader data were collected by Victor Nivo (Ver. 3.0, Perkin Elmer). Cell images were collected by Image Lab Touch software (Ver. 6.1, BioRad). Gel images were collected by ImageStudio (Ver. 5.2.5, LI-COR Biosciences). Flow cytometry data was collected on BD FACStation™ software (Ver. 6.1x , BD Biosciences).
Data analysis	GraphPad Prism v8, R (R-project) v3.5.1, PinAPL-Py v2.9, Xcalibur (Thermo Scientific, v3.0.63), FastQC v0.11.8, Trimmomatic v0.32, STAR v2.7.0, HTSeq v0.14, DESeq2 v1.20, FlowJo v8.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Any data generated or analyzed during this study, associated protocols, materials within the manuscript, and public databases (GSE163162, GSE145789) are included in the article and related supplementary information or are available from the corresponding author.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical analyses were used to predetermine sample size. Sample sizes for all experiments were chosen based on previous publications as cited in the text to account for intrinsic variability of the experiments.
Data exclusions	None
Replication	All experiments were replicated as indicated in the text and supporting information.
Randomization	No randomization was applied in these experiments. Specific controls for covariates in the cell-based assays were not relevant as identical protocols and reagents were used within each assay across all samples except for the variable being tested.
Blinding	No blinding was applied in these experiments. All data acquisition and analyses were performed automatically where possible by software described above and in the text and supporting information.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Anti-heparan sulfate monoclonal IgM antibody (AMSBio #370255-1, Clone F58-10E4); Anti-human D-heparan sulfate monoclonal IgG antibody (AMSBio #370260-S, clone F69-3G10); Anti-JHDM1B (KDM2B) rabbit polyclonal antibody (MilliporeSigma #09-864, Lot #3173542); Anti-beta-Actin mouse monoclonal antibody (Sigma #A2228-200UL, Clone AC-74); Anti-Antithrombin III goat polyclonal antibody (anti-AT pAb, R&D Systems #AF1267); Donkey Anti-Mouse IgG IRDye 800CW secondary antibody (LICOR #926-32212), Donkey Anti-Rabbit IgG IRDye 680LT secondary antibody (LICOR #926-68023); Donkey Anti-Goat IgG Alexa Fluor 488 secondary antibody (Invitrogen #A-11055); Goat Anti-mouse IgM Alexa Fluor 488 secondary antibody (Invitrogen #A21042); 6X-His Tag mouse monoclonal antibody (HIS.H8) biotin (Invitrogen #MA1-21315-BTIN), rabbit polyclonal anti-SULF1 (SAB1410410, Sigma), goat anti-rabbit HRP antibody (Cell Signaling Technologies, 7074S), rabbit monoclonal anti-H3K4me3 antibody (04-745, Millipore), rabbit polyclonal anti-H3K36me2 (ab9049, Abcam).
Validation	All antibodies were commercially avaialbe and used for applications previously validated by the manufacterers and in relevant citations on the manufacter's websites or in house. Anti-heparan sulfate monoclonal IgM antibody (AMSBio #370255-1, Clone F58-10E4): Manufacturer validated for Western blotting using liposome-incorporated membrane heparan sulfate proteoglycan (F58) from human fetal lung fibroblasts Anti-human D-heparan sulfate monoclonal IgG antibody (AMSBio #370260-S, clone F69-3G10): Manufacturer validated for Western blotting using heparitinase-digested medium heparan sulfate proteoglycans (F69) from human lung fibroblasts Anti-JHDM1B (KDM2B) rabbit polyclonal antibody (MilliporeSigma #09-864, Lot #3173542): Manufacturer validated for Western blotting using KLH-conjugated linear peptide corresponding to human JHDM1B Anti-beta-Actin mouse monoclonal antibody (Sigma #A2228-200UL, Clone AC-74): Manufacturer validated for Western blotting using Beta-actin N-terminal peptide. Anti-Antithrombin III goat polyclonal antibody (anti-AT pAb, R&D Systems #AF1267): Manufacturer validated for Western blotting

using mouse myeloma cell line NS0-derived recombinant human Serpin C1/Antithrombin-III (His33-Lys464).
Donkey Anti-Mouse IgG IRDye 800CW secondary antibody (LICOR #926-32212): Manufacturer validated for Western secondary detection using Ms IgG.
Donkey Anti-Rabbit IgG IRDye 680LT secondary antibody (LICOR #926-68023): Manufacturer validated for Western secondary detection using Rb IgG.
Donkey Anti-Goat IgG Alexa Fluor 488 secondary antibody (Invitrogen #A-11055): Manufacturer validated for Western secondary detection using Gt IgG.
Goat Anti-mouse IgM Alexa Fluor 488 secondary antibody (Invitrogen #A21042): Manufacturer validated for Western secondary detection using Ms IgM.
6X-His Tag mouse monoclonal antibody (HIS.H8) biotin (Invitrogen #MA1-21315-BTIN): Manufacturer validated for Western blotting using 6x His synthetic peptide.
Rabbit polyclonal anti-SULF1 (SAB1410410, Sigma): Manufacturer validated for Western blotting using SULF1 full-length human protein
Goat anti-rabbit HRP antibody (Cell Signaling Technologies, 7074S): Manufacturer validated for Western secondary detection using Rb IgG.
Rabbit monoclonal anti-H3K4me3 antibody (04-745, Millipore): Manufacturer validated for Western blotting using a peptide containing the sequence [RTtrimKQ] in which lysine 4 is trimethylated on human histone H3.
Rabbit polyclonal anti-H3K36me2 (ab9049, Abcam): Manufacturer validated for Western blotting using a synthetic peptide within Human Histone H3 aa 1-100 (di methyl K36) conjugated to keyhole limpet haemocyanin.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T (ATCC® CRL-3216™), A375 (ATCC® CRL-1619™), HeLa (ATCC® CCL-2™)
Authentication	Cells were used as received from ATCC without further authentication.
Mycoplasma contamination	All cell lines tested negative for Mycoplasma.
Commonly misidentified lines (See ICLAC register)	The study did not use any commonly misidentified cell lines.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	GSE145789 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE145789) GSE163162 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE163162)
Files in database submission	ChIP_A375_WT_H3K4me3_rep1 ChIP_A375_WT_H3K4me3_rep2 ChIP_A375_WT_H3K36me2 ChIP_A375_WT_Input ChIP_A375_KDM2BKO_H3K4me3_rep1 ChIP_A375_KDM2BKO_H3K4me3_rep2 ChIP_A375_KDM2BKO_H3K36me2 ChIP_A375_KDM2BKO_Input
Genome browser session (e.g. UCSC)	https://genome.ucsc.edu/s/mahoeksema/KDM2B_Ryan

Methodology

Replicates	Duplicates were used for ChIP-seq experiments
Sequencing depth	file total reads uniquely mapped reads ChIP_A375_WT_H3K4me3_rep1 13002463 11163786 75bp SR ChIP_A375_WT_H3K4me3_rep2 17967162 15111652 75bp SR ChIP_A375_WT_H3K36me2 18417053 14511235 75bp SR ChIP_A375_WT_Input 11362348 8656329 75bp SR ChIP_A375_KDM2BKO_H3K4me3_rep1 14995184 12872334 75bp SR ChIP_A375_KDM2BKO_H3K4me3_rep2 15841688 13706284 75bp SR ChIP_A375_KDM2BKO_H3K36me2 18684404 14570684 75bp SR ChIP_A375_KDM2BKO_Input 7451556 5720458 75bp SR
Antibodies	H3K4me3 (Millipore, 04-745, lot 3135978) H3K36me2 (Abcam, ab9049, lot GR3258133-1)

Peak calling parameters	Normalized tag counts were computated using HOMER annotatePeaks.pl script with parameter "-norm 1e7". Histone marks were counted within +/- 3000 bp around KDM2B peaks
Data quality	Regions were called based on following parameters (poisson p-value: 1.00e-04 and 4 fold enrichment): file peaks ChIP_A375_WT_H3K4me3_rep1 24682 ChIP_A375_WT_H3K4me3_rep2 27658 ChIP_A375_WT_H3K36me2 87395 ChIP_A375_KDM2BKO_H3K4me3_rep1 22920 ChIP_A375_KDM2BKO_H3K4me3_rep2 21393 ChIP_A375_KDM2BKO_H3K36me2 36481
Software	ChIP-seq data was mapped to the hg38 genome using bowtie2 v2.2.92 and tag directories were made using HOMER v4.11.3. KDM2B ChIP-seq data was downloaded from GEO: GSE1089294. The binding sites of KDM2B were identified using HOMER 'findPeaks -style factor' and overlayed with H3K4me3 and H3K36me2 signal in A375 wild-type or KDM2BC5 cells using HOMER annotatePeaks.pl with parameters "-norm 1e7 -size 6000 -hist 25" to help compute the histograms of 25-bp bins within +/- 3000 bp of KDM2B peaks.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cell lines (obtained from ATCC) were grown in monolayer culture, were washed with PBS, lifted using Cell Dissociation Buffer (Gibco) or 10 mM EDTA, and incubated with relevant antibodies or protein ligands as described in the Materials and Methods section before flow cytometry analysis.
Instrument	BD FACSCalibur (#343202) was utilized for analysis. BDInflux sorter (#646500) was utilized for cell sorting.
Software	BD FACStation™ v6.1x for Mac OS® X was used to collect flow cytometry data. FlowJo Analytical Software (v8.0, Tree Star Inc.) was used to analyze the data.
Cell population abundance	10,000 cells/events were analyzed per sample. Forward and side scatter gating were used to check purity.
Gating strategy	Cells were gated according to forward and side scattering.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.