PNAS
www.pnas.org

Supplementary Information for
THAP1 Modulates Oligodendrocyte Maturation by Regulating ECM Degradation in Lysosomes

Dhananjay Yellajoshyula[1*], Samuel S. Pappas[4,5], Abigail Rogers[1], Biswa Choudhury[6], Xylena Reed[2], Jinhui Ding[2], Mark R. Cookson[2], Vikram Shakkottai[5], Roman Giger[3] and William T. Dauer[4,5,7*]

Corresponding authors: William Dauer, Dhananjay Yellajoshyula
Email: william.Dauer@UTSouthwestern.edu; dyellaj@med.umich.edu

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Datasets S1 to S3 (excel files).

Supplementary Materials & Methods

Generation of Mice

The breeding strategy used to generate derive all primary OPC cells and conditional null mice was as follows: *Thap1*$^{+/-}$*, Olig2-Cre*$^{+}$*,Tg*^{GUS} × *Thap1*$^{flox/flox}$*.* This breeding strategy produced the following offspring ± Tg^{GUS}: *Thap1*$^{flox/+}$; *Thap1*$^{flox/-}$; *Thap1*$^{flox/+}$*, Cre*$^{+}$; and *Thap1*$^{flx/-}$*, Cre*$^{+}$ (CKO). Breeding strategy for generating O-CKO expressing MOBP-eGFP was as follows: *Thap1*$^{+/-}$*; Olig2-Cre*$^{+}$*;Tg*^{GUS} × *Thap1*$^{flox/flox}$*;Tg*$^{MOBP\text{-}eGFP}$*.* This breeding strategy produced- control (*Thap1*$^{+/flx}$), O-CKO (*Thap1*$^{flx/-}$*; Olig2-Cre*$^{+}$*;Tg*$^{MOBP\text{-}eGFP}$) and O-CKO-GUS (*Thap1*$^{flx/-}$*; Olig2-Cre*$^{+}$*;Tg*$^{MOBP\text{-}eGFP}$*;Tg*^{GUS}). Age and sex-matched littermate mice were used for all experiments. Primers used for genotyping in this study (*Thap1*, *Cre* & *Tg*^{GUS}) are listed in Table S1.

RNA extraction, qRT-PCR and RNAseq

Total RNA extraction from OL cultures for qRT-PCR and RNAseq analysis was done using NucleoSpin® RNA (Takara) and Trizol (Thermofisher) for mouse cerebral cortex as per manufacturers instructions. cDNA synthesis from total RNA was done using MMLV Reverse Transcriptase (Takara) as per manufacturers instructions. Quantitative real time PCR (qRT–PCR) was performed with the StepOnePlus System (ABI) and 2x SYBR Power Mix (ABI). RNAseq was performed on Illumina HISeq2500 (Illumina) platform at the neurogenetics laboratory, NIA. RNAseq libraries were made using the TruSeq Stranded Total RNA Library prep kit with rRNA removal mix Gold (Illumina, 20020598). Libraries were quantified and normalized using the KAPA Library Quantification kit for Illumina Platforms (KAPA Biosystems, KR0405) and sequenced on an Illumina HiSeq 2500 using HiSeq SBS kit v4 250 cycle kit (Illumina, FC-401-4003). A standard Illumina pipeline was used to generate fastq files. RNAseq Analyses was performed using DESEQ.

Immunostaining, Imaging & Analysis

OL Culture: OPC or OL cells were permeabilized with PBS +1% BSA + 0.33% Saponin, followed by overnight incubation in the same solution. **Tissue:** All CNS tissue sections were stained as free floating sections (40µm^2) in PBS + 1% BSA + 0.1% Trixton-X-100 as previously described (1). Antibodies used for this study are listed in SI Index Table S2. **Imaging:** Fluorescent images were collected with a 10x or 20x objective using Zeiss Axio-Observer microscope with Apotome.2. Low magnification images for the entire coronal section was acquired using a 4x objective on a Zeiss Axioskop 2 plus. Image processing, intensity and cell number measurements and image stitching was done using image J.

Electron Microscopy

EM was performed as previously described (1). Briefly, P21 mice were anesthetized and perfused with EM perfusion solution (3% paraformaldehyde, 2.5% glutaraldehyde in 0.1M phosphate buffer). Following perfusion, brains were dissected and postfixed at 4°C overnight in perfusion solution. Genu of CC was dissected, processed and sectioned at EM core facility, Emory University. All EM images were acquired using JEOL JSM 1400 at the University of Michigan, MIL core services. Axon caliber and g-

ratio (ratio of the inner axonal diameter to total (including myelin) outer diameter) were calculated as previously described (1, 2).

Derivation of OPC from NSC cells

OPC used in this study were derived from primary NSC isolated from the sub ventricular zone (SVZ) of P7 mouse pups as previously described (1). Routine propagation of the OPC was done as a monolayer on poly-L-ornithine (0.1mg/ml) and laminin (5µg/ml) coated dishes in OPC expansion media (SATO medium supplemented with 20µg/ml PDGF and FGF2; (1)) or differentiated on poly-L-ornithine (0.1mg/ml) and poly-D-lysine (0.1mg/ml) coated coverslips or dishes in oligodendrocyte differentiation medium (SATO media with 40µg/ml T3) (1).

β-glucuronidase activity assay - MUG assay

Tissue or cell pellets were lysed in MUG lysis buffer (0.2M Sodium acetate buffer, pH 4.5;10 mM EDTA; 0.1% Trixton-X-100) to obtain total cell homogenate. Equal concentrations of serially diluted lysate (in the range of 0- 60 µg/ml) from control and *Thap1* cKO genotypes was incubated with 10 mM 4-methylumbelliferyl β-D-glucuronide (MUG, Sigma) in 0.2M Sodium acetate buffer, pH 4.5 at 37°C for one hour. The reaction was stopped with equivalent volume of 0.2M sodium citrate followed by which the fluorescence (excitation = 360nm; emission = 460nm) was measured using Biotek Synergy HT microplate reader. Based on fluorescence reading from purified β-glucuronidase standard (Sigma), we determined the β-glucuronidase activity / hour and β-glucuronidase activity / mg of the lysate.

Quantitative Chromatin immunoprecipitation (qCHIP)

qCHIP was performed as previously described (1, 3). Sheared chromatin (sonicated to 200–500 bp) from 2×10^6 mouse OL cells was incubated with 2.5µg of the antibody pre-incubated with Dynabeads (Invitrogen). After washing, elution and cross-link reversal, DNA from each ChIP sample and the corresponding input sample was purified and analyzed further using qPCR. Each ChIP sample and a range of dilutions of the corresponding input sample (0.01 – 2% input) were quantitatively analyzed with gene-specific primers using the StepOnePlus System (ABI) and SYBR qPCR Powermix (ABI). Antibodies used for qCHIP are listed in table S2. Primers used for qChIP analysis are listed in table S1.

Gene Ontology enrichment analysis

Candidate THAP1-dependent DEGs identified from RNAseq catalysis were analyzed for enrichment of biological pathways by gene ontology (GO) enrichment analyses using the following web based applications: GENEONTLOGY http://geneontology.org/docs/go-enrichment-analysis/ and DAVID https://david.ncifcrf.gov/home.jsp. Top GO enriched ($p<0.01$ and > 2 fold enrichment) for GO-Biological processes and GO-Molecular function was used for this study.

GRIL LC/MS analyses

GAG Extraction: A small fraction of homogenized cell pellet or media supernatant was used for estimating protein concentration (BCA assay, this value was used to normalize

the amount of GAG present in the samples) and the remaining material was digested with Pronase (Sigma-Aldrich) in phosphate-buffered saline at 37°C for 24h. Pronase digested samples were further filtered using 45micron syringe filter followed by anion-exchange chromatography using freshly packed DEAE-Sephacel column. Negatively charged GAG samples will bind to the column was washed with 10mL of DEAE-wash solvent (50mM NaOAc, 150mM NaCl, pH-6.0). Total GAG was eluted with GAG-elution buffer (50mM NaOAc, 2M NaCl, pH-6.0). Finally the samples were desalted using PD10 desalting column and lyophilized (4). **GRIL LC/MS:** The total GAGs thus extracted was digested with Chondroitinase ABC, Heparin lyases I, II, and III or Hyaluronidase to enzymatically depolymerize CS, HS and HA chains respectively as previously described (4). Corresponding disaccharides were reductively coupled with the respective disaccharides, spiked with known amount of $^{13}C_6$-Aniline tagged with disaccharide standard mixture prior to running on LC-MS. LC/MS analysis was done using C18 reversed-phase column using a mixture of di-butyl amine and 70% methanol as eluent. The mass spec was run at negative ionization mode (40). The final values of CS, HS or HA GAGs thus obtained was normalized to the total protein concentration.

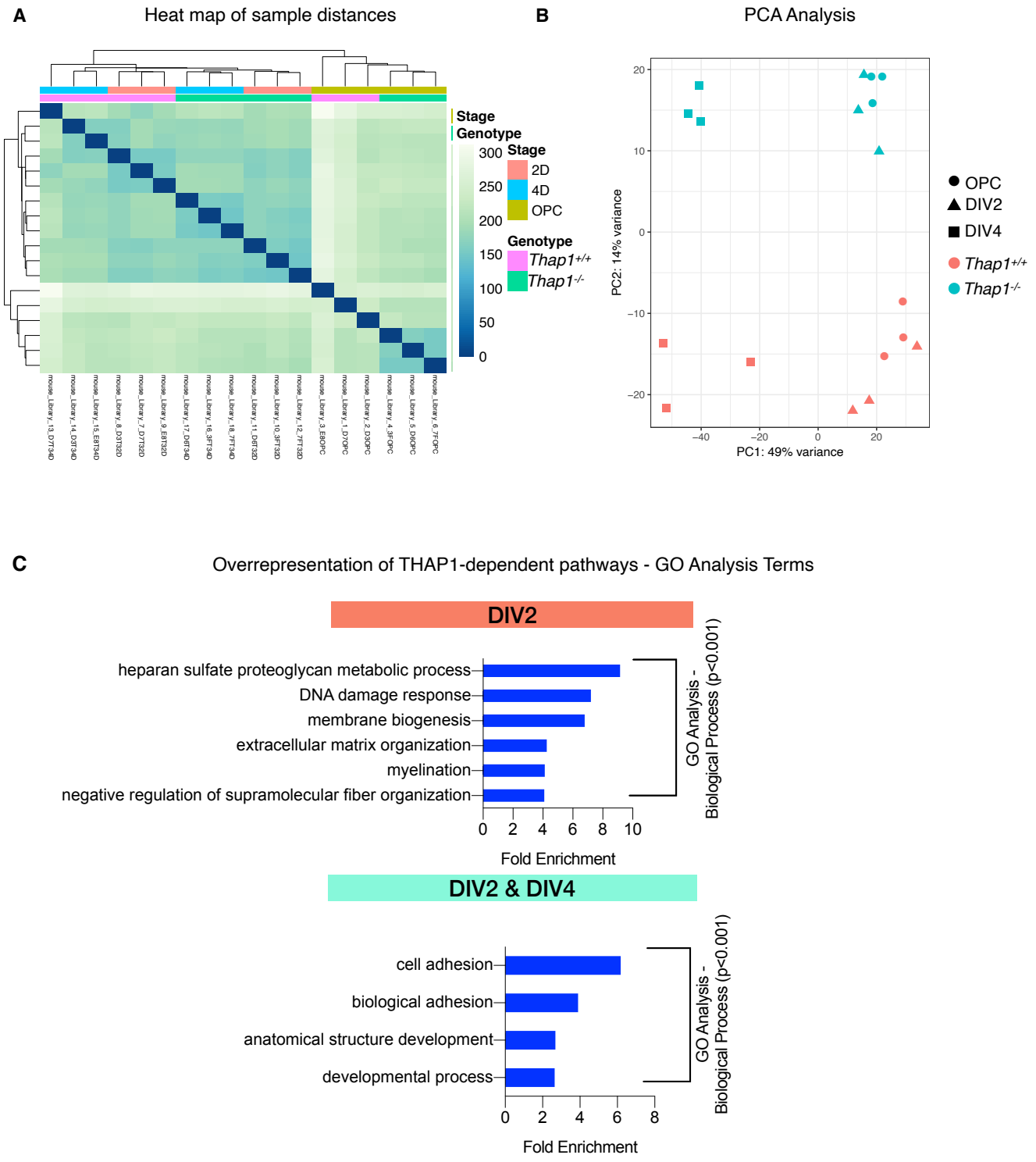
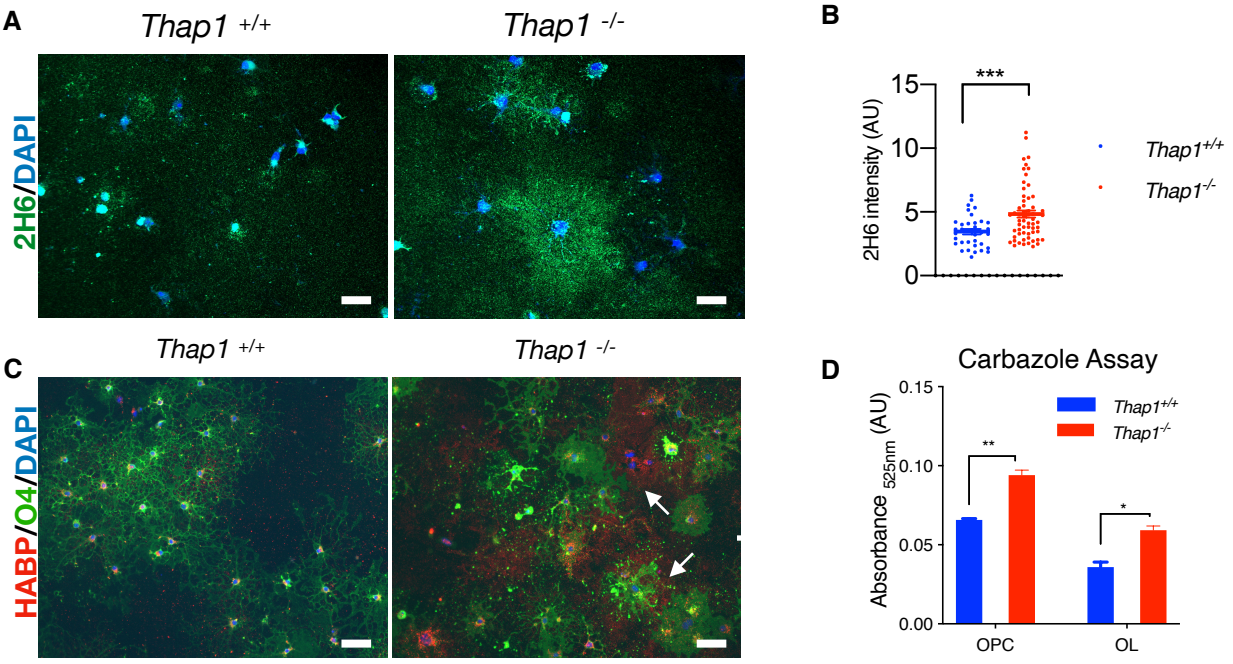


Figure S1. THAP1 loss disrupts the extracellular matrix transcriptome

(A) Heatmap for sample distances and (B) principal component analysis (PCA) showing the variance in the samples used for RNAseq analysis: 3 clonal lines for each genotype (*Thap1*$^{+/+}$ and *Thap1*$^{-/-}$) from 3 stages of OL cultures: OPC, DIV2 & DIV4. (C) *Thap1*$^{+/+}$ vs *Thap1*$^{-/-}$ differentially expressed genes (DEG) selective for either DIV2 (703) or present at both DIV2 & DIV4 (277) used for enrichment analysis of GO (Gene Ontology) terms to identify overrepresented biological pathways. Graphs show the most significantly overrepresented GO terms (p<0.001, sorted as fold overrepresented) in OL from DIV2 (GENEONTOLOGY, GO - Biological Process), and from the DIV2 & DIV4 overlapping group (DAVID, GO - GO - Biological process).



E Secreted GAG / Media (µg/mg)

	OPC		OL	
	Thap1$^{+/+}$	*Thap1*$^{-/-}$	*Thap1*$^{+/+}$	*Thap1*$^{-/-}$
CS	73.53	134.07	9.40	23.56
HS	3.67	7.55	0.44	1.01
HA	1.40	2.61	0.86	2.42

F GAG / Homogenate (µg/mg)

	OPC		OL	
	Thap1$^{+/+}$	*Thap1*$^{-/-}$	*Thap1*$^{+/+}$	*Thap1*$^{-/-}$
CS	151.64	405.42	169.35	550.80
HS	46.85	171.20	79.74	121.03
HA	161.77	186.62	290.16	448.82

G Secreted Chondroitin Sulfate GAG / Media (µg/mg)

	OPC		OL	
	Thap1$^{+/+}$	*Thap1*$^{-/-}$	*Thap1*$^{+/+}$	*Thap1*$^{-/-}$
C-4S	63.15	123.88	7.77	19.76
C-6S	7.44	6.34	1.00	2.48
C-2,6S	0.82	1.17	0.13	0.44
C-4,6S	0.79	1.05	0.09	0.15

H Chondroitin Sulfate GAG / Homogenate (µg/mg)

	OPC		OL	
	Thap1$^{+/+}$	*Thap1*$^{-/-}$	*Thap1*$^{+/+}$	*Thap1*$^{-/-}$
C-4S	134.74	376.49	150.48	472.26
C-6S	13.62	13.40	15.21	48.29
C-2,6S	0.25	1.72	0.28	10.14
C-4,6S	2.09	5.00	2.33	5.22

Figure S2. *Thap1* null oligodendrocytes accumulate and secrete excess glycosaminoglycans: (A) Representative images (scale bar 50µm) and (B) corresponding quantification for *Thap1*$^{+/+}$ and *Thap1*$^{-/-}$ OL cells (DIV4) immunostained with 2H6 (C-4S-GAG; green) under non-permeabilizing conditions. *Thap1*$^{-/-}$ OL exhibit significant accumulation of CS-GAG (*Thap1*$^{+/+}$ = 3.45 AU ± 0.21; *Thap1*$^{-/-}$ = 4.84 AU ± 0.28; t-test; $t_{(93)} = 3.45$, $p<0.0001$. (C) Representative images demonstrating accumulation of HA-GAG (stained using hyaluronan binding protein, HABP; red) in *Thap1*$^{-/-}$ OL cultures (+T3-DIV4). Arrows indicate surrounding regions of select OL with HA-GAG accumulation. (D) Carbazole assay (demonstrating GAG content) of total cell homogenate extracted from *Thap1*$^{+/+}$ and *Thap1*$^{-/-}$ OPC and OL (+T3-DIV4). Absorbance at 525 nm for OPC; *Thap1*$^{+/+}$ = 0.066 AU ± 0.0006; *Thap1*$^{-/-}$ = 0.094 AU ± 0.0022; t-test: $t_{(2)} = 12.70$; $p = 0.006$ & OL - *Thap1*$^{+/+}$ = 0.036 AU ± 0.0023; *Thap1*$^{-/-}$ = 0.059 AU ± 0.0019; t-test: $t_{(2)} = 7.96$; $p = 0.015$. Table displaying the estimated amount of (E-F) CS, HS and HA GAGs (µg/mg) and (G-H) composition of mono (C-4S, C-6S) and bi-sulfated (C-2,6S and C-4,6S) CS-GAGs (µg/mg) (E,G) secreted in the media or (F,H) from cell homogenate from *Thap1*$^{+/+}$ and *Thap1*$^{-/-}$ OPC and OL (+T3-DIV4) using GRIL LC/MS analysis.

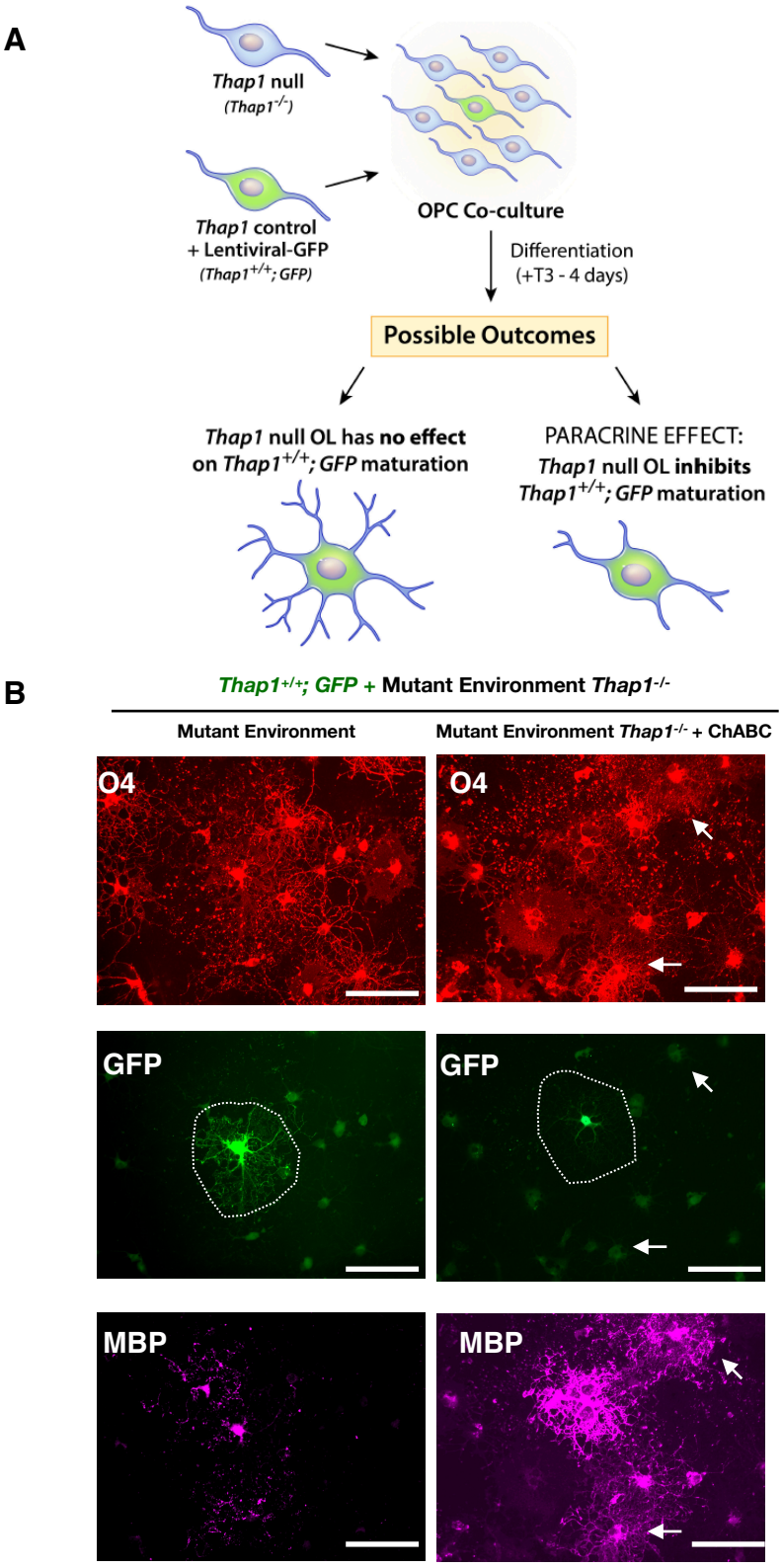


Figure S3. Excess glycosaminoglycan secretion by *Thap1* null cells impairs oligodendroglial maturation via a non-cell autonomous mechanism

(A) Schematic illustrating co-culture experimental paradigm to test for paracrine effects of GAGs secreted by the OL lineage. *Thap1* null (*Thap1*$^{-/-}$) OL were co-cultured with LV-GFP labelled control (*Thap1*$^{+/+}$*; GFP*) OPC at 10:1 ratio in differentiation media for four days (+T3). The two possible outcomes from this experimental paradigm are depicted in the illustration. (B) Representative images (scale bar 100 µm) of *Thap1*$^{+/+}$*;GFP OL* cultured in mutant environments for 4 days (+T3) without (left panel) and with ChABC (0.1U/ml) treatment (right panel). Cultures are stained for O4, MBP, & DAPI. Arrows in the right panels indicate examples of *Thap1*$^{-/-}$ OL that are expressing MBP upon treatment with ChABC

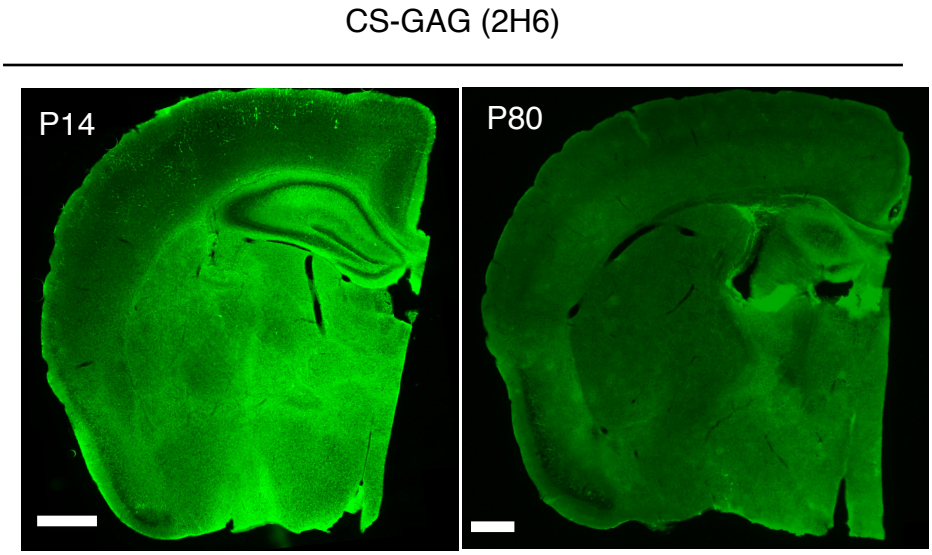


Figure S4. CS-GAG staining in juvenile and adult CNS tissue.
Representative image (scale bar - 500µm) demonstrating C-4S GAG (2H6) staining for control mouse P14 and P80 tissue highlighting the strong decrease in the CS-GAG staining in adult CNS.

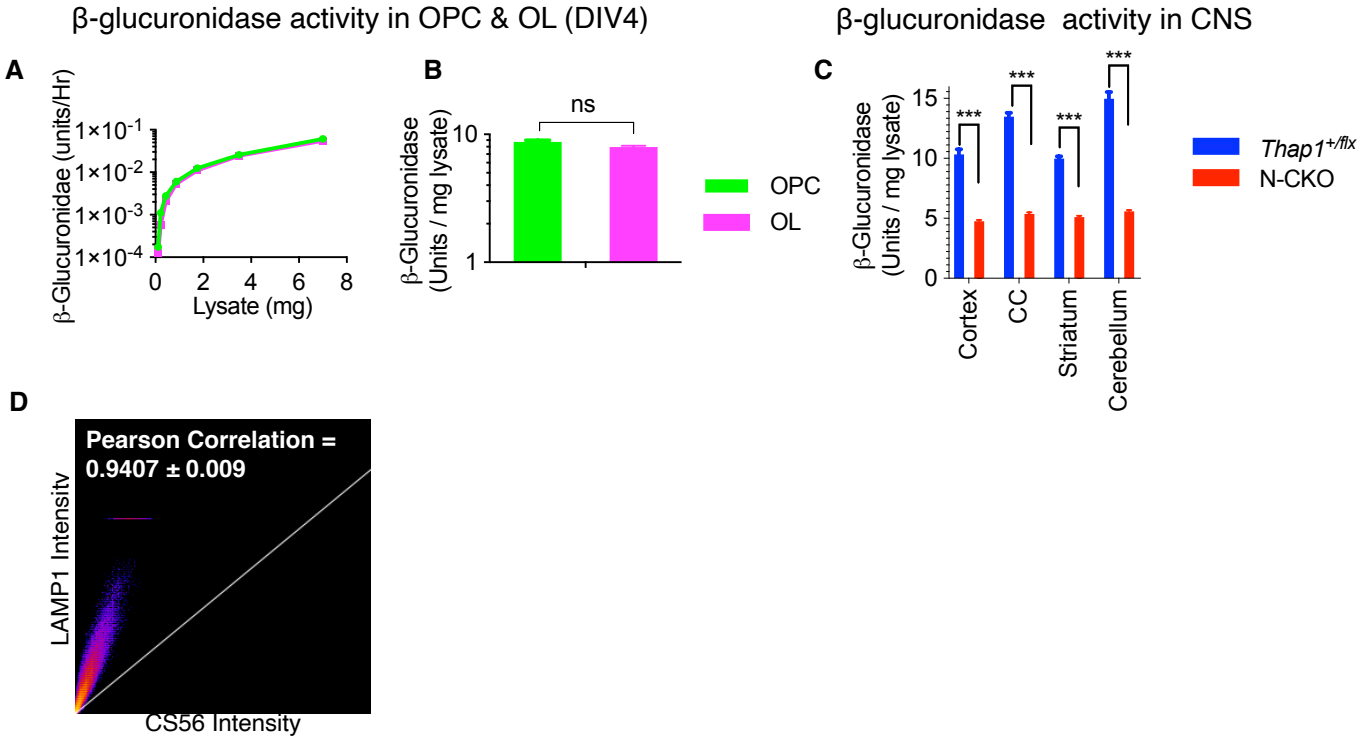


Figure S5. β-glucuronidase activity during oligodendrocyte differentiation.

(A) β-glucuronidase activity (Units / Hr) for *Thap1*$^{+/+}$ OPC and OL (+T3-DIV4) lysate (0 - 8 µg; x-axis; STAR Methods). (B) Normalized β-glucuronidase activity (Units/mg Lysate) for OPC = 8.71 U / mg ± 0.27 and OL = 7.97 U / mg ± 0.18; t-test: $t_{(2)} = 2.18$; $p = 0.16$). (C) β-glucuronidase activity in multiple brain regions (cerebral cortex, corpus callosum, striatum and cerebellum) of P21 THAP1 N-CKO (*Thap1*$^{flx/-}$; nestin-*Cre*$^{+}$) and control (*Thap1*$^{+/flx}$) mice. Bar graph shows mean ± SEM values of normalized β-glucuronidase activity Units/Lysate (mg). Cerebral cortex - *Thap1*$^{+/flx}$ = 10.32 U / mg ± 0.43 ; N-CKO = 4.76 U / mg ± 0.106; t-test: $p < 0.0001$; corpus callosum - *Thap1*$^{+/flx}$ = 13.49 U / mg ± 0.33 ; N-CKO = 5.38 U / mg ± 0.129; t-test: $p < 0.0001$; striatum - *Thap1*$^{+/flx}$ = 9.98 U / mg ± 0.183 ; N-CKO = 5.10 U / mg ± 0.10; t-test: $p < 0.0001$ and cerebellum - *Thap1*$^{+/flx}$ = 14.97 U / mg ± 0.552 ; N-CKO = 5.58 U / mg ± 0.117; t-test: $p < 0.0001$) from P21 mice. (D) Pearson's coefficient value shows ~ 95% colocalization (R = 0.94 ± 0.009; multiple region of interest (ROI) for N=50 cells) for CS-GAG (CS-56) and lysosomes (LAMP1) in OL (DIV4). Graph demonstrating co-localization of CS-56 and LAMP1 pixels for the image represented in FIg. 4N calculated using Image J.

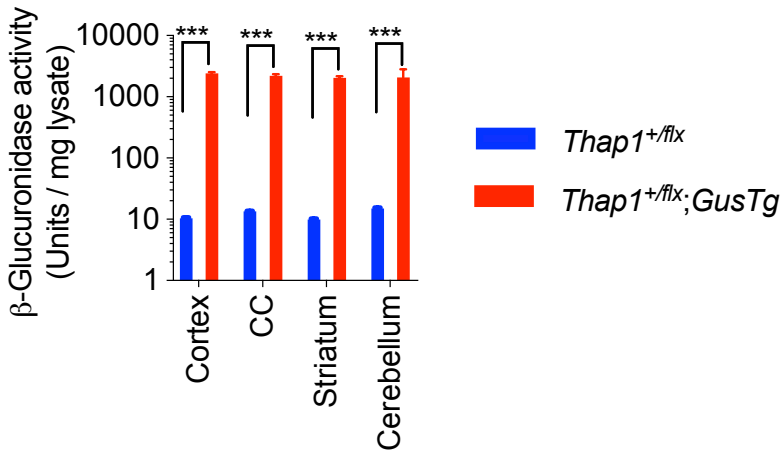


Figure S6. β-glucuronidase overexpression in CNS tissue.

β-glucuronidase activity in multiple brain regions (cerebral cortex, corpus callosum, striatum and cerebellum) of P21 *Thap1*$^{+/flx}$ and *Thap1*$^{+/flx}$;Tg^{GUS} mice. Bar graph shows mean ± SEM values of normalized β-glucuronidase activity Units/Lysate (mg). Cerebral cortex - *Thap1*$^{+/flx}$ = 10.32 U / mg ± 0.43 ; *Thap1*$^{+/flx}$;Tg^{GUS} = 2402 U / mg ± 120.6; t-test: $p < 0.0001$; corpus callosum - *Thap1*$^{+/flx}$ = 13.49 U / mg ± 0.33 ; *Thap1*$^{+/flx}$;Tg^{GUS} = 2208 U / mg ± 140.2; t-test: $p < 0.0001$; striatum - *Thap1*$^{+/flx}$ = 9.98 U / mg ± 0.183 ; *Thap1*$^{+/flx}$;Tg^{GUS} = 2056 U / mg ± 110.3; t-test: $p < 0.0001$ and cerebellum - *Thap1*$^{+/flx}$ = 14.97 U / mg ± 0.552 ; *Thap1*$^{+/flx}$;Tg^{GUS} = 2078 U / mg ± 724.9; t-test: $p < 0.0001$.

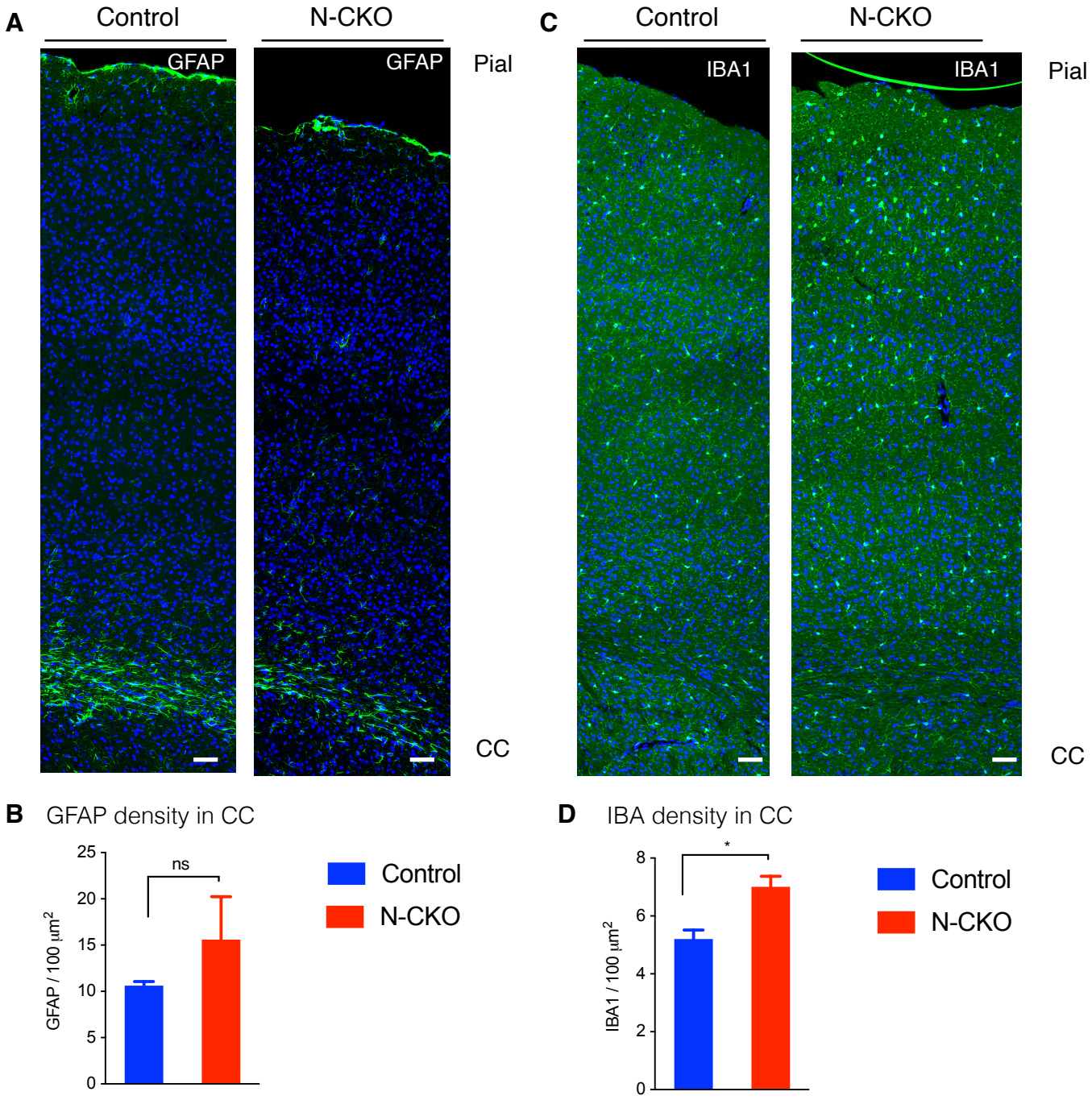


Figure S7. Astrocyte and microglia cells comparable in whole in whole brain *Thap1* CNS. Representative images (scale bar 100µm) and corresponding quantification for P80 cerebral cortex from control (*Thap1*$^{+/flx}$) and N-CKO (*Thap1*$^{flx/-}$;Nestin-cre) immunostained with (A-C) GFAP and (C-D) IBA1. The graphs represent the density (number/100µm^2; y-axis) of (B) GFAP (Control = 10.62 cells/100µm^2 ± 0.45; N-CKO = 15.61 cells/100µm^2 ± 4.63; t-test; $t_{(3)}$ = 1.42, p = 0.25) and (C) IBA1 (Control = 5.2 cells/100µm^2 ± 0.31; N-CKO = 7.008 cells/100µm^2 ± 0.365; t-test; $t_{(3)}$ = 3.7, p = 0.034).

Table S1. Primer list and sequences used for the study

Genotyping Primers		
Gene	Forward	Reverse
Thap1	GCATAGGACAGAGCCTTTCAG	GATGCCAATACCTGATTGGAG
	TGCTGGGTGTTGGAAAATAA	
Cre	CTAGGCCACAGAATTGAAAGATCT	GTAGGTGGAAATTCTAGCATCATCC
	GCGGTCTGGCAGTAAAAACTATC	GTGAAACAGCATTGCTGTCACTT
Tg^{GUS}	CTA GGC CAC AGA ATT GAA AGA TCT	GTA GGT GGA AAT TCT AGC ATC ATC C
	CTG TGG CTG TCA CCA AGA GC	GGA CAC TCA TCG ATG ACC AC
$Tg^{MOBP\text{-}eGFP}$	CCTACGGCGTGCAGTGCTTCAGC	CGGCGAGCTGCACGCTGCGTCCTC
ChIP Primers		
Gene	Forward	Reverse
Gusb	GGGACCTGGAAGAGTCACAT	AGCCTTGGGGCCTCTAGATA
qRT-PCR Primers		
Gene	Forward	Reverse
Gusb	TCATGACGAACCAGTCACCG	CGGTTTCGTTGGCAATCCTC
Rpl19	TCAGGCTACAGAAGAGGCTTGC	ACAGTCACAGGCTTGCGGATG

Table S2: Antibodies used in the study

Antibodies used for immunostaining		
Antibody	Source	Catalog
Rabbit anti-NG2	Millipore	AB5320
Rat anti-MBP	Millipore	MAB386
Goat Anti-OLIG2	R&D Systrems	AF2418
Mouse anti-CS56	Thermofisher	MA1-83055
Mouse anti-2H6	Cosmo Bio LTD	CAC-NU-07-001
Rat anti-LAMP1	Santacruz	sc-19992
Mouse anti-O4	Hybridoma	(Dr. Roman Giger, U. Michigan)
Rabbit anti-GFAP	Agilent	Z033429-2
Rabbit anti-IBA1	Thermofisher	PA5-27436
Antibodies used for chromatin immunoprecipitation (ChIP)		
Goat anti-THAP1	Santacruz	sc-98174
Rabbit anti-YY1	Santacruz	sc-281
Rabbit anti-H3K4me3	Active Motif	91264
Normalized anti-Rabbit IgG	Santacruz	sc-2027

Additional data

Dataset S1 (separate file, excel). *Thap1*$^{+/+}$ Vs *Thap1*$^{-/-}$ differentially regulated genes (DEG) in OPC and OL (DIV2 & DIV4).

Dataset S2 (separate file, excel). Gene Ontology analysis for *Thap1*$^{+/+}$ Vs *Thap1*$^{-/-}$ differentially regulated genes (DEG) in OPC and OL (DIV2 & DIV4).

Dataset S3 (separate file, excel). THAP1 regulated genes common to ECM metabolism.

SI References

1. D. Yellajoshyula *et al.*, The DYT6 Dystonia Protein THAP1 Regulates Myelination within the Oligodendrocyte Lineage. *Dev Cell* **42**, 52-67 e54 (2017).
2. J. J. Winters *et al.*, Congenital CNS hypomyelination in the Fig4 null mouse is rescued by neuronal expression of the PI(3,5)P(2) phosphatase Fig4. *J Neurosci* **31**, 17736-17751 (2011).
3. D. Yellajoshyula, E. S. Patterson, M. S. Elitt, K. L. Kroll, Geminin promotes neural fate acquisition of embryonic stem cells by maintaining chromatin in an accessible and hyperacetylated state. *Proc Natl Acad Sci U S A* **108**, 3294-3299 (2011).
4. R. Lawrence *et al.*, Evolutionary differences in glycosaminoglycan fine structure detected by quantitative glycan reductive isotope labeling. *J Biol Chem* **283**, 33674-33684 (2008).