Supporting Information for:

Mapping posttranscriptional regulation of the human glycome uncovers microRNA defining the glycocode.

Praveen Agrawal[1‡], Tomasz Kurcon[1‡], Kanoelani T. Pilobello[2‡], John F. Rakus[1‡], Sujeethraj Koppolu[1], Bianca Batista[2], William S. Eng[1], Ku-Lung Hsu[3], Yaxuan Liang[1], and Lara K. Mahal[1,2,3]∗

Includes:

Supplementary Figures 1-9
Supplementary Tables 1-3
Supplementary Methods and Sequences.
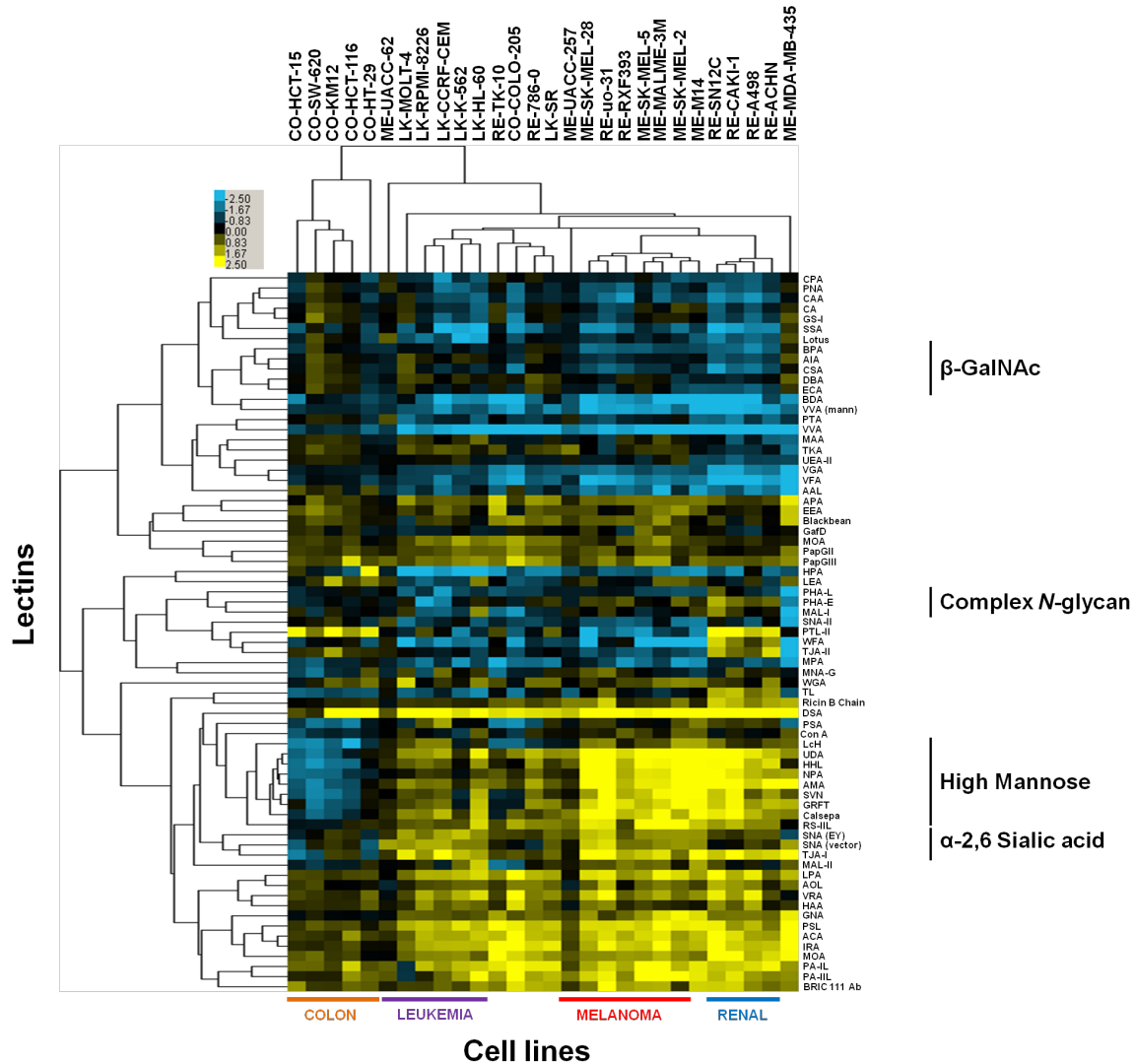
Supplementary Figures

Figure S1. Ratiometric comparison of 28 NCI-60 cell lines from 4 tissue types. Equal amounts (1.5 μg by protein) of Cy5-labeled samples and Cy3-labeled reference were analyzed on the lectin microarray. Median-normalized $\log_2$ ratios (S/R) were hierarchically clustered using centered Pearson correlation coefficient as the distance metric and average linkage analysis (n=71 lectins). Heat map is shown. Yellow = $\log_2(S/R) > \log_2(S_{median}/R_{median})$, Blue = $\log_2(S_{median}/R_{median}) > \log_2(S/R)$.
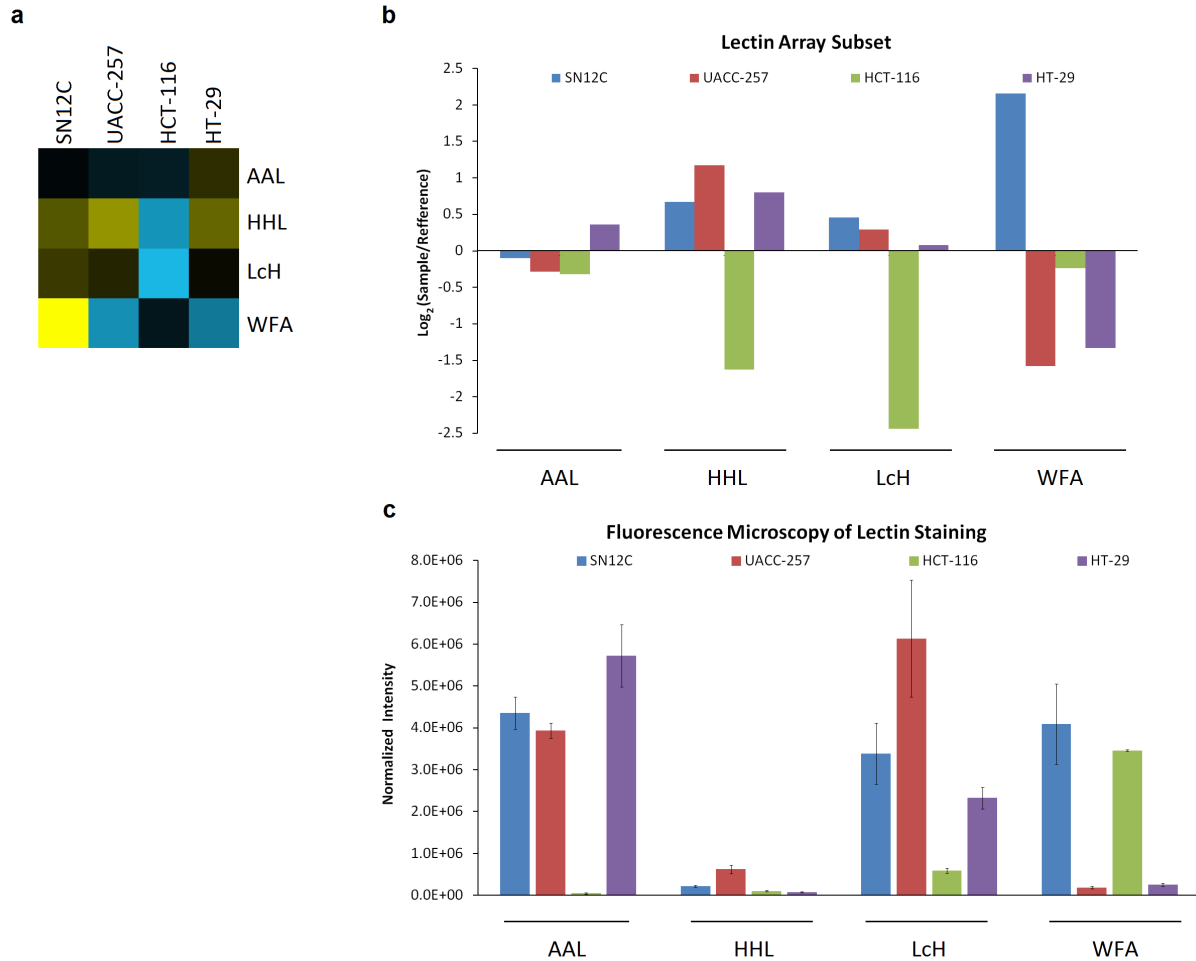
Figure S2 Validation of lectin microarray data via fluorescence microscopy. (*a*) Subset of lectin microarray data from four cell lines used for validation. Data is from the expanded NCI-60 dataset shown in Figure 1. (*b*) Graphical representation of log$_2$ intensity values represented in (*a*). (*c*) SN12C (renal), UACC-257 (melanoma), HCT-116 (colon), HT-29 (colon) cells were grown to ~80% confluency, fixed with 4% paraformaldehyde and stained with biotinylated lectins (AAL, HHL, LcH, WFA, 10 µg/ml in PBS) followed by Cy5-streptavidin (1:100, Invitrogen). Fluorescence intensity was visualized by fluorescence microscopy. Graph represents the average fluorescence for 6 areas per image chosen in brightfield and normalized to cell count (>29 cells/image, brightfield) for 4 random images for each cell line/lectin combination (n=24). Standard deviation is shown.

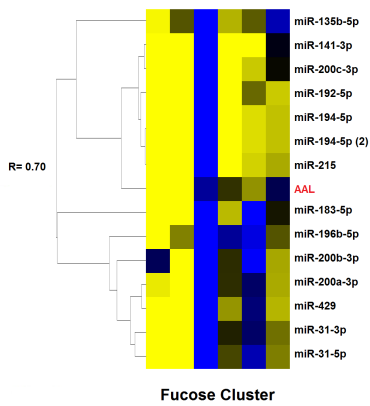**a**   High Mannose Clusters
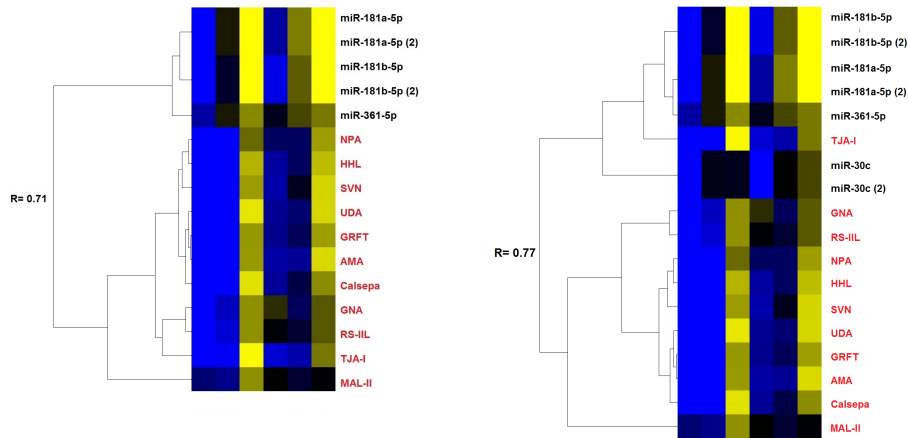
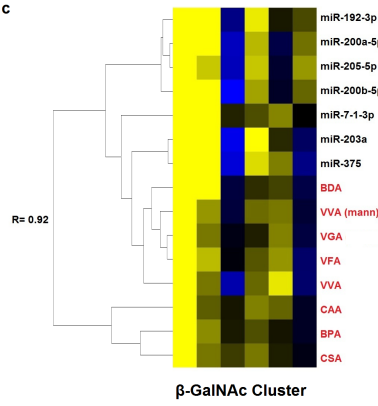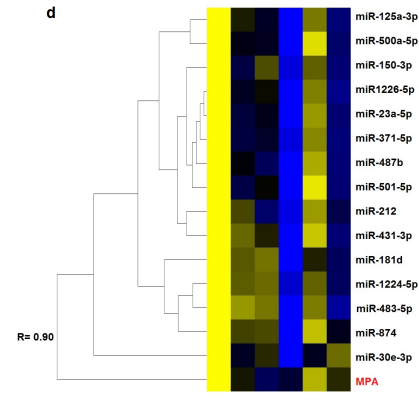Pearson corelation: Uncentered   Centered

b   **c**   **d**

Fucose Cluster   β-GalNAc Cluster   Tn/T-Antigen Cluster

e   **f**
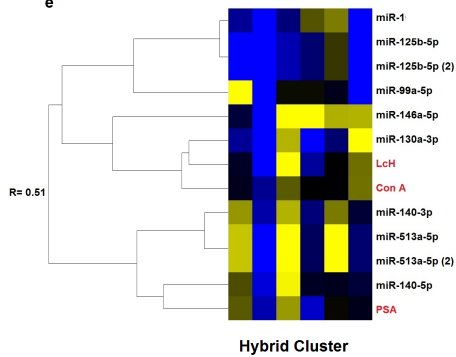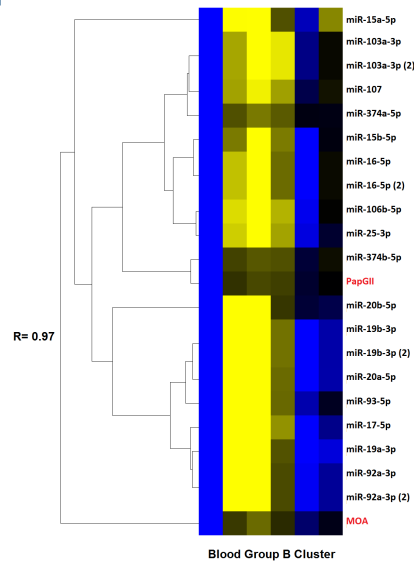
Hybrid Cluster   Blood Group B Cluster

Figure S3. MiRNA/Lectin Clusters. (*a*) High mannose clusters: Uncentered Pearson Cluster (as in Fig. 2c, R=0.71, one-tailed *P*=0.06) and Centered Pearson cluster containing miRNA-30c (R=0.77, one-tailed *P*<0.05). (*b*) Fucose cluster (R=0.70, one-tail *P*=0.060). (*c*) β-GalNAc cluster (R=0.92, one-tailed *P*<0.05). (d) Tn/T-Antigen cluster (R=0.90, one-tailed *P*<0.05). (*e*) Hybrid cluster (R= 0.51, one-tailed *P*=0.15). (*f*) Blood group B cluster (R= 0.97, one-tailed *P*<0.05). *P*-values were calculated using the Pearson correlation coefficient R and n=6. Lectins are indicated in red.
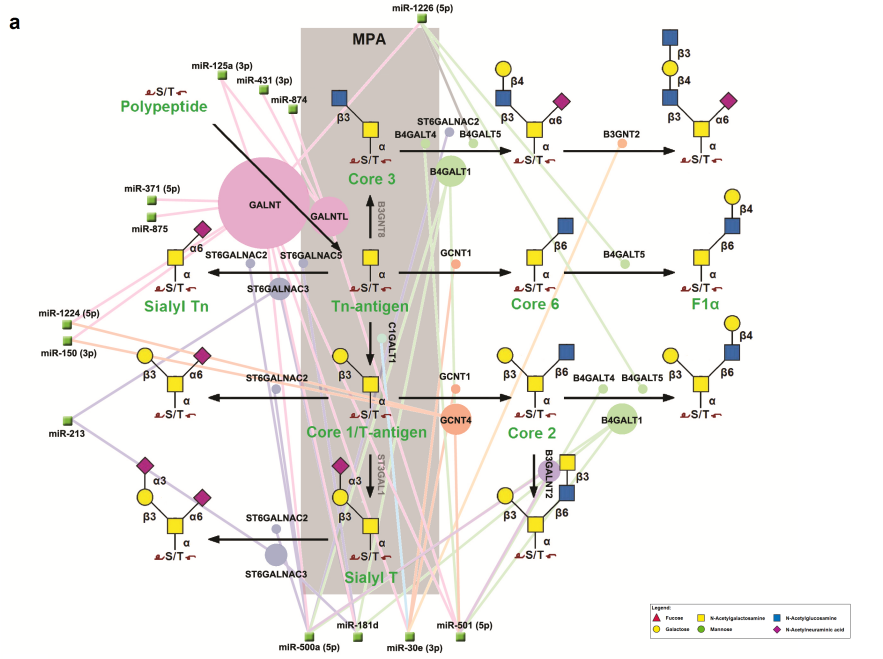
Figure S4. Detailed representation of additional glycan biosynthetic networks derived from lectin/miRNA clusters. (*a*) Tn/T network, (*b*) hybrid network, (*c*) blood group B network. Grey boxes: epitopes recognized b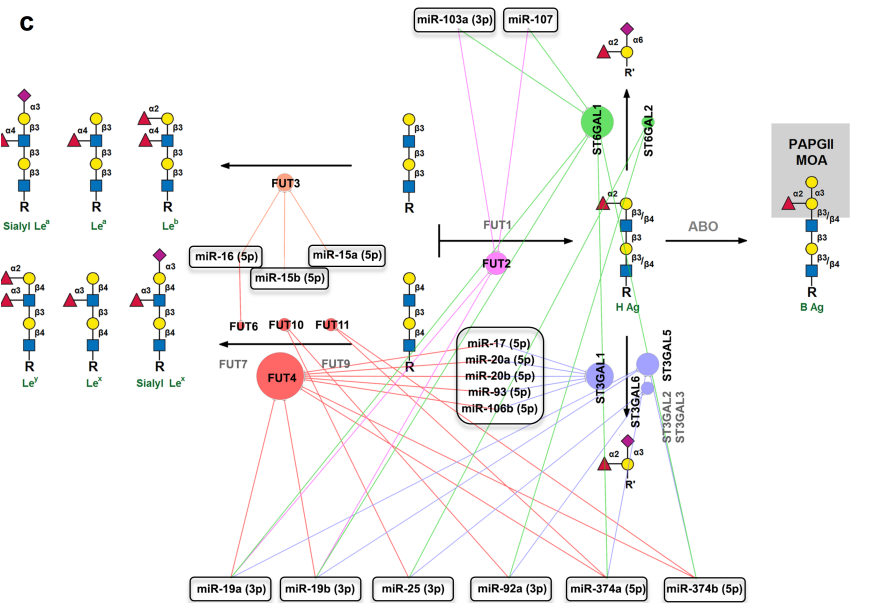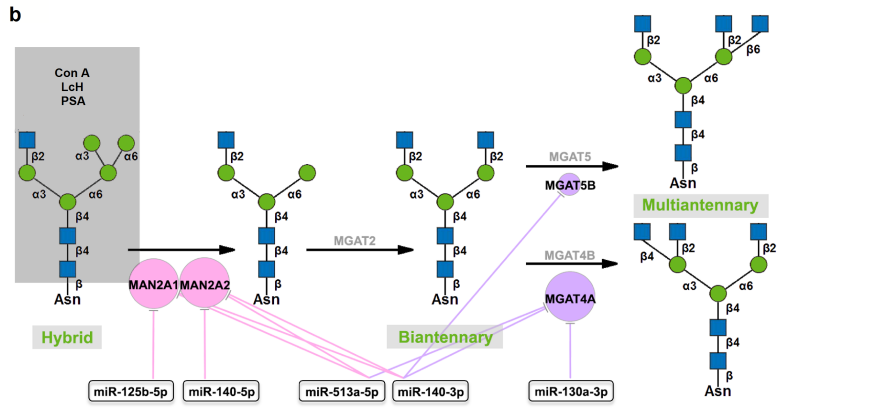y lectins. Bubbles: predicted miRNA targets, the bubble size reflects number of miRNA targeting gene. Lines: connect miRNA with targets, Genes in silver: genes in the pathway that are not targeted. (*a*) Tn-/T-antigen network contains multiple miRNAs predicted to target enzymes involved in branching at the 6-position of GalNAc. Modification at this position abrogates binding of MPA. Due to the large number of predicted targets from GALT and GALNTL family members each family are represented by one bubble. (*b*) Hybrid network is composed of lectins with overlapping specificity for hybrid and biantennary *N*-linked glycan structures (Con A, PSA and LcH). α-mannosidase II family members (*MAN2A1*, *MAN2A2*) are common relevant predicted target of the miRNAs in the cluster and a regulatory switch between hybrid and complex *N*-glycans. (*c*) Blood group network consists of multiple miRNAs targeting enzymes diverging substrate pools away from formation of the blood group B epitope recognized by PapGII (37) and MOA, lectins from this cluster.

a

HHL

PSA

b
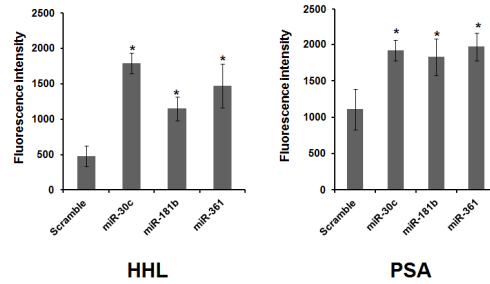
	GNA		HHL		NPA		PSA		LcH Control	
Treated	Avg. Int. (a.u.)	StDev (±)	Avg. Int. (a.u.)	StDev (±)	Avg. Int. (a.u.)	StDev (±)	Avg. Int. (a.u.)	StDev (±)	Avg. Int. (a.u.)	StDev (±)
Scramble mimic	473	133	475	145	850	121	1109	281	1147	418
miR-30c-5p	969	200	1793	147	2120	343	1919	142	1312	513
miR-181b-5p	820	123	1152	169	1507	211	1827	252	1068	400
miR-361-5p	1383	200	1472	312	1601	198	1968	193	1278	432

c

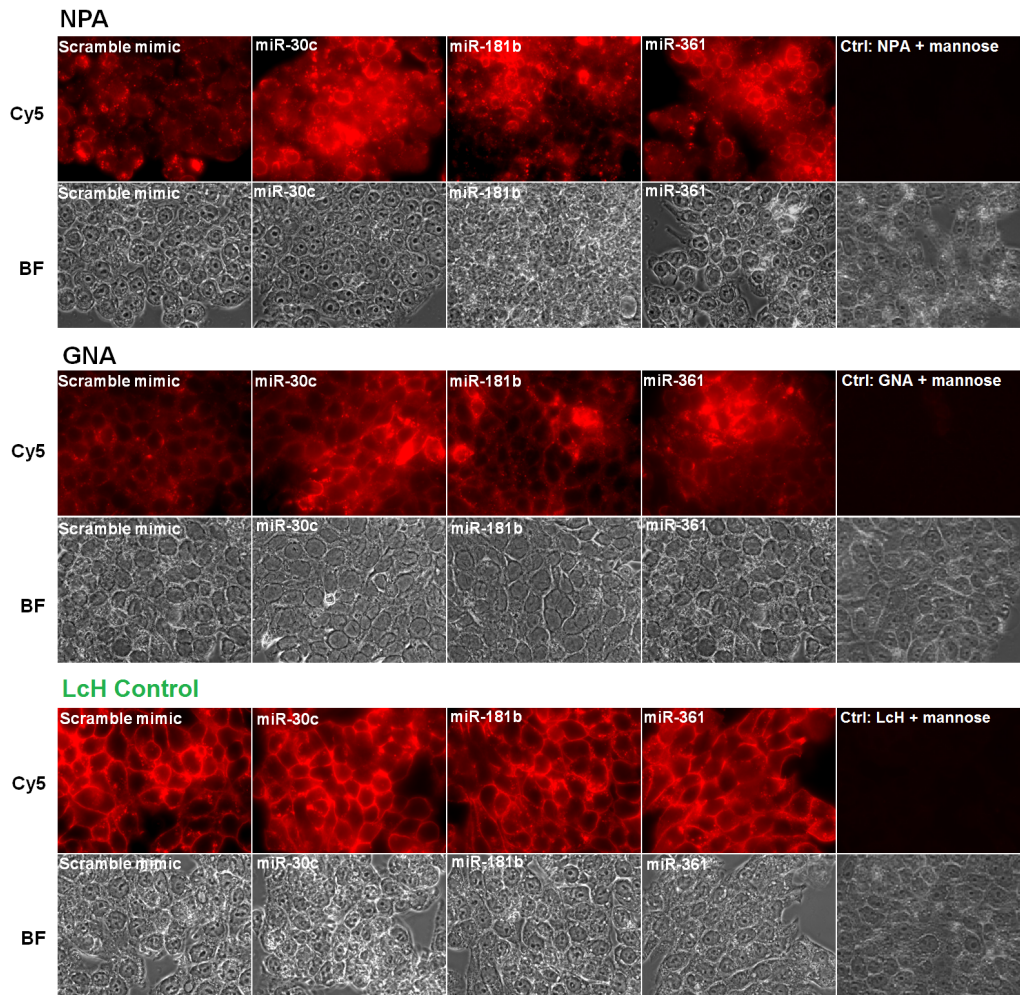**Figure S5. MiR-30c, -181b-5p and -361-5p increase cell surface high mannose.** HT-29 cells were treated with miR-30c, -181b-5p (miR-181b), -361-5p (miR-361) mimics or a control (scramble mimic, 50 nM) for 96 hour, fixed and labeled with biotinylated lectins/ Cy5-streptavidin and visualized by fluorescence microscopy. (*a*) Quantitation of staining with HHL and PSA. (*b*) Summary table of lectin

staining following treatment (scramble in bold blue). For both (*a and b*) data from 10 images per replicate (3 biological replicates, 30 images total, 180 random areas) were averaged to generate graphs. Areas with clear fluorescent aggregates were excluded from our analysis. Error bars represent standard deviations. *P*-values were calculated using the two-tailed unpaired Student's *t*-test with equal variances for comparison to scramble control, $*P < 0.05$. LcH is a specificity control for PSA staining of high mannose. LcH and PSA both bind core fucosylated *N*-linked glycans but PSA has much higher affinity for high mannose epitopes (Tateno, H., Nakamura-Tsuruta, S., and Hirabayashi, J., *Glycobiology* **2009**, *19*. 527–536) and thus is more sensitive to changes in high mannose levels as observed herein. (*c*) Representative fluorescence and brightfield images for staining with NPA, GNA and LcH are shown. For all lectins, in preincubation with monosaccharide inhibited binding.
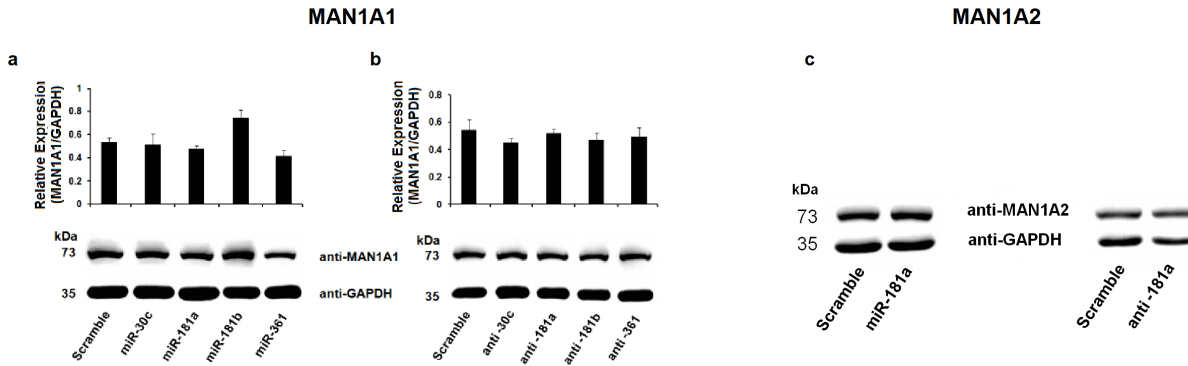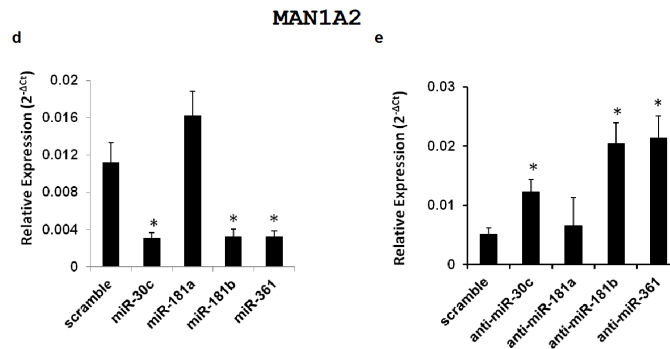
Figure S6. Analysis of *MAN1A1* and *MAN1A2* expression in cells transfected with miR-30c, -181a, -181b and miR-361 or scramble. (*a-b*) *MAN1A1* protein expression as assessed by Western blot analysis in HT-29 following 72 h treatment with miRNA mimics (*a*) or inhibitors (*b*). Graphs show average signal normalized to GAPDH of three biological replicates. Representative images corresponding to the graphs are shown. (*c*) *MAN1A2* protein expression following 72 h treatment with miR-181a mimics or inhibitors as assessed by Western blot analysis. No statistical difference in *MAN1A2* expression was observed upon treatment. Image shown is representative of 3 independent experiments. (*d-e*) RT-PCR analysis of *MAN1A2* expression in SN12C following 72 h treatment with miR-30c, -181a, -181b and miR-361 or scramble mimics (*d*) or corresponding inhibitors (*e*). Graphs show average relative expression normalized to *GAPDH* of three biological replicates. For all graphs, error bars denote standard deviation and P-values <0.05 using the Student's t-test comparison with scramble control are denoted by *.
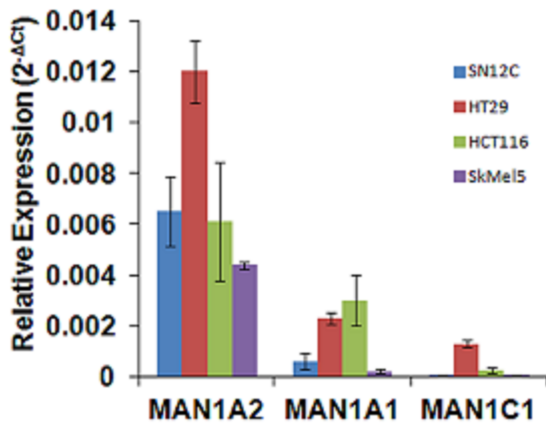
Figure S7. RT-qPCR analysis of the endogenous levels of *MAN1A2*, *MAN1A1* and *MAN1C1* in NCI-60 cell lines HT-29 and HCT116, SN12C and SkMel-5. Graphs show relative expression normalized to *GAPDH*. Error bars denote standard deviation.
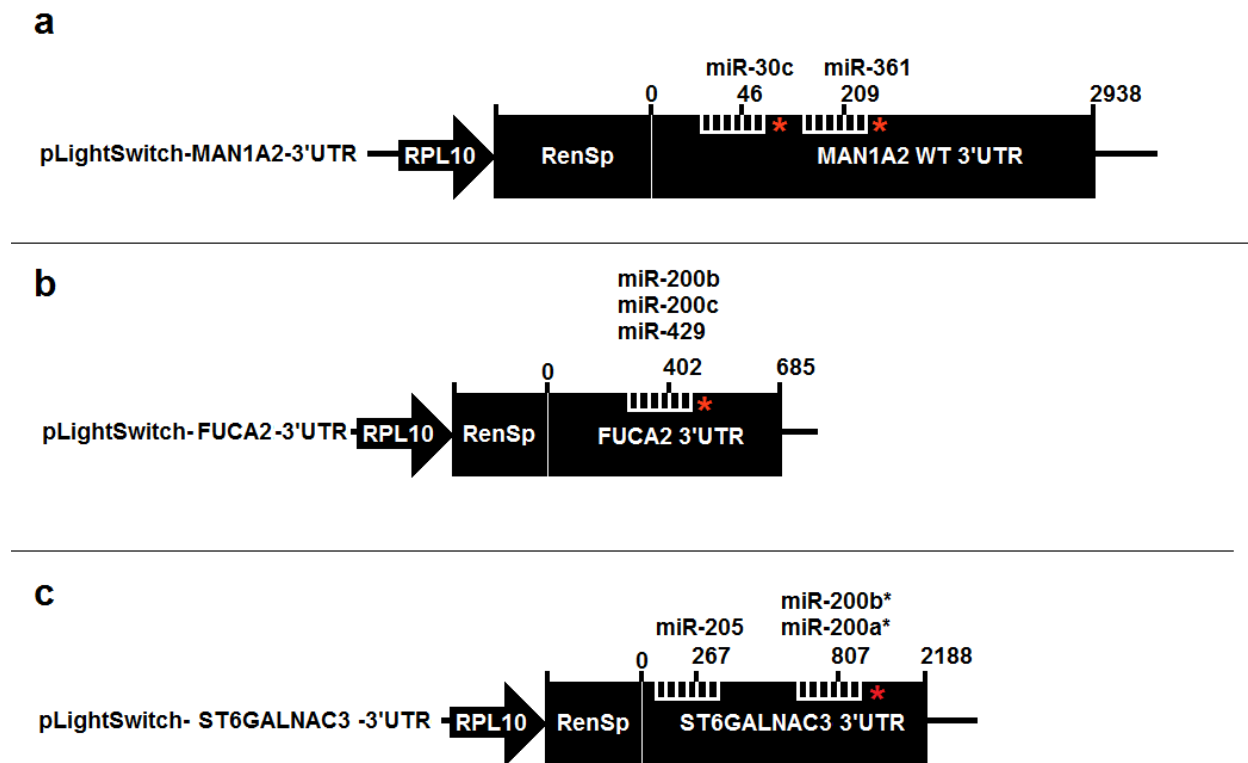
Figure S8. Schematic representations of luciferase constructs. (a) *MAN1A2*, (b) *FUCA2*, (c) *ST6GALNAC3*. Mutated miRNA binding sites are indicated with a red asterisk. See Supporting Information Table S3 for details on constructs.
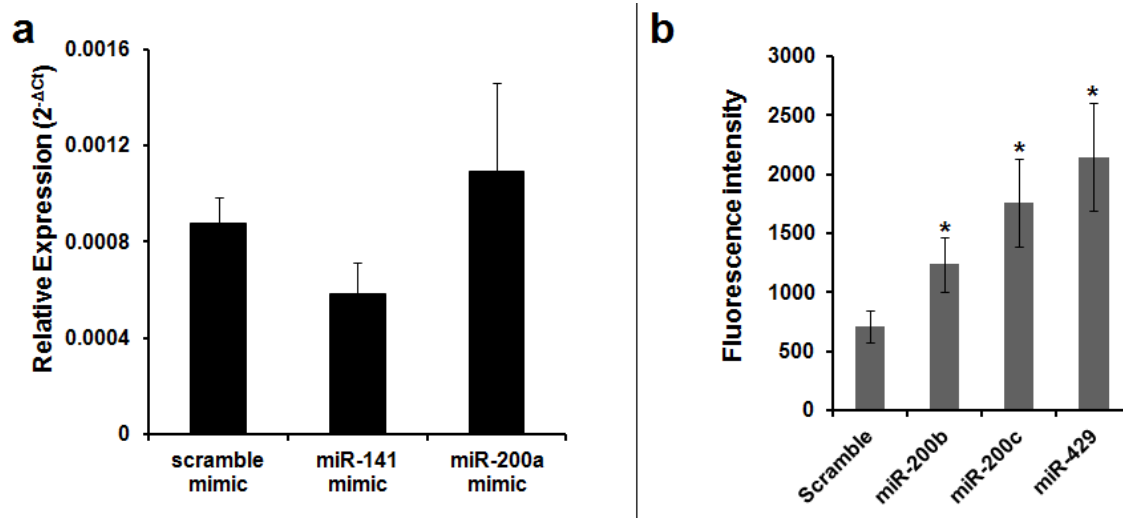
Figure S9. Effects of miR-200f on *FUCA2* and fucose expression. (*a*) Effects of miR-200a and miR-141 on *FUCA2* gene expression in HT-29. HT-29 cells were treated with miRNA or scramble mimics (50 nM) for 72 h. *FUCA2* mRNA levels were analyzed in RT-qPCR as previously described. Graphs show average relative expression normalized to *GAPDH* of three biological replicates. Error bars denote standard deviation. No significant difference was observed in *FUCA2* mRNA levels for miR-200a or -141 treated samples compared to scramble using the Student's t-test. (*b*) Quantitative analysis of AAL staining of HT-29 treated with miR-200b, -200c, -429 or scramble for 96 hours. Graph shows average data from 6 random regions/image, 10 images per replicate, 3 biological replicates (30 images total, 180 random areas). Error bars represent standard deviations. *P*-values were calculated using the two-tailed unpaired Student's *t*-test with equal variances for comparison to scramble control, $*P < 0.05$.

Supplementary Tables

Table S1. **Lectin print conditions (Original dataset).** Abbreviations: Gal: Galactose, Fuc: Fucose, Man: Mannose, Lac: Lactose, GalNAc: *N*-acetylgalactosamine, GlcNAc: *N*-acetylglucosamine.

| Lectin source	Lectin abbreviation	Print concentration (µg/ml)	Print monosaccharide
Abrus precatorius	APA	500	Gal
Anguilla anguilla	AAA	1000	Fuc
Arachis hyogaea	PNA	500	Gal
Artocarpus integrifolia	AIA	500	Gal
Bauhinia purpurea	BPA	500	Gal
Bryonia dioica	BDA	500	Gal
Canavalia ensiformis	Con A	500	Man
Cancer antennarius	CCA	500	Lac
Caragana arborescens	CAA	500	Gal

Cicer arietinum	CPA	1000	Lac
Colchicum autumnale	CA	500	Gal
Cystisus scoparius	CSA	500	Gal
Datura stramonium	DSA	500	Lac
Dolichos biflorus	DBA	500	Gal
Erythrina cristagalli	ECA	500	Gal
Eunonymus europaeus	EEA	500	Lac
Galanthus nivalis	GNA	1000	Man
Griffonia simplicifoia	GS-I	500	Gal
Griffonia simplicifoia	GS-II	500	GlcNAc
Salvia sclarea	SSA	2000	Lac
Helix pomatia	HPA	500	Gal
Lens culinaris	LcH	500	Man
Limax flavus	LFA	500	Lac
Lotus tetragonolobus	Lotus (LTL)	500	Fuc
Lycopersicon esculentum	LEA	500	GlcNAc
Limulus polphemus	LPA	500	Lac
Maackia amurensis	MAA	500	Lac
Narcissus pseudonarcissus	NPA	1000	Man
Persea americana	PAA	500	GlcNAc
Phaseolus lunatus	LBA	1000	Gal
Phaseolus vulgaris Erythroagglut*inin*	PHA-E	500	Lac
Phaseolus vulgaris Leucoagglutinin	PHA-L	500	Gal
Pisum sativum	PSA	500	Man
Aspergillus oryzae	AOL	500	Fuc
Psophocarpus tetragonolobus	PTA	500	Gal
Robinia pseudoacacia	RPA	500	GlcNAc
Sambucus nigra	SNA	500	Gal
Sambucus nigra	SNA-II	500	Gal
Sambucus nigra	SNA	500	Gal
Solanus tuberosum	STA	500	GlcNAc
Sophora japonica	SJA	500	Gal
Trichosanthes kirilowii	TKA	500	Gal
Triticum vulgare	WGA	500	GlcNAc
Tulipa sp.	TL	1000	GlcNAc
Ulex europaaeus I	UEA-I	500	Fuc
Ulex europaaeus II	UEA-II	1000	GlcNAc
Uritica dioica	UDA	1000	GlcNAc

Vicia graminea	VGA	500	Gal
Vicia villosa	VVA	500	Gal
Vicia villosa	VVA (mann)	500	Man
Vigna radiata	VRA	500	Gal
Vicia faba	VFA	500	Gal
Wisteria floribunda	WFA	500	Gal
Black bean	Blackbean	1000	Lac
Ricin B chain	Ricin B Chain	1000	Lac
Hippeastrum Hybrid	HHL	500	Man
Maackia amurensis-I	MAL- I	1000	Lac
Maackia amurensis-II	MAL-II	500	Lac
Psophocarpus tetragonolobus	PTL-I	1000	Gal
Psophocarpus tetragonolobus	PTL-II	1000	Gal
Cyanovirin	CVN	1000	Man
Scytovirin	SVN	500	Man
Griffithsin	GRFT	500	Man
Cholera Toxin from Vibrio cholerae	Cholera Toxin	2000	Lac
Amaranthus Caudatus	ACA	500	Lac
Marasmium oreades	MOA	500	Gal
Maclura pomifera	MPA	500	Gal
Polyporus squamosus	PSL	500	Gal
Trichosanthes japonica	TJA-I	1000	Lac
Trichosanthes japonica	TJA-II	500	Lac
Aleuria aurantia	AAL	500	Fuc
antibody	7-LE Ab	500	Lac
antibody	2-25 Ab	500	Lac
Allium moly	AMA	1000	Man
Calystegia Sepium	Calsepa	500	Man
Homarus Americanus	HAA	1000	Gal
Marasmium oreades	MOA	500	Gal
Iris Hybrid	IRA	1000	Gal
Morniga sp.	MNA-G	1000	Gal
Bacteria	GafD	500	GlcNAc
Bacteria	PA-IL	500	Gal
Bacteria	PA-IIL	500	Fuc
Bacteria	RS-IIL	500	Man
Galectin-1	Galectin-1	500	Lac
antibody	BRIC 111 Ab	500	Lac

Table S2. **Lectin microarray print conditions (Expanded NCI-60 dataset).** Abbreviations: Gal: Galactose, Fuc: Fucose, Man: Mannose, Lac: Lactose, GalNAc: *N*-acetylgalactosamine, GlcNAc: *N*-acetylglucosamine.

Lectin source	Lectin abbreviation	Print concentration (µg/ml)	Print monosaccharide
Abrus precatorius	APA	500	Gal
Anguilla anguilla	AAA	1000	Fuc
Bryonia dioica	BDA	500	Gal
Canavalia ensiformis	Con A	500	Man
Cancer antennarius	CCA	500	Lac
Caragana arborescens	CAA	500	Gal
Cicer arietinum	CPA	1000	Lac
Colchicum autumnale	CA	500	Gal
Cystisus scoparius	CSA	500	Gal
Datura stramonium	DSA	500	Lac
Dolichos biflorus	DBA	500	Gal
Eunonymus europaeus	EEL	500	Lac
Galanthus nivalis	GNA	1000	Man
Helix pomatia	HPA	500	Gal
Salvia sclarea	SSA	2000	Lac
Limax flavus	LFA	500	Lac
Limulus polphemus	LPA	500	Lac
Maackia amurensis	MAA	500	Lac
Narcissus pseudonarcissus	NPA	1000	Man
Persea americana	PAA	500	GlcNAc
Phaseolus lunatus	LBA	1000	Gal
Psophocarpus tetragonolobus	PTA	500	Gal
Sambucus nigra	SNA	500	Gal
Sambucus nigra	SNA-II	500	Gal
Trichosanthes kirilowii	TKA	500	Gal
Tulipa sp.	TL	1000	GlcNAc
Ulex europaaeus II	UEA-II	1000	GlcNAc
Uritica dioica	UDA	1000	GlcNAc
Vicia graminea	VGA	500	Gal
Vicia villosa	VVA (mann)	500	Gal
Vigna radiata	VRA	500	Gal
Vicia faba	VFA	500	Gal
Black bean	Blackbean	1000	Lac
Ricin B chain	RCA	1000	Lac

Maackia amurensis-I	MAL- I	1000	Lac
Maackia amurensis-II	MAL-II	500	Lac
Psophocarpus tetragonolobus	PTL-I	1000	Gal
Psophocarpus tetragonolobus	PTL-II	1000	Gal
Cyanovirin	CVN	1000	Man
Scytovirin	SVN	500	Man
Robinia pseudoacacia	RPA	500	GlcNAc
Cholera Toxin from Vibrio cholerae	Cholera Toxin	2000	Lac
Amaranthus Caudatus	ACL	500	Lac
Allium sativum	ASA	1000	Man
Marasmium oreades	MOA	500	Gal
Maclura pomifera	MPA	500	Gal
Trichosanthes japonica	TJA-I	1000	Lac
Trichosanthes japonica	TJA-II	500	Lac
Allium moly	AMA	1000	Man
Calystegia Sepium	Calsepa	500	Man
Homarus Americanus	HAA	1000	Gal
Iris Hybrid	IRA	1000	Gal
Morniga sp.	MNA-G	1000	Gal
Aleuria aurantia	AAL	500	Fuc
Amaranthus Caudatus	ACL	500	Lac
Bauhinia purpurea	BPA	500	Gal
Dolichus biflorus	DBA	500	Gal
Erythrina cristagalli	ECA	500	Gal
Griffonia simplicifoia	GS-I	500	Gal
Griffonia simplicifoia	GS-II	500	GlcNAc
Hippeastrum hybrid	HHL	500	Man
Artocarpus integrifolia	AIA	500	Gal
Lens culinaris	LCA	500	Man
Lycopersicon esculentum	LEA	500	GlcNAc
Lotus tetragonolobus	Lotus (LTL)	500	Fuc
Phaseolus vulgaris erythroagglutinin	PHA-E	500	Lac
Phaseolus vulgaris leucoagglutinin	PHA-L	500	Gal
Pisum sativum	PSL	500	Man
Arachis hyogaea	PNA	500	Gal
Glycine max	SBA	500	Gal
Sophora japonica	SJA	500	Gal
Solanus tuberosum	STA	500	GlcNAc
Ulex europaaeus I	UEA-I	500	Fuc
Vicia villosa	VVA	500	Gal
Wisteria floribunda	WFA	500	Gal

Triticum vulgare	WGA	500	GlcNAc
Euonymus europaeus	EEL	500	Lac
Ricinus Communis	RCA	500	Gal
Iberis amara	IAA	500	GalNAc
Mangifera indica	MIA	500	GalNAc
Salvia horminum	SHA	500	GalNAc
Phytolacca americana	PWM	500	Lac
Agaricus bisporus	ABA	500	Gal
Aegopodium podagraria	APP	500	GalNAc
Bacteria	GafD	500	GlcNAc
Bacteria	PA-IL	500	Gal
Bacteria	RS-IIL	500	Man
Bacteria	PA-IIL	500	Fuc
antibody	7-LE Ab	500	Lac
antibody	2-25 Ab	500	Lac
antibody	STn 219	0.0025	Lac

Table S3. Primers, microRNA mimics and inhibitors and constructs used in this study. Construct sequences were confirmed by standard sequencing (Genewiz). $^{¥}$*FUCA2* 3'-UTR sequence is shown below.

	Gene/construct	Direction or Catalog #	Primer sequence
mRNA qPCR	GAPDH	Fwrd	5'-GAAGGTGAAGGTCGGAGTC-3'
		Rev	5'-GAAGATGGTGATGGGATTTC-3'
	MAN1A2	Fwrd	5'-CCTAGATGGGCAAAGATGGATT-3'
		Rev	5'-TGTGTTAAAGGCAGGAAGGAGTT-3'
	MAN1A1	Fwrd	5'-TACGCTTTGTTGGTGGACTACTCT-3'
		Rev	5'-GTTCCTTCCAATACCATTTTTCAT-3'
	MAN1C1	Fwrd	5'-CAGCAGCATCTTGGCGGAGTTTG-3'
		Rev	5'-CTGGGCTGAGGAAGTTGGGGTAGA-3'
	FUCA2	Fwrd	5'-GAGGTTCTGTGGTCGGATGG-3'
		Rev	5'-GCCACCATGCTTACAGATGC-3'
microRNA qPCR	miR-30c	Fwrd	5'-GTGTAAACATCCTACACTCTC-3'
	miR-181a	Fwrd	5'-AACATTCAACGCTGTCGGTGA-3'
	miR-181b	Fwrd	5'-AACATTCATTGCTGTCGGTGG-3'
	miR-361	Fwrd	5'-GGGGGTTATCAGAATCTCCA-3'
	miR-16	Fwrd	5'-GCAGCAGCACGTAAATATTGGCGTA-3'

Luciferase Constructs	pLightSwitch-MAN1A2 3'UTR		Purchased from SwitchGear Genomics
	pLightSwitch-MAN1A2 3'UTR 30c Mutant	Fwrd	5'-ACCTGTGTTTCGCTCATATGGACCACT-3'
		Rev	5'-GAGAATGGTCCTTCTGGAACTGCTT-3'
	pLightSwitch-MAN1A2 3'UTR 361-5p Mutant	Fwrd	5'-CAAAACATGCTATAACATGTGTAGGA-3'
		Rev	5'-GTCAGGTATAAAATGGAATAATTTCAAAG-3'
	pLightSwitch-FUCA2 3'UTR¥	Fwrd	5'-AACTCGAGGGCTCTAGCCCTGACTAA-3'
		Rev	5'-AAGCTAGCTTCCAGTCCTGTAAGCTC-3'
	pLightSwitch-FUCA2 3'UTR 200b/200c/429 Mutant	Fwrd	5'-CATTGCTAGTAAAGTCTTTTTTGTGCC-3'
		Rev	5'-CATTTCCCAGAAAATATCCCCATATT-3'
	pLightSwitch-ST6GALNAC3-3'UTR		Purchased from SwitchGear Genomics
	pLightSwitch-ST6GALNAC3 3'UTR -200b-5p Mutant	Fwrd	5'-GCTACTTAACAT GTAAGAT ACCAACTTCAAC-3'
		Rev	5'-ATTTTTGGTTGAGTGCTAAATATTTCAGAAG-3'
			mature microRNA sequence
miRIDIAN microRNA Mimics	Negative Control #1 (Scramble)	CN-001000-01-05	
	Positive Control (GAPDH)	CP-001000-02-05	
	hsa-miR-30c	C-300542-05-0005	UGUAAACAUCCUACACUCUCAGC
	hsa-miR-181a-5p	C-300552-05-0005	AACAUUCAACGCUGUCGGUGAGU

	hsa-miR-181b-5p	C-300554-07-0005	AACAUUCAUUGCUGUCGGUGGGU
	hsa-miR-361-5p	C-300663-05-0005	UUAUCAGAAUCUCCAGGGGUAC
	hsa-miR-200b-3p	C-300582-07-0005	UAAUACUGCCUGGUAAUGAUGA
	hsa-miR-200c-3p	C-300646-05-0005	UAAUACUGCCGGGUAAUGAUGGA
	hsa-miR-429	C-300722-05-0005	UAAUACUGUCUGGUAAAACCGU
	hsa-miR-200b-5p	C-301144-01-0005	CAUCUUACUGGGCAGCAUUGGA
miRIDIAN microRNA Hairpin Inhibitors	Negative Control#1 (Scramble)	IN-001005-01-05	
	Positive Control (GAPDH)	IP-004000-01-05	
	hsa-anti-miR-30c	IH-300542-07-0005	UGUAAACAUCCUACACUCUCAGC
	hsa-anti-miR-181a-5p	IH-300552-07-0005	AACAUUCAACGCUGUCGGUGAGU
	hsa-anti-miR-181b-5p	IH-300554-08-0005	AACAUUCAUUGCUGUCGGUGGGU
	hsa-anti-miR-361-5p	IH-300663-07-0005	UUAUCAGAAUCUCCAGGGGUAC

Supplementary Methods and Sequences:

RT-qPCR conditions for mRNA amplification:
Step one-95 for 10 min
Step two and three-95 for 15 sec; 60 for 1 min *40 cycles
Step four- 95 for 15 sec
Step five- 60 for 15 sec to 95 for 15 sec (RAMP 20 min)
Step six- 95 for 15 sec
Step seven- 4 for 1 min

Sequence of FUCA2 3'-UTR cloned from HT-29:
5'-
GGCTCTAGCCCTGACTAATGTGATCTAAAGTGCAGCAGAGTGGCTGATGCTGCAAGTTATGTCTAAGGCTAGGAACT
ATCAGGTGTCTATAATTGTAGCACATGGAGAAAGCAAATGTAAAACTGGATAAGAAAATTATTTTGGCAGTTCAGCC
CTTTCCCTTTTTTCCCACTAAATTTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTA
AAGTCTCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTCACATTATAGTAGCAAGGAATTG
GTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATGATTATATAGTTATGCATCACTTAAT

ATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAATTTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATC
CTAGATGGCATAGCCTACTACACACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCAT
GTTACTGAATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAGGTACAGTA
AAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAGCTTACAGGACTGGAA-3'