Supporting Information

Mechanistic insights into carbohydrate distribution on cell membranes revealed by dSTORM imaging

*Junling Chen,* [a] [b] *Jing Gao,* [a] *Mingjun Cai,* [a] *Haijiao Xu,* [a] *Junguang Jiang,* [a] *Zhiyuan Tian,* [c] and *Hongda Wang* * [a]

[a] State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, Jilin 130022, P.R. China.

[b] University of Chinese Academy of Sciences, Beijing 100049, P.R. China.

[c] School of Chemistry and Chemical Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

* Email: hdwang@ciac.ac.cn
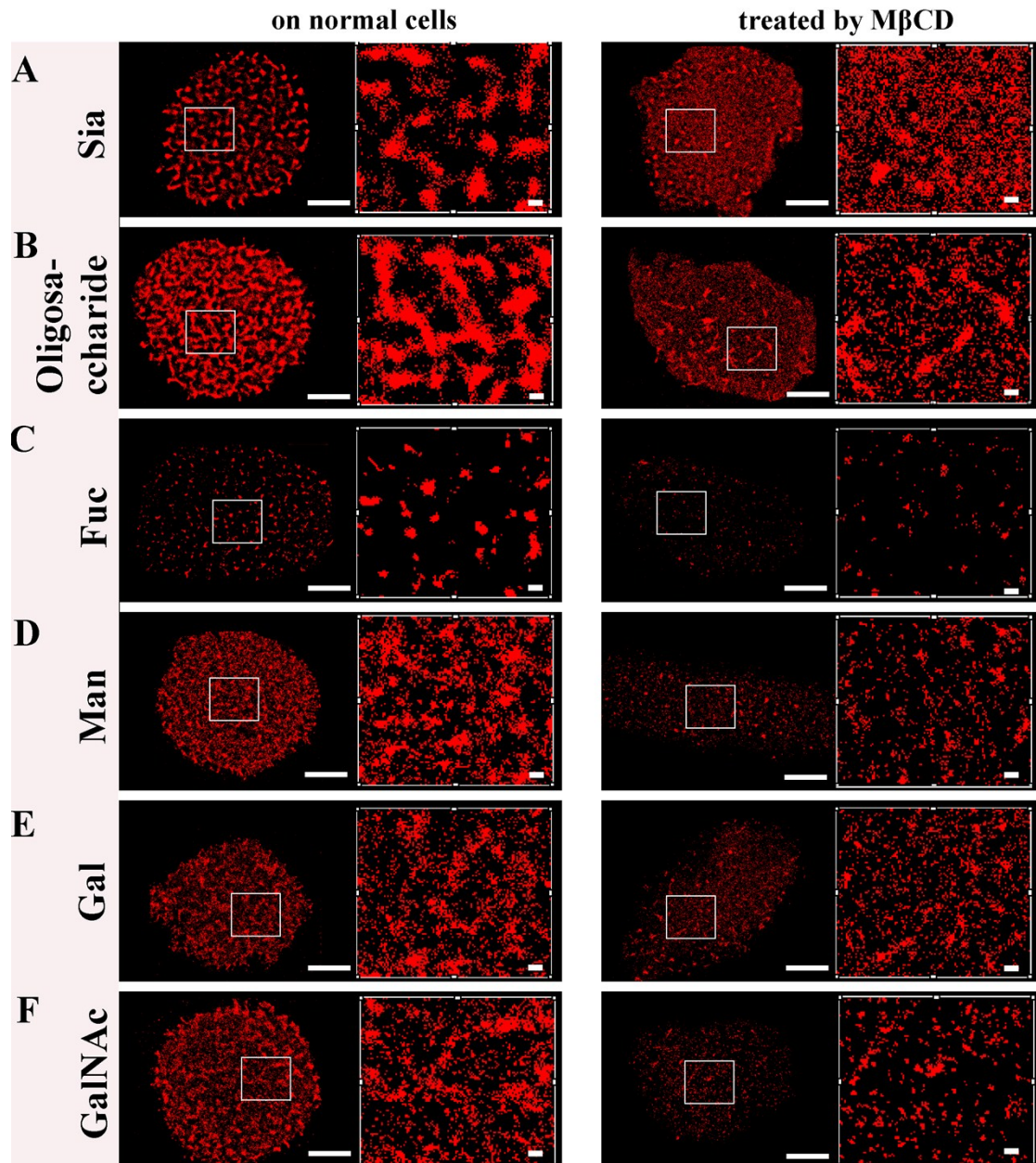
Figure S1. Systemic imaging comparison of the distribution of carbohydrates between normal cell membranes and membranes treated with MβCD. (A-F) dSTORM imaging of all carbohydrates on the treated membranes (left) and normal cell membranes (right). Scale bars are 5 μm. The enlarged images more clearly display the morphological differences of carbohydrates. Scale bars are 500 nm.
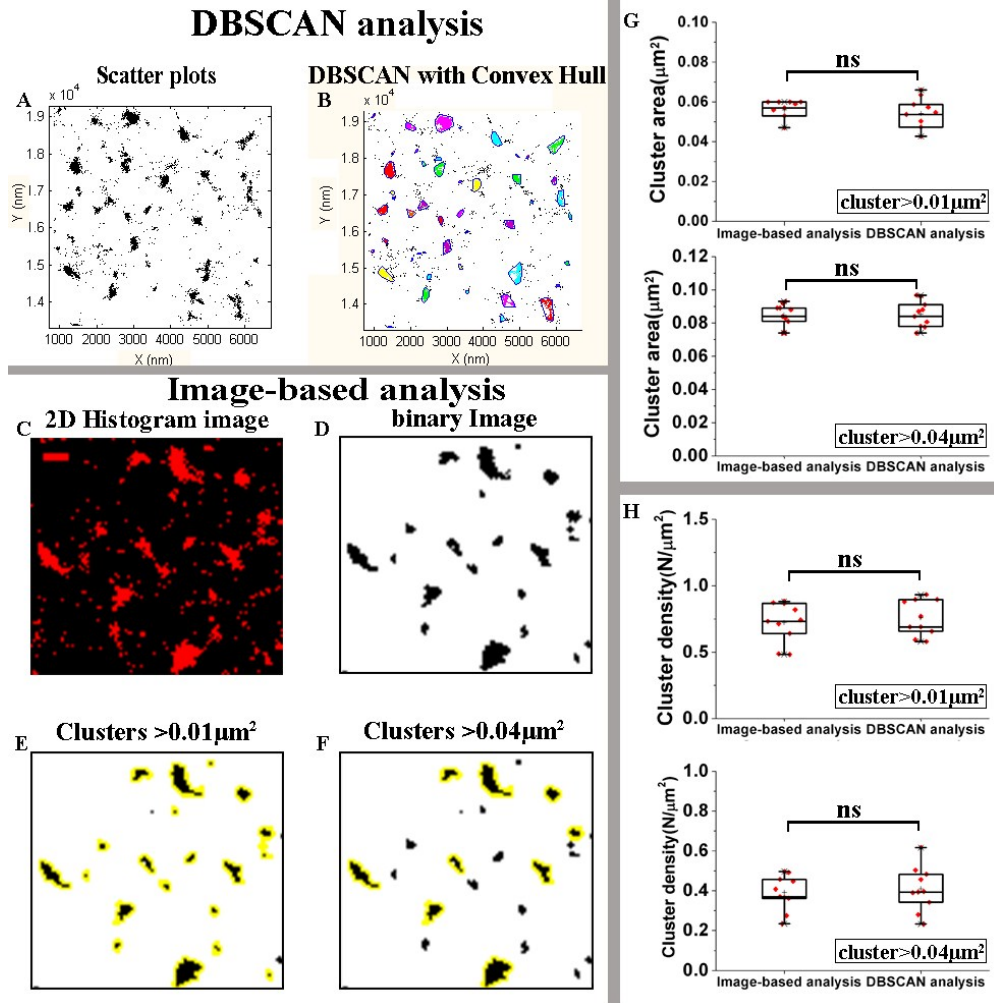
Figure S2. Fuc clusters were characterized by DBSCAN analysis and image-based analysis. (A) dSTORM image of Fuc generated as scatter plots in MatLab. (B) The qualified clusters were identified by DBSCAN analysis. **(**C) The original reconstruction dSTORM image in the view mode of 2D histogram. Scale bar is 2 μm. (D) The binary image after "*Remove outliers*". (E and F) The images of quantified clusters (outlined in yellow). (G and H) Box plots showing compared median average cluster area (G) and cluster density (H) from these two analysis methods, with setting threshold of cluster size is 0.01 μm² (in upper image) and 0.04 μm² (in lower image). The boxes represent the Interquartile Range (IQR, 25th to 75th percentile of the dataset). The lines inside box represent the median values. The whiskers represent 1 to 99% of the data. '+' represents the mean value. All statistical analyses were acquired from ten cells in three independent experiments. Statistical comparisons were performed by the Mann Whitney test, 'ns' means no significance.
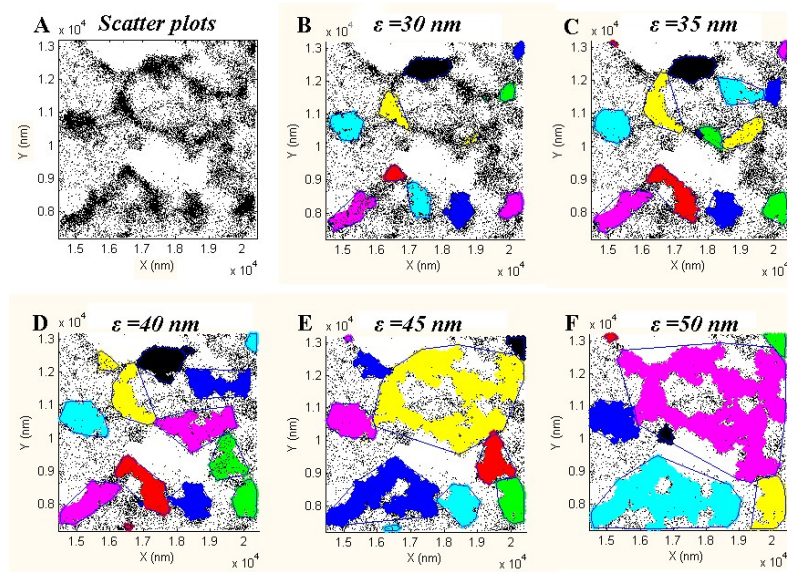
Figure S3 (A) dSTORM image of the distribution of oligosaccharide on Vero apical membrane in the view mode of scatter plots. (B-H) The images of quantified clusters with convex hull identified by DBSCAN analysis with setting different search radius (ε), including ε=30 nm (B), ε=35 nm (C), ε=40 nm (D), ε=45 nm (E), ε=50 nm (F).
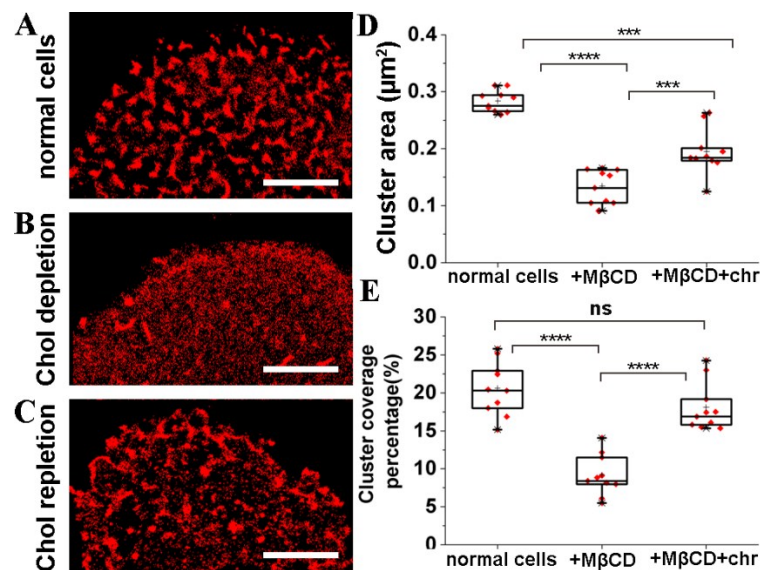

Figure S4. Effects of cholesterol depletion and repletion on the distribution of Sia. (A-C) The reconstructed dSTORM images of Sias on normal cell membrane (A), the membrane with depletion of cholesterol (B) and the membrane with repletion of cholesterol (C). (D and E) Box plots depicting the median cluster area (D) and cluster coverage percentage (E) of Sia clusters on normal membrane and membranes with depletion and repletion of cholesterol. The boxes represent the Interquartile Range (IQR, 25th to 75th percentile of the dataset). The whiskers represent 1 to 99% of the data. '+' represents the mean value. All statistical analyses were acquired from ten cells in three independent experiments. Statistical comparisons were performed by the Mann Whitney test. "***" means P<0.001, "****" means P<0.0001, 'ns' means no significance.
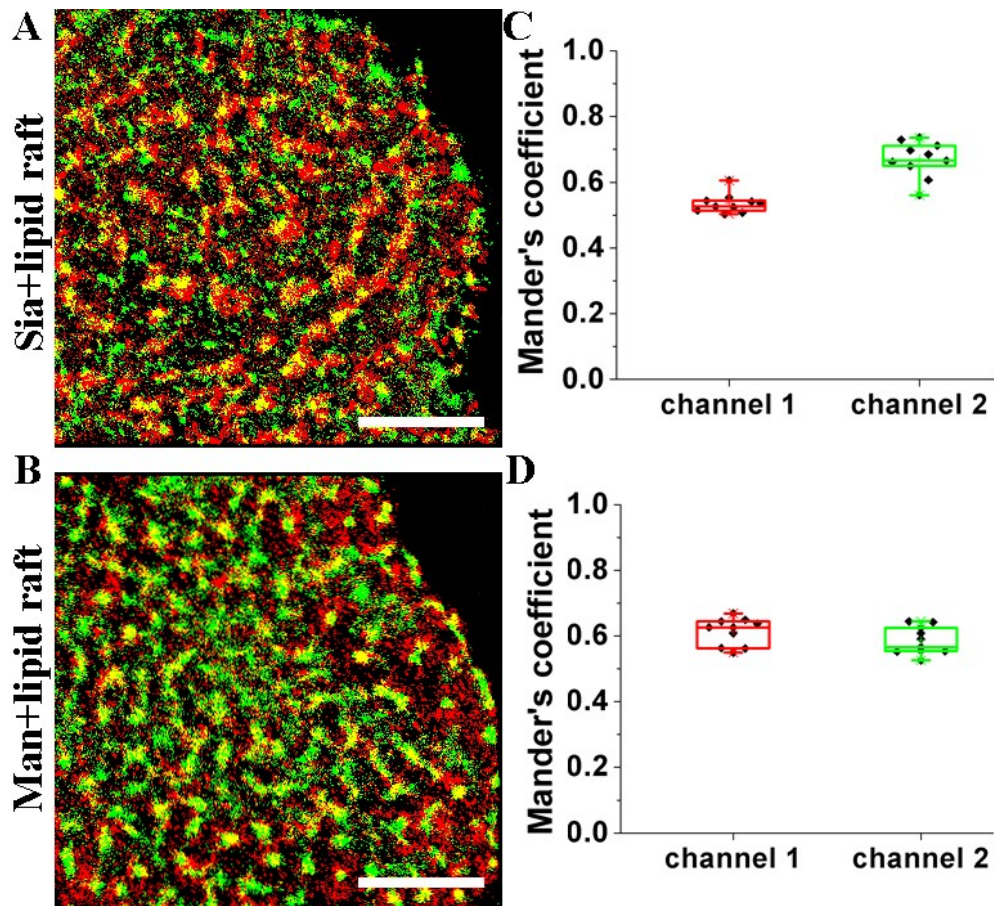
Figure S5. Colocalized distribution of carbohydrates and lipid rafts on Vero membranes. (A) The merged image of Sias (red color) and lipid rafts (green color) acquired from dual-color dSTORM imaging. (B) The merged image of Mans (red color) and lipid rafts (green color). (C and D) Box plots showing the median Mander's coefficient characterize the colocalization of Sia (C) or Man (D) to lipid raft. The boxes represent the Iinterquartile Range (IQR, 25th to 75th percentile of the dataset). The lines inside box represent the median values. The whiskers represent 1 to 99% of the data. '+' represents the mean value. All statistical analyses were acquired from ten cells in three independent experiments.
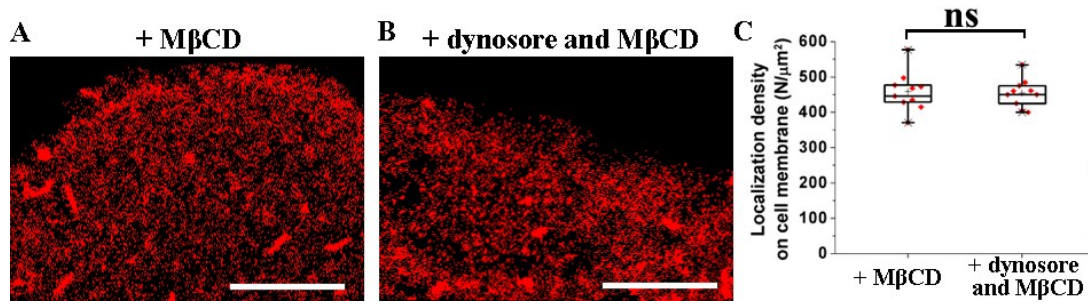
Figure S6. Effects of MβCD treatment on localization density of Sia on cell membrane in the presence of dynosore or not. (A and B) The reconstructed dSTORM images of Sias on the membrane with only MβCD treatment (A) and with both MβCD and dynasore treatment (B). Scale bars are 5 μm. (C) Box plot of localization density on these two classes of treated membranes. The boxes represent the Interquartile Range (IQR, 25th to 75th percentile of the dataset). The lines inside box represent the median values. The whiskers represent 1 to 99% of the data. '+' represents the mean value. All statistical analyses were acquired from ten cells in three independent experiments. Statistical comparisons were performed by the Mann Whitney test, 'ns' means no significance.
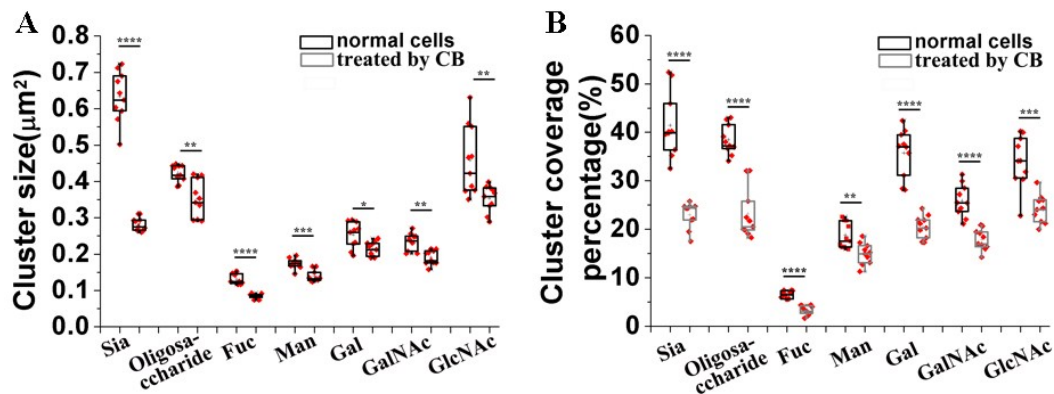
Figure S7. Quantified comparisons of average cluster area (A) and cluster coverage percentage (B) on normal and treated membranes showing the morphological changes caused by treating with CB. The boxes represent the Interquartile Range (IQR, 25th to 75th percentile of the dataset). The lines inside box represent the median values. The whiskers represent 1 to 99% of the data. '+' represents the mean value. All statistical analyses were acquired from ten cells in three independent experiments. Statistical comparisons were performed by the Mann Whitney test. "*" means P<0.05, "**" means P<0.01, "***" means P<0.001, "****" means P<0.0001, and 'ns' means no significance.

on normal cells   treated by lactose
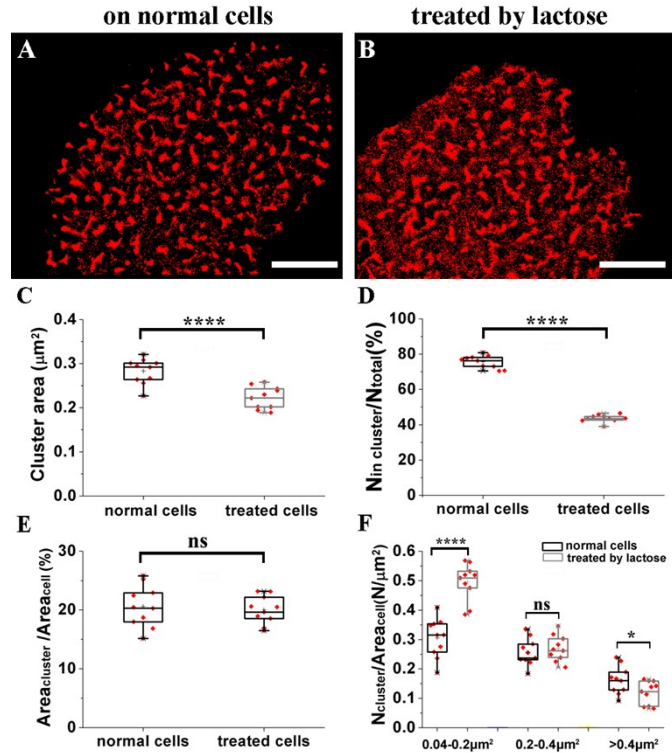
Figure S8. Statistical analyses of the effect of galectins on the organization of Sia on Vero apical membranes. (A-B) Comparative dSTORM images of Sia on normal Vero apical membranes (A) and on membranes treated with lactose (B). Scale bars are 5 μm. (C-G) Box plots of the distribution of average cluster area (C), ratio of the number of localizations in clusters to total localizations on the entire cell membrane (D), percentage of cluster coverage percentage on cell membrane (E), number of clusters belonging to different classes on unit cell membrane (F) on normal cell membranes (black boxes) and on treated membranes (gray boxes). The boxes represent the Interquartile Range (IQR, 25th to 75th percentile of the dataset). The lines inside box represent the median values. The whiskers represent 1 to 99% of the data. '+' represents the mean value. All statistical analyses were acquired from ten cells in three independent experiments. Statistical comparisons were performed by the Mann Whitney test. "*" means P<0.05", ****" means P<0.0001, and 'ns' means no significance.
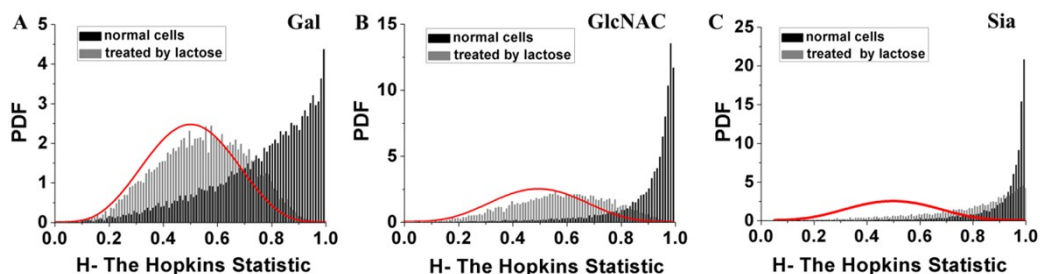

Figure S9. Average results of modified Hopkin's test for spatial randomness on centroids of the examined carbohydrates domains on normal and treated membranes with lactose, including Gal (A), GlcNAc (B) and Sia (C). Red curve of the function showing a random distribution, with centering at the Hopkin's statistic 0.5 on the abscissa.
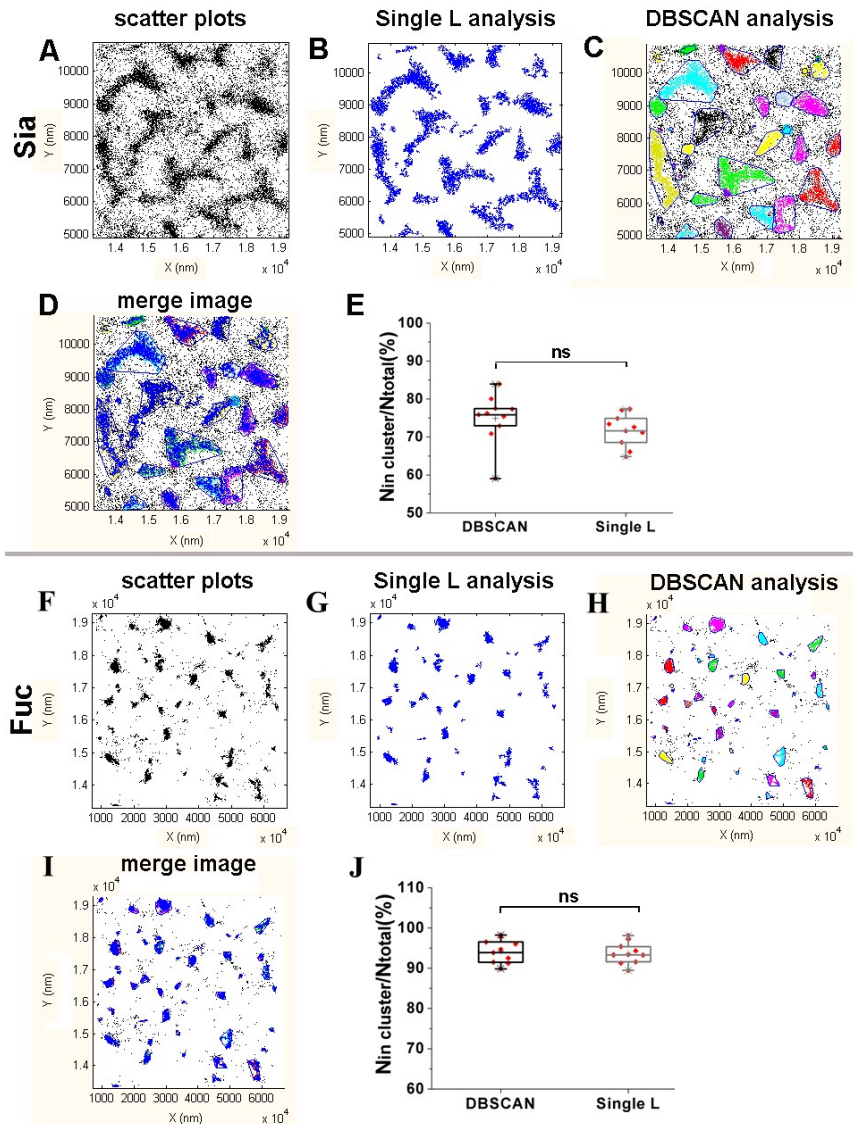
Figure S10. The ratio of localizations in clusters to total localizations on normal cell membrane characterized by two analysis methods. (A) dSTORM image of Sia generated as scatter plots in MatLab. (B) Image of quantified Sia clusters acquired by Single L analysis. (C) Image of quantified Sia clusters acquired by DBSCAN analysis. (D) The merged image of B and C. (E) Box plots depicting the median ratio of localizations in clusters to total localizations on cell membrane acquired from these two analysis methods. (F-J) The corresponding analysis results of Fuc. The boxes represent the Interquartile Range (IQR, 25th to 75th percentile of the dataset). The lines inside box represent the median values. The whiskers represent 1 to 99% of the data. '+' represents the mean value. All statistical analyses were acquired from ten cells in three independent experiments. Statistical comparisons were performed by the Mann Whitney test, "ns" means no significance.
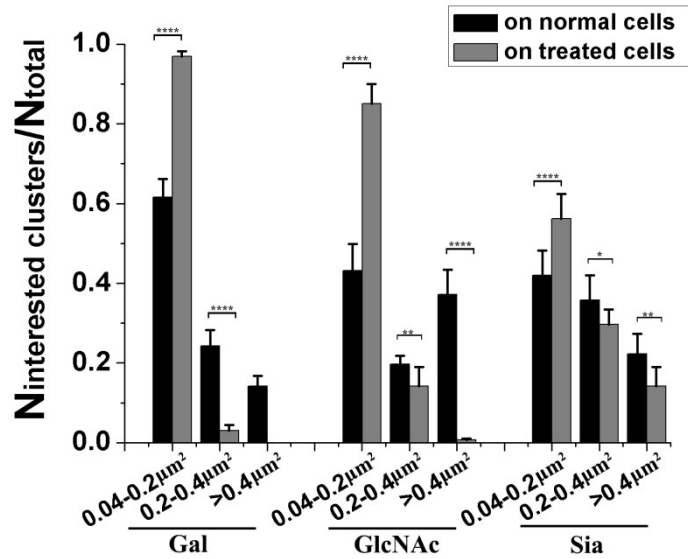
Figure S11. Statistical analyses of the effect of galectins on the organization of Gal, GlcNAc and Sia on Vero apical membranes. The ratio of the cluster number of each group to the total number of clusters on normal cell membranes (black columns) and on treated membranes (gray columns). All statistical analyses are acquired from ten cells in three independent experiments. All results are the means ± S.D. Statistical comparisons were performed by the Mann Whitney test. "*"means P<0.05, "**" means P<0.01, and "****" means P<0.0001.
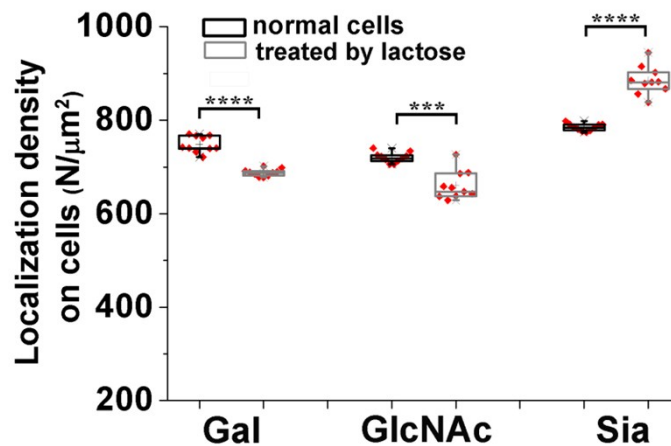
Figure S12. Comparative analyses of the localization density on cell membrane of Gal, GlcNAc and Sia between on normal and the treated membranes by lactose. The boxes represent the Interquartile Range (IQR, 25th to 75th percentile of the dataset). The lines inside box represent the median values. The whiskers represent 1 to 99% of the data. '+' represents the mean value. All statistical analyses were acquired from ten cells in three independent experiments. Statistical comparisons were performed by the Mann Whitney test. "***" means P<0.001, and "****" means P<0.0001.

13. Materials and Methods

13.1 Cell culture

Vero cells (African green monkey kidney cells) were purchased from the Shanghai Institute of Biological Sciences, and were cultured in minimum essential medium (MEM, HyClone) containing 10% fetal bovine serum (FBS, HyClone), 100 μg/mL streptomycin and 100 U/mL penicillin, in a 5% CO_2 environment at 37°C. For dSTORM imaging, the cells were seeded on clean cover slips (22 mm × 22 mm, Fisher) placed in a culture dish with ~1.5 mL culture medium for ~24 hours.

13.2 Sample preparation

13.2.1 Preparation of specific probes for labeling different carbohydrates and membrane proteins

In our research, according to the binding specificity of lectins (information provided by the manufacturer's introduction), several lectins are selected to recognize their specific carbohydrates. Maackia amurensis lectin (MAL, from Sigma) with high affinity of sialic acid linked to galactose by an α2-3 linkage (Sia)[1, 2], lectin from Phaseolus vulgaris (PHA-L, from EY laboratories) binding to oligosaccharide (pentasaccharide sequence Galβ1-4GlcNAcβ1-2(Galβ1-4GlcNAcβ1-6) Manα1-R (the so-called "2,6-branch"))[3, 4], wheat germ agglutinin (WGA, from Sigma) that mainly interacts with high affinity with N-acetyl-D-glucosamine (GlcNAc) and its β-(1→4)-linked oligosaccharides[3, 5]; lectin from Anguilla anguilla (eel) (AAA, from Sigma) for fucose (Fuc)[3, 6], lectin from Morniga M (MNA-M, from EY laboratories) whose binding specificities of mannose (Man)[7], erythrina cristagalli lectin (ECL, from Sigma) for D-Galβ1-4GlcNAc (Gal)[3], and lectin from glycine max (SBA) possessing high affinity for N-acetylgalactosamine (GalNAc)[5, 8]. We linked these lectins to Alexa647 to act as unique probes for imaging the corresponding carbohydrates with a similar protocol in previous study[9]. Simply, the lectin solution reacted with Alexa647 (Molecular Probes) by rocking at 27°C for ~3 hours with gentle vortex. Then, using a PD Spin Trap G-25 filtration column (GE Healthcare) with 1x PBS to remove the excrescent free Alexa647. Finally, according to the Beer-Lambert, the labeling ratio was calculated law by measuring the absorbance of sample at 280 nm (WGA) and 650 nm (maximum absorbance of Alexa647).

Similarly, EGF (from PeproTech) or band 3 monoclonal antibody (BIII 136, from Santa Cruz) was conjugated to Alexa647 to specifically label membrane proteins (EGFR or band 3).

13.2.2 Sample preparation for imaging carbohydrates

In view of successful protocols of sample preparation in the previous imaging[9], we washed the cultured cells with 1× PBS for 3 times, then fixed cells with 4% paraformaldehyde (PFA) at 37°C for 40 min. Next, washing again to stain cells with Alexa647-linked lectin solution (50 μL) at 4°C for 10 min. Then, washing 4 times with1× PBS to remove the excess solution. For imaging, the small coverslip containing cells was gently sealed onto the microscope slide (24 mm × 50 mm, Fisher) where imaging buffer (containing 140 mM beta-mercaptoethanol (βME), 0.5 mg/mL of glucose oxidase and 40 μg/mL of catalase) dropped.

13.2.3 Sample preparation for dual-color super-resolution imaging

After cell being fixed, samples were staining with 50 μL lipid raft markers (10 μg/mL, cholera toxin subunit B (recombinant)-Alexa Fluor 555 conjugates, Molecular Probes) for 30 min, then after washing, cells were labeled with Alexa647-linked lectins as described above.

13.2.4 Treatment with lactose

After cells had attached to the small glass cover slip, we treated them with growth medium containing 150 mM lactose for a minimum of 24 h to disrupt glycan-dependent interactions mediated by cell surface galectins. The targeted cells are treated by the same protocols of fixation, staining and sealing for dSTORM imaging.

13.2.5 Cholesterol depletion and repletion

For cholesterol depletion, before being fixed, cells cultured on a clean slide were washed 3X with pre-warmed PBS and incubated in 5 mM MβCD (Sigma) for 30 min to extract the cholesterol. Then, fixation, washing, staining and sealing as above description.

For cholesterol repletion, depleted cells were incubated with a solution containing cholesterol/ MβCD complex (5 mM MβCD, at 1:10 molar ratio) for 1 h. The preparation of cholesterol/MβCD solution is similar as previously described procedure[10]. In brief, a small volume of cholesterol stock solution (50 mg/mL in chloroform: methanol (1:1, V:V)) was added to a glass tube and evaporated under a gentle stream of nitrogen. Then, 5 mM MβCD solution was added to the cholesterol. After being vortexed and sonicated, the solution was incubated in a rotating

water bath at 37°C overnight. Then, with the same protocols of fixation, washing, staining and sealing, the samples were prepared well.

13.2.6 Disruption of the actin cytoskeleton

To visualize changes in the organization of carbohydrates upon impairing the actin cytoskeleton, we treated cultured cells with 10 µg/mL Cytochalasin B (CB, from Sigma) for 30 min. Then, the cells were treated with the same procedures of fixation, washing, staining and sealing before imaging.

For inhibition of endocytosis by dynasore as a control, cells were treated with 80 µM dynasore (Sigma-Aldrich) for 30 min to inhibit glycoconjugates endocytosis, before treating with the above procedure of actin cytoskeleton.

13.3. dSTORM imaging

dSTORM imaging was performed on an inverted Nikon Ti-E microscope with an oil-immersion objective (100 ×, 1.49 NA, Nikon, Japan). The sample was excited with a 640 nm laser (single color imaging) with TIRF illumination which can significantly decrease the background noise around single molecules. During dual-color dSTORM imaging, a set of optical elements, including an excitation filter (zet405/488/561/647x, 25 mmR, Chroma), a dichroic mirror (zt405/488/561/647rpc, 25.5 mm × 36 mm × 1 mm, Chroma) and an emission filter, was used for sample illumination and imaging. We imaged lipid rafts with exciting 561 nm laser following carbohydrates with 647 nm laser, which can avoid the bleaching of Alexa647 by 561 nm laser illumination. With an electron multiplying charge coupled device (EMCCD, Photometrics, Cascade II) camera, a imaging sequence of 5000 frames were acquired at a rate of 25 Hz to reconstruct a super-resolution image through the QuickPALM in ImageJ. In the course of imaging, the z-drift was eliminated by a focus lock, and the x-y drift was reduced by stabilizing the sample with two clips and corrected by embedding 100 nm diameter TetraSpeck microspheres (Invitrogen) as fiducial markers.

13.4. Data analysis

13.4.1 Reconstruction of dSTORM image

Firstly, we performed QuickPALM[11] in Image J (developed by National Institutes of Health (NIH)) to analyze a frames sequence for a reconstructed dSTORM image. After background subtraction,

through setting the minimum SNR (commonly 6 or 8) and maximum FWHM (commonly 3), fluorescence peaks were identified in each frame and least-squares fitting was performed with an elliptical Gaussian function to localize the positions of particles. Then, with determining centroid positions of peaks and further rejecting the poor fit and asymmetric PSFs, we obtained a localization data of single fluorescent molecules. Finally, a reconstructed image was generated with different view modes, for example, as 2D histogram for view the morphologies of carbohydrates or proteins or 2D Gaussian rendering of molecular positions (2D particle intensity-8 bit) for Mander's coefficients analysis of colocalization or scatter plots for analysis the number of localizations in the region of interest (ROI).

13.4.2 Cluster analysis

DBSCAN (Density Based Spatial Clustering of Applications with Noise) was a common algorithm used to describe clustering in the dSTORM dataset. We implemented it on our dSTORM data of Fuc and oligosaccharide in SuperCluster[12] (Jan. 2014 release, a custom program written in MatLab) as previously described, by defining search radius for cluster identification (ε) (40-50nm for Fuc, 30 nm-50 nn for oligosaccharide) and minimum localizations/cluster (20 for Fuc and oligosaccharide). Because of large cluster size and irregular shape, oligosaccharide clusters cannot be accurately identify by DBSCAN (Figure S3). Therefore, DBSCAN analysis is not suitable for most types of carbohydrates which mostly distributed in large clusters with irregular shapes.

The method of image-based cluster analysis of carbohydrate cluster was similar as that for GlcNAc[9], seen from Figure S2. In brief, we first applied "Remove Outliers" (the function in ImageJ) to the original reconstructed dSTORM image to acquire a clear image where clusters markedly distribute. Then, after converting the image to binary one, we implemented *"Analyze Particles"* in ImageJ to identify the qualified clusters by setting the threshold (>0.04 μm^2 for analysis of carbohydrates) (sometimes independent clusters need to be delimitated by cutting off slight connectivity (*i.e.,* 4- and 8-connected)). Finally, through analyzing the identified clusters, we can obtained some parameters of each cluster and a summary data, including cluster area, perimeter, circularity, the total number of clusters and the total area of clusters on the examined cell membrane.

The modified Hopkin's test for spatial randomness by comparing the distances between random data points and nanodomain centroids to the actual distances. Supercluster program

was also used to run this analysis, and code is available at the UNM STMC website software page (http://stmc.health.unm.edu)[12]. For a random distribution, the function is shaped like a normal "bell" curve centered at the Hopkin's statistic 0.5 on the abscissa; for clustering, the plot is shifted to the right.

We performed the pairwise cumulative distribution function (CDF) analysis in the Supercluster as previously described[12]. Based on the distribution of pairwise distances for imaging of single Alexa647-linkd lectins on glass under identical imaging condition of cell imaging and the size of ROI (3×3-4×4 μm^2), the minimum and maximum distance were set as 20 nm and 1000 nm to analyze carbohydrates data. During analysis, the pairwise distribution function (PDF) was generated from these pair-wise distances. Through placing the pair-wise distances within 100 equally sized bins over the analysis distance, the pair-wise culumlative distribution function (CDF) was calculated. Then, after normalization to the total number of distances, all of the counts in the CDFs sum to 1.

13.4.3 Calculation of the localization density on the cell membrane

To achieve the localization density on the cell membrane, we imported the x and y coordinates of total localizations acquired from "QuickPALM" analysis in ImageJ to MATLAB and obtained ascatter plots, then we gained the total number of localizations in the ROI or entire cell membrane by removing points that were not localized in the ROI or cell membrane. Finally, we accounting the localization density in the ROI or in entire cell membrane by acquiring the area of the ROI or cell membrane in ImageJ.

13.4.4 Calculation of the ratio of the number of localizations in clusters to the total localizations on cell membrane

The number of localizations in cluster can be acquired as one result from the DBSCAN analysis of carbohydrate clusters, or via "*Single L*" (a custom written code in MatLab). The DBSCAN method is same as above description in cluster analysis. For "*Single L*" method, similarly with the method of calculation of the localization density on cell membranes, after generation of the scatter plots, boxed regions (2×2-3×3 μm) were selected to obtain the total number of localizations in this region. Then, running "*Single L*" to extract clusters without single localizations and obtain total number of localizations in cluster. So that, we can calculated the ratio of the number of localizations in clusters to the total localizations on the cell membrane.

13.4.5 Colocalization analysis

During using the Mander's coefficient to test the colocalization, the reconstructed dSTORM images in two channels (view mode in 2D particle intensity-8 bit) were imported to the Image-Pro Plus (IPP, developed by Media Cybernetics) and measured the mander's coefficients of two channels via "co-localization" of the ROI.

The code of "Single L":

```
function [L,Morethanzero,index2]=singleL(X,r)

%X is data of interest, r is the radius for the function calculation

%L is the L value generated for each point at r. Morethanzero means the x-y coordinates of localizations whose L is larger than zero.

[D,N] = size(X');

A=(max(X(:,1))-min(X(:,1)))*(max(X(:,2))-min(X(:,2)));%A is the area

X2 = sum(X'.^2,area

distance = sqrt(repmat(X2,N,1)+repmat(X2',1,N)-2*X*X');% distance is distance matrix

index= find(distance<r);

delta=zeros(N,N);

delta(index)=1;

L=sqrt(A*(sum(delta)-ones(1,N))/(pi*N))-(r*ones(1,N));

index1=find(L<=0);

index2=find(L>0);

X(index1,:)=[];

Morethanzero=X;
```

References

1. W.-C. Wang and R. Cummings, *Journal of Biological Chemistry*, 1988, **263**, 4576-4585.
2. R. Knibbs, I. Goldstein, R. M. Ratcliffe and N. Shibuya, *Journal of Biological Chemistry*, 1991, **266**, 83-88.
3. D. R. Dorscheid, A. E. Conforti, K. J. Hamann, K. F. Rabe and S. R. White, *The Histochemical journal*, 1999, **31**, 145-151.
4. T. W. Hamelryck, M.-H. Dao-Thi, F. Poortmans, M. J. Chrispeels, L. Wyns and R. Loris, *Journal of Biological Chemistry*, 1996, **271**, 20479-20485.

5. K. Yamamoto, S. Ito, F. Yasukawa, Y. Konami and N. Matsumoto, *Analytical biochemistry*, 2005, **336**, 28-38.

6. A. B. Boraston, V. Notenboom, R. A. J. Warren, D. G. Kilburn, D. R. Rose and G. Davies, *Journal of molecular biology*, 2003, **327**, 659-669.

7. H. Wang, L. Obenauer-Kutner, M. Lin, Y. Huang, M. J. Grace and S. M. Lindsay, *Journal of the American Chemical Society*, 2008, **130**, 8154-8155.

8. P. Altevogt, M. Fogel, R. Cheingsong-Popov, J. Dennis, P. Robinson and V. Schirrmacher, *Cancer research*, 1983, **43**, 5138-5144.

9. J. Chen, J. Gao, J. Wu, M. Zhang, M. Cai, H. Xu, J. Jiang, Z. Tian and H. Wang, *Nanoscale*, 2015, **7**, 3373-3380.

10. A. Christian, M. Haynes, M. Phillips and G. Rothblat, *Journal of lipid research*, 1997, **38**, 2264-2272.

11. R. Henriques, M. Lelek, E. F. Fornasiero, F. Valtorta, C. Zimmer and M. M. Mhlanga, *Nature methods*, 2010, **7**, 339-340.

12. M. S. Itano, M. S. Graus, C. Pehlke, M. J. Wester, P. Liu, K. A. Lidke, N. L. Thompson, K. Jacobson and A. K. Neumann, *Frontiers in physics*, 2014, **2**.