

Supplementary Table 2.

Selected clusters with structural similarities. Cluster identifiers, number of glycans within each cluster, core structure of cluster, glycan length (number of monosaccharides in cluster) are listed for structurally similar glycan clusters. Glycan architecture occurs as a linear tree (-), branched tree (+) or linear/branched together (*). Localization of core carbohydrate substructures describes where the common sub-structure occurs; at core (directly after the spacer), at terminal part, or varying location. Pearson correlation coefficients for glycans are summarized by within-cluster range of values. Expected natural glycan linkage is described as Glycosphingolipid (GSL); *N*-linked (N); *O*-linked glycan (O).

Cluster ID	Number of glycans	Key Carbohydrate Substructure	Range of glycan length	architecture	Localization of key carbohydrate substructure	Range of pairwise correlation	Expected natural linkage
1	6	Gal β 1-3GlcNAc β 1-3Gal β 1-4Glc β	4-6	*	core	0.78-0.98	GSL
9	8	[Gal β 1-4GlcNAc β] _n	2-6	-	varying	0.48-0.96	N
10	3	[GlcNAc β 1-4] _n	3-6	-	varying	0.91-0.96	N
15	3	3'-O-Su-Gal β 1-4	2	-	terminal	0.69-0.84	GSL, N
16	5	Gal β 1-4Glc or GlcNAc β	2	-	core	0.15- 0.62	GSL, N
17	4	6'-O-Su-Gal β 1-4GlcNAc	2	-	varying	0.43-0.81	N

26	4	4'-O-Su-Galβ1-4GlcNAc	2	-	varying	0.78-0.98	N
28	2	[Glcα-] _n	3,4	-	varying	0.94	GSL
30	3	Galα1-3(Fuca1-2)Galβ	3-4	-	varying	0.87-0.97	GSL, N
31	3	GalNAcα1-3(Fuca1-2)Galβ	3-4	-	varying	0.59-0.92	GSL, N
32	2	Neu5Gcβ or Neu5Acβ 2-6 Galβ1-4GlcNAcβ	3	-	core	0.87	N
33	4	Galβ1-3GalNAcα-	2-4	*	varying	0.66-0.86	O
34	8	Fuca1-3, Fucβ1-3 or Galβ1-4GlcNAcβ	1-4	*	core	0.44-0.90	N
35	2	Neu5Acβ2-6GalNAcα	2,4	-	varying	0.67	O
36	2	Galα1-4Galβ1-4GlcNAcβ	3,4	*	varying	0.64	N
37	2	Galβ1-4GlcNAcβ1-3GalNAcα	3,5	*	varying	0.82	O
38	4	(6,6'Su ₂ ; Neu5Acα2-6; Galα1-3; Galα1-4) to Galβ1-4GlcNAcβ	2-4	*	varying	0.25-0.63	N
40	3	Fuca/Neu5Acα1/2-2/3/4-Galβ1-3GlcNAcβ	3	*	core	0.46-0.64	N
42	2	Neu5Acα2-3Galβ1-4GlcNAcβ	3	-	varying	0.70	N
44	2	GalNAcα	4	+	varying	0.96	N, O
47	2	GlcNAcβ1-6Galβ1-3GlcNAcβ	3,4	-	varying	0.69	N