11-29-2011

Using LectinMicroarrays to Identify Regulatory Mechanisms for Mammalian Glycosylation

John F. Rakus
Marshall University, rakus@marshall.edu

NEW YORK UNIVERSITY

Using Lectin Microarrays to Identify Regulatory Mechanisms for Mammalian Glycosylation

John Rakus, Ph.D.

Department of Chemistry
Advisor: Dr. Lara Mahal
November 29, 2011

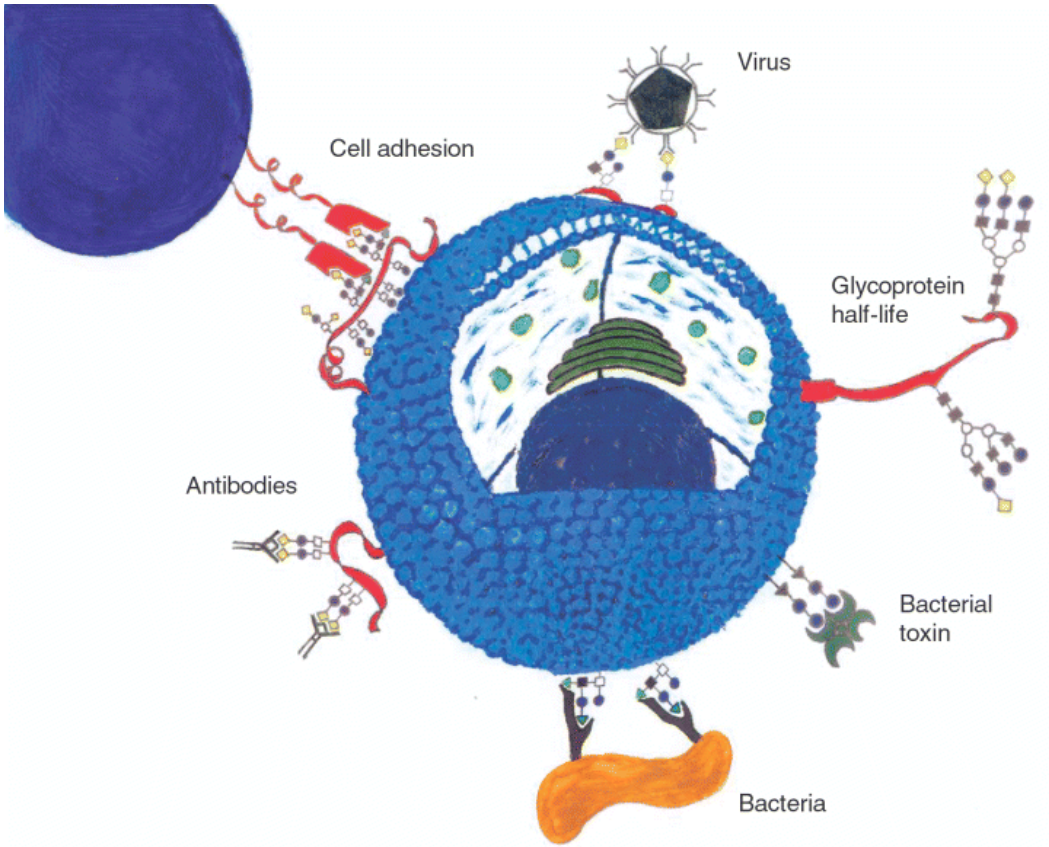
Agenda

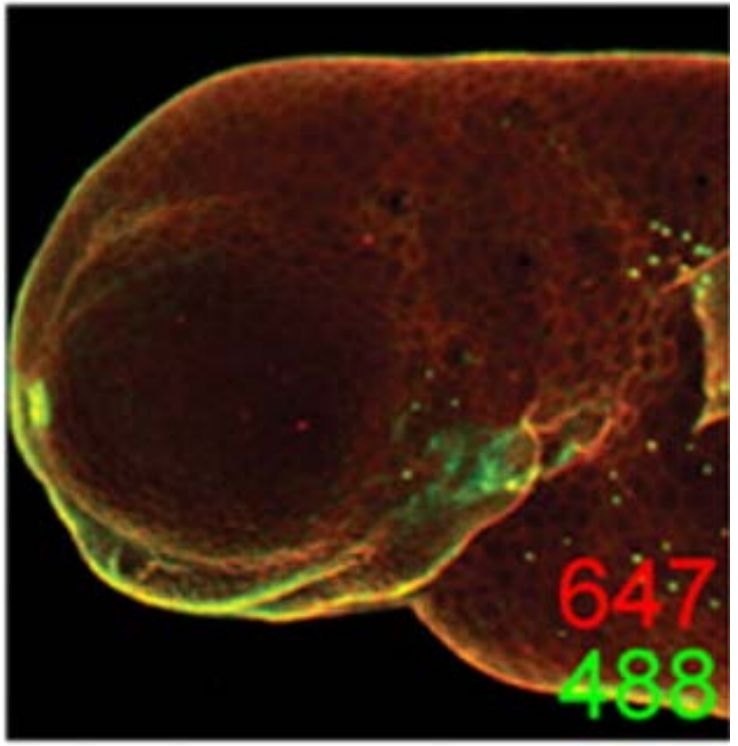
- **Carbohydrates are a diverse and critically important class of biological macromolecules**
- **Microarray analysis of the glycome**
- **Methods to evaluate and integrate microarray data sets**
- **Mechanisms for genetic regulation of the glycome**

Agenda

- **Carbohydrates are a diverse and critically important class of biological macromolecules**
- **Microarray analysis of the glycome**
- **Methods to evaluate and integrate microarray data sets**
- **Mechanisms for genetic regulation of the glycome**

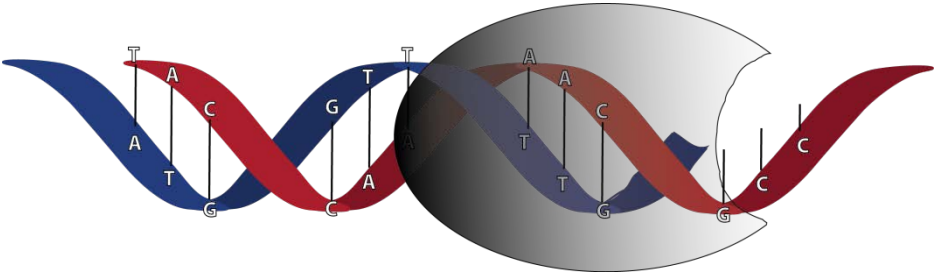
Carbohydrates are pervasive and involved in many cellular interactions

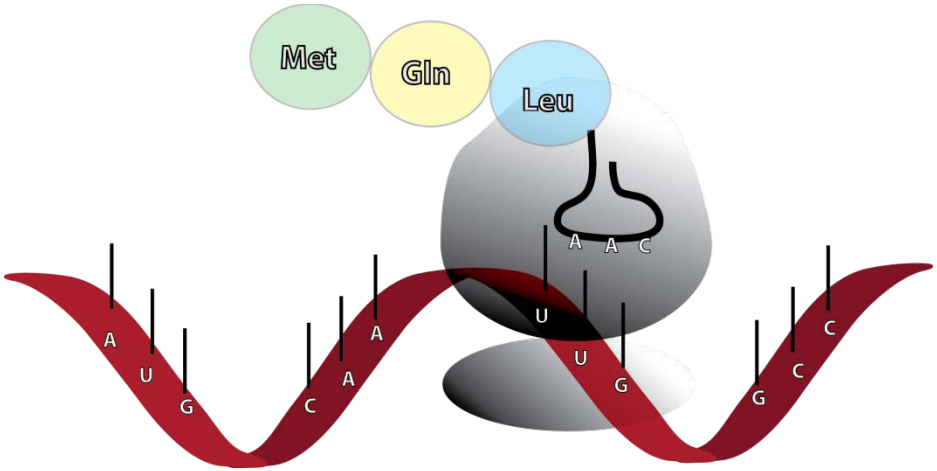




Holgersson et al, *Immuno Cell Biol*, 2005
Baskin et al, *PNAS*, 2007

Nucleic acids and proteins are synthesized with a defined template and dedicated polymerases





Macromolecule: Nucleic acid

Polymerase: DNA Pol or RNA Pol

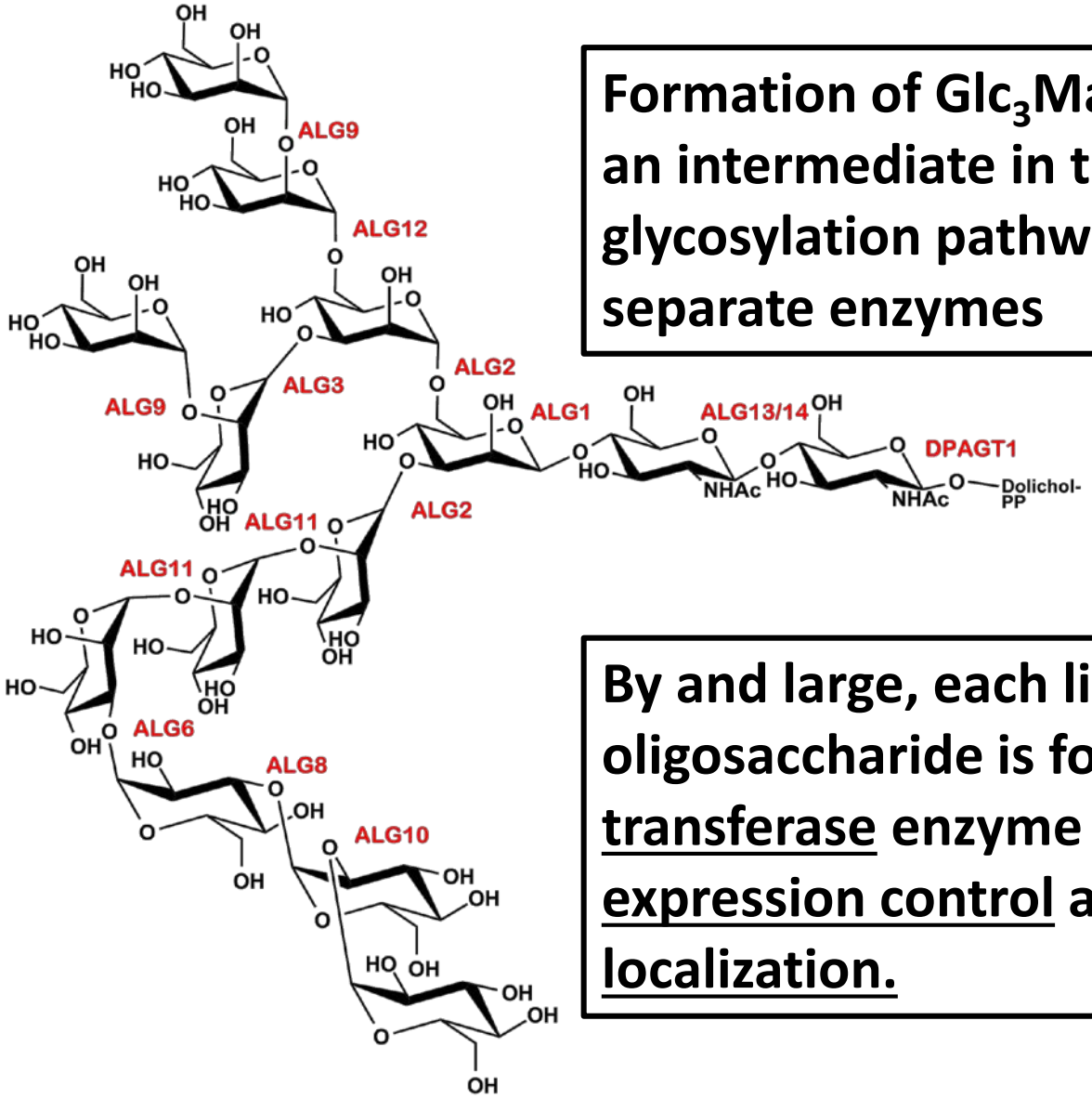
Template: DNA strand

Macromolecule: polypeptide

Polymerase: Ribosome

Template: mRNA strand

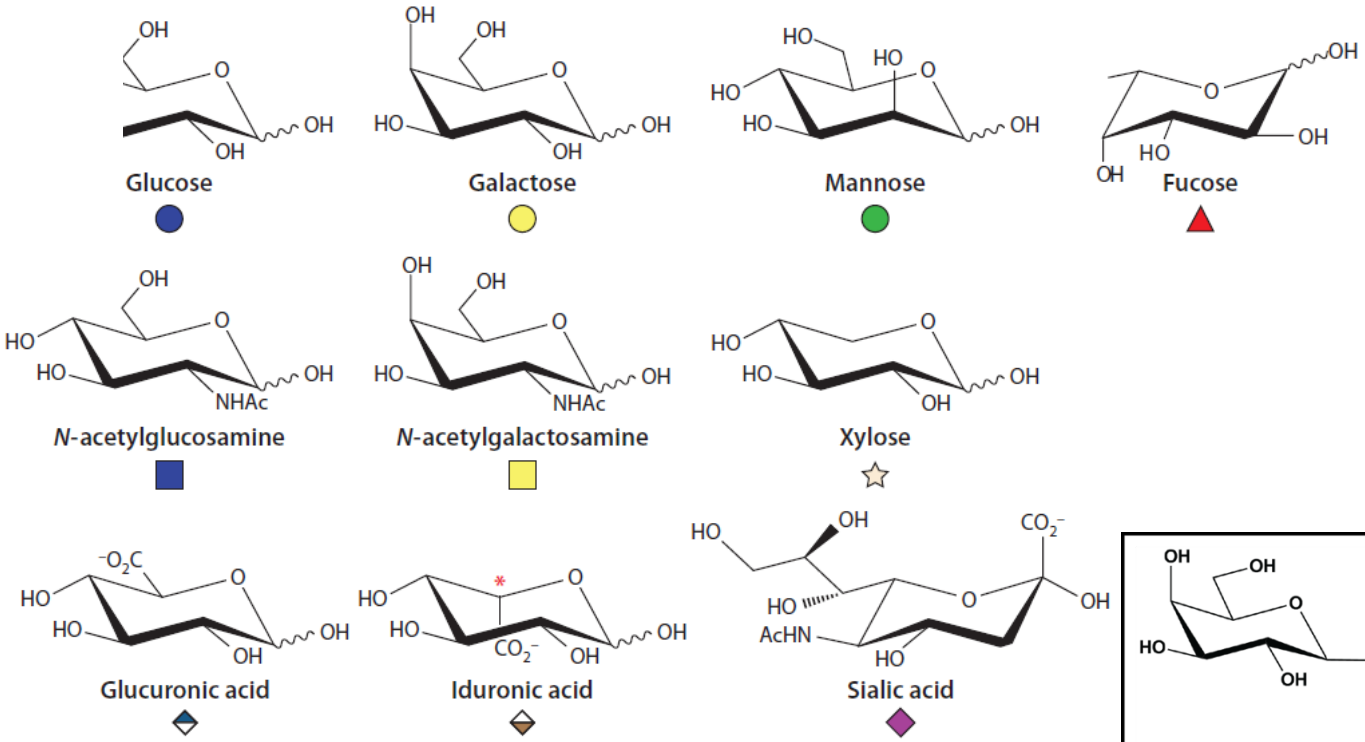
Glycan biosynthesis lacks a dedicated polymerase and genetic template

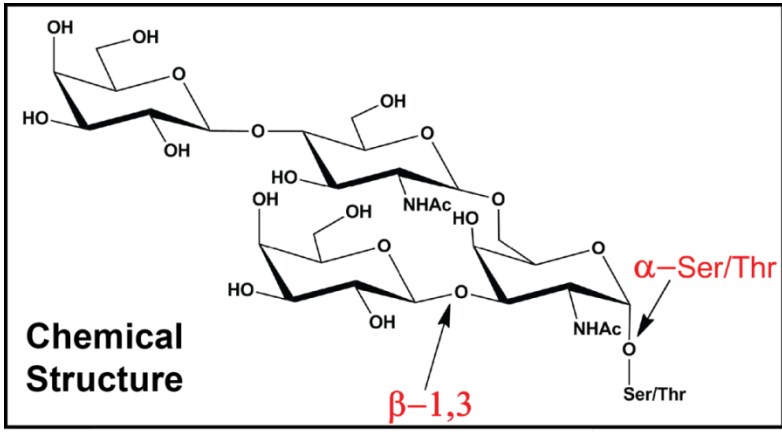


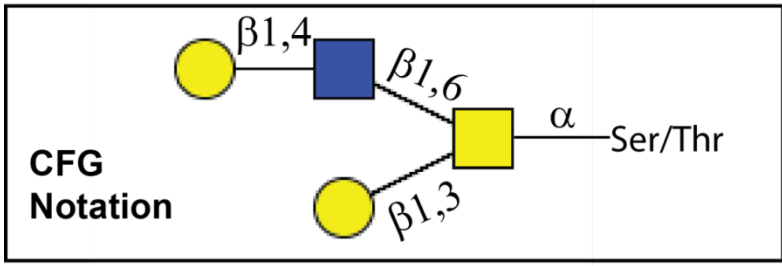
Formation of $Glc_3Man_9GlcNAc_2$-DolPP, an intermediate in the *N*-linked glycosylation pathway, requires 12 separate enzymes

By and large, each linkage in an oligosaccharide is formed by a specific transferase enzyme with specific expression control and subcellular localization.

Consortium for Functional Glycomics (CFG) Notation



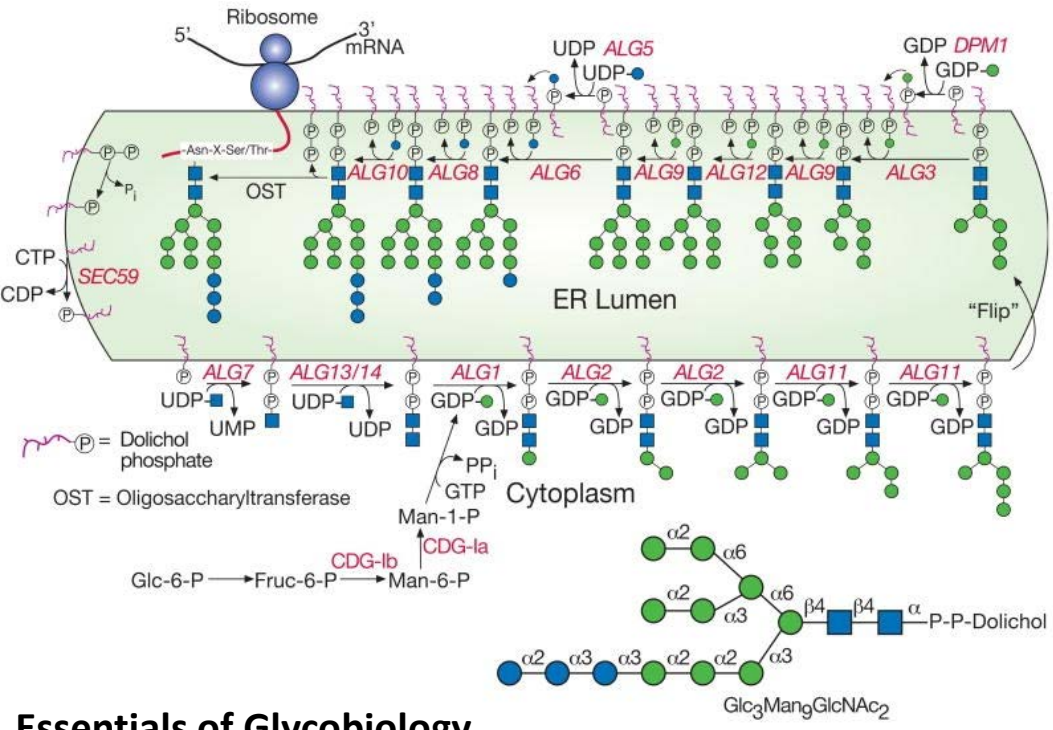




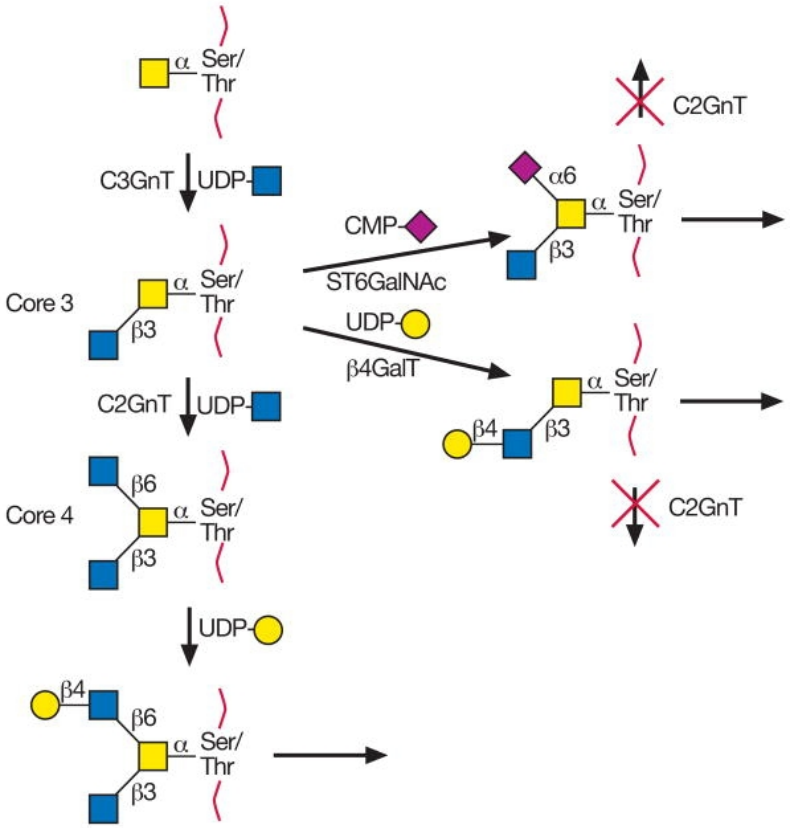
Rakus and Mahal, *Ann Rev Anal Chem*, 2011

There are two primary glycosylation pathways

***N*-linked glycosylation** occurs in the ER and Golgi and involves construction of a lipid-linked 14-mer precursor before being transferred to an Asn residue and further modified to form the final structure. Modified proteins have N-x-S/T consensus sequence



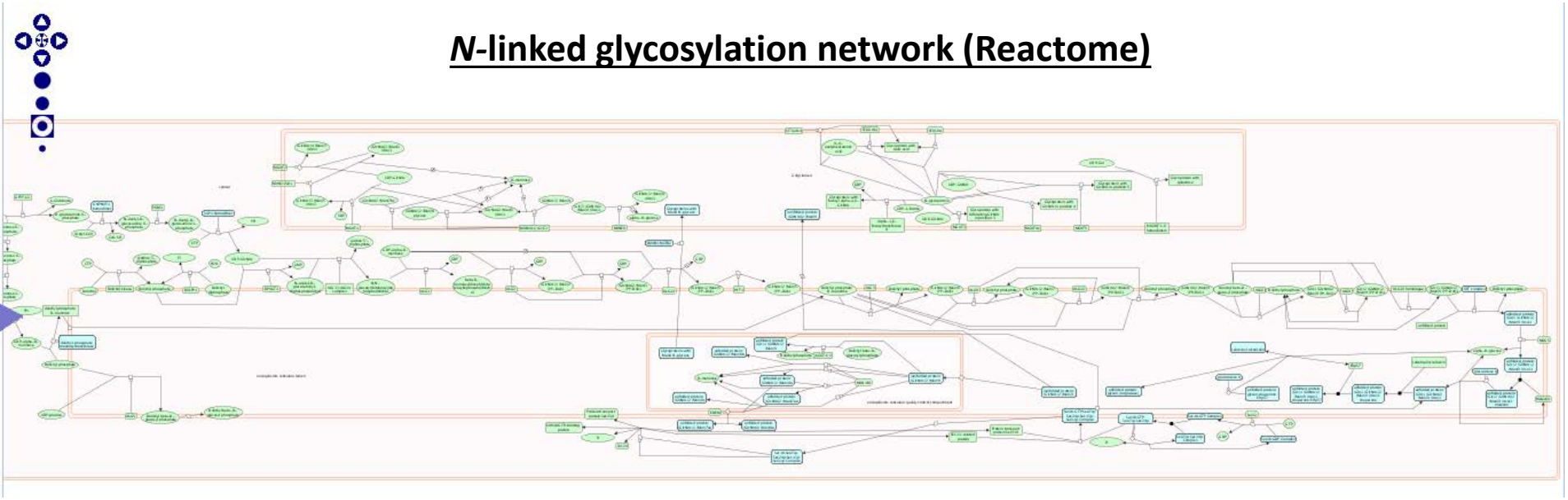
***O*-linked glycosylation** occurs in the Golgi apparatus and involves transfer of a monosaccharide directly to a Ser/Thr residue by a specific ppGalNacT followed by further elaboration. No known consensus sequence



Essentials of Glycobiology

Carbohydrate synthetic regulation

***N*-linked glycosylation network (Reactome)**



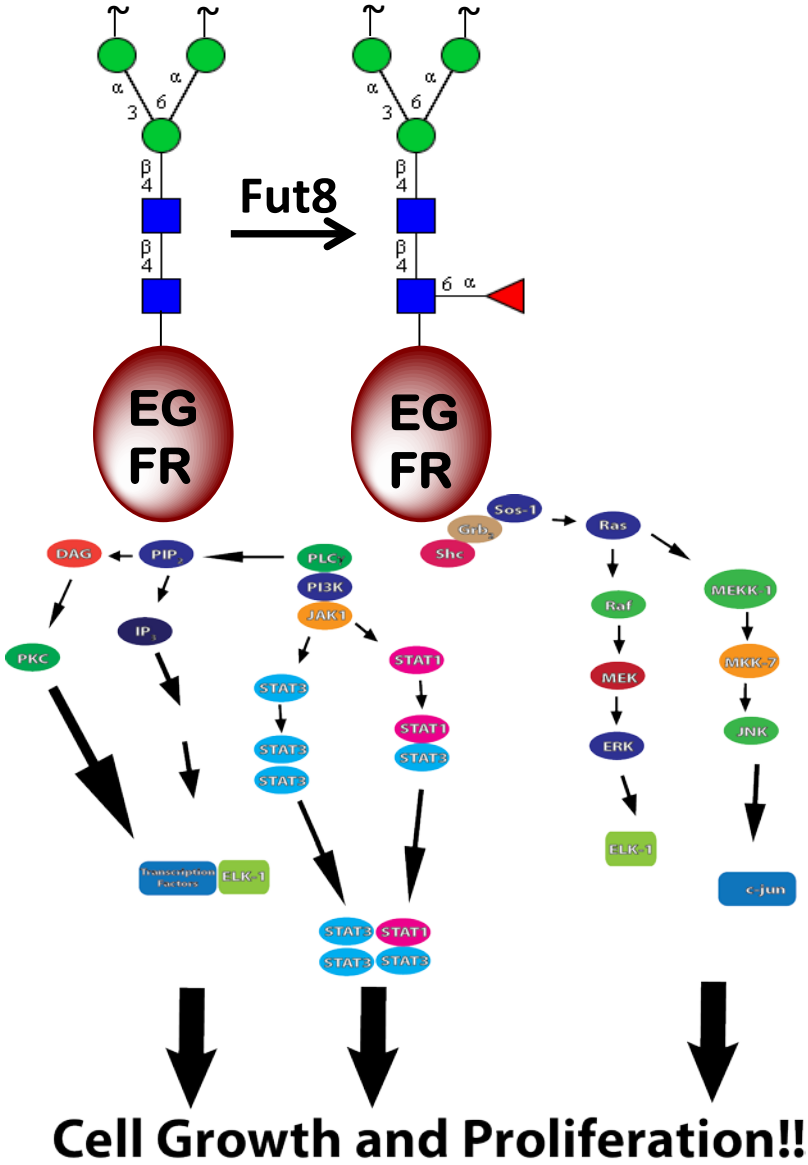
- **Synthesizing a glycome requires a large commitment of cellular resources**
- **Many glycosylation enzymes (glycosyltransferases and glycosidases), sugar transporter and metabolic proteins, and regulation elements (over 120 identified as of 2011)**
- **Can we study the genome-wide regulation of this process?**

Glycosylation in hPSC differentiation

- **Human pluripotent stem cells (hPSCs) express greater levels of α2,6 sialic acid and α1,2 fucose over differentiated cell types**
- **hPSCs express transferases *FUT1, FUT2, FUT10 (*α1,2 linkages*), ST3GAL2, ST6GAL1***
- **hPSCs repress transferases *POFUT1; POFUT2; ST3GAL1,3-6; ST6GALNAC6; ST8SIA1,4***
- **Genetic regulation of pluripotency appears to involve regulation of genes which model specific sialic and fucose epitopes on the cell surface**

Wang et al, *Cell Res,* 2011

Core fucosylation of EGFR by Fut8 stimulates EGF-depending signaling



Wang et al, *JBC*, 2006
Matsumoto et al, *Cancer Sci*, 2008

Differential core fucosylation in cancer

Increase in core fucosylation has been observed in…

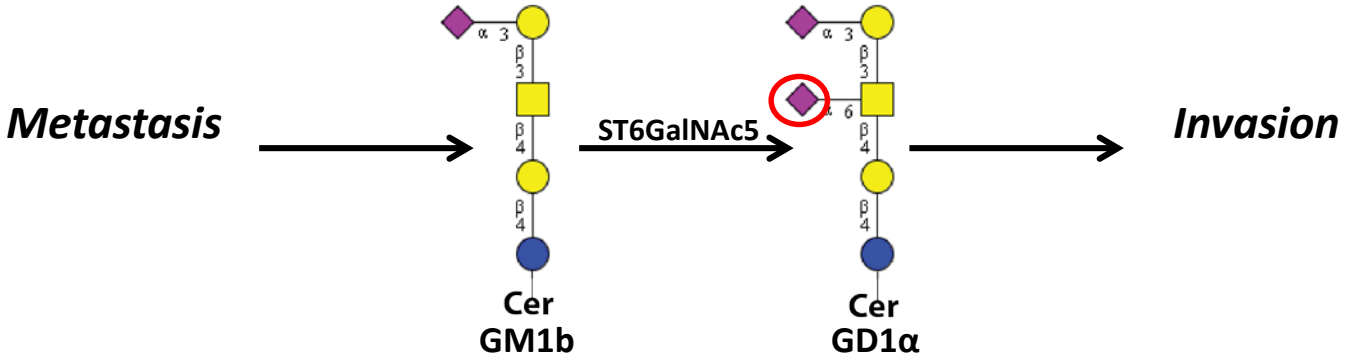
- **Prostate cancer (Saldova et al, *Glycobiology*, 2011)**
- **Pancreatic cancer (Sarrats et al, *Proteomics Clin Appl*, 2010)**
- **Stomach cancer (Bones et al, *JPR*, 2011)**

Decrease in core fucosylation has been observed in…

- **Lung cancer (Arnold et al, *JPR*, 2011)**

Some disease-specific roles for glycans

- The α2,6-sialyltransferase ST6GalNAc5 is upregulated in metastatic breast cancer, allowing greater adhesion to brain cells and passage across the blood-brain barrier



- Glycan-based biomarkers have been investigated in various cancers and schizophrenia revealing certain epitopes are more expressed in disease systems

Sialyl Lewis^X Core α1,6 fucose

- The glycome is differentially regulated in cancer, therefore cancer is a logical system in which to study variations in glycome genetic regulation

Agenda

- Carbohydrates are a diverse and critically important class of biological macromolecules
- **Microarray analysis of the glycome**
- Methods to evaluate and integrate microarray data sets
- Mechanisms for genetic regulation of the glycome

Model System: The NCI-60 Cell Panel

- **NCI-60: 60 cell lines for screening of potential cancer therapeutics**
- **Vary in tissue type, metastasis, individual of origin**
- **CellMiner: open source database containing mRNA, miRNA and protein array data, genetic mapping, pharmacological and mutational analysis**

The NCI-60 panel

LEUKEMIA
- CCRF-CRM
- HL-60
- K-562
- MOLT-4
- RPMI-8226
- SR

LUNG
- A549
- EKVX
- HOP-62
- HOP-92
- NCI-H226
- NCI-H23
- NCI-H322M
- NCI-H460
- NCI-H522

COLON
- COLO 205
- HCC-2998
- HCT-116
- HCT-15
- HT29
- KM12
- SW-620

CNS
- SF-268
- SF-295
- SF-539
- SNB-19
- SNB-75
- U251

MELANOMA
- LOX IMVI
- MALME-3M
- M14*
- MDA-MB-435*
- SK-MEL-2
- SK-MEL-28
- SK-MEL-5
- UACC-257
- UACC-62

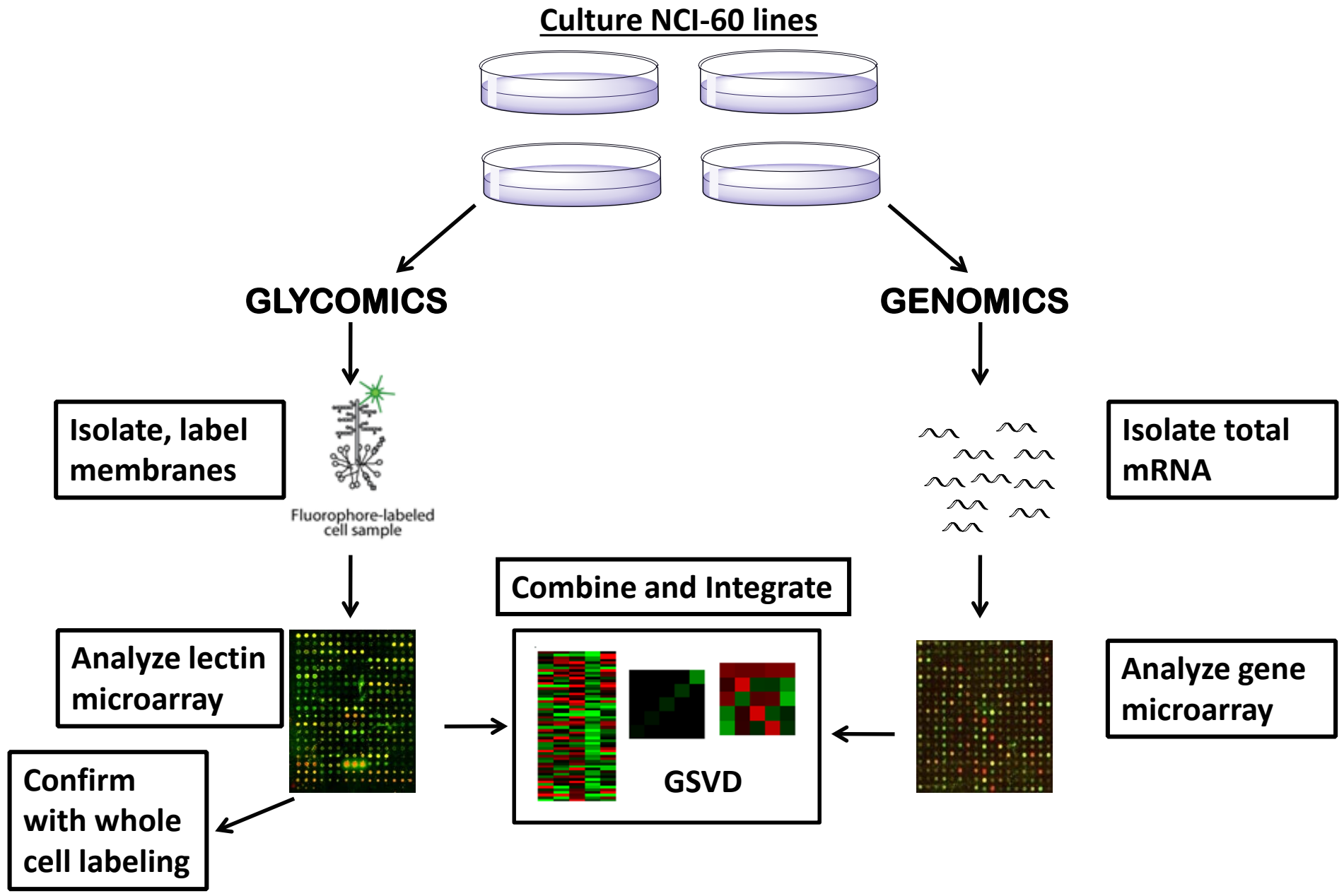
OVARIAN
- IGR-OV1
- OVCAR-3
- OVCAR-4
- OVCAR-5
- OVCAR-8
- NCI/ADR-RES
- SK-OV-3

PROSTATE
- PC-3
- DU-145

RENAL
- 786-0
- A498
- ACHN
- CAKI-1
- RXF 393
- SN12C
- TK-10
- UO-31

BREAST
- MCF7
- MDA-MB-231
- MDA-MB-468
- HS 578T
- BT 549
- T-47D

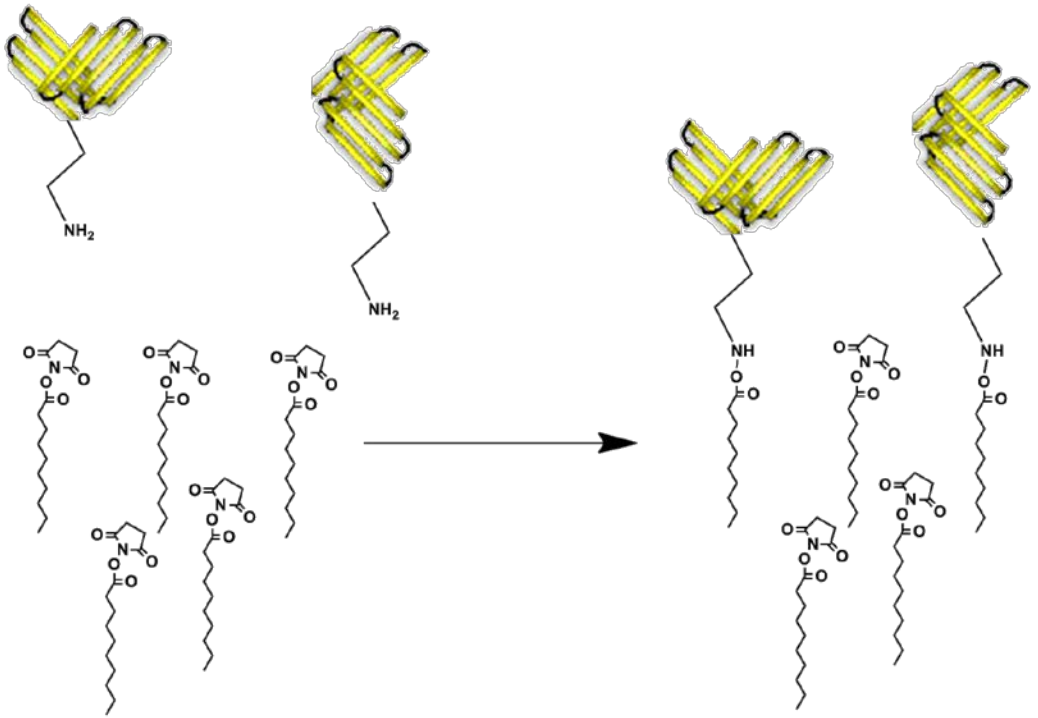
Experimental Outline



Lectins

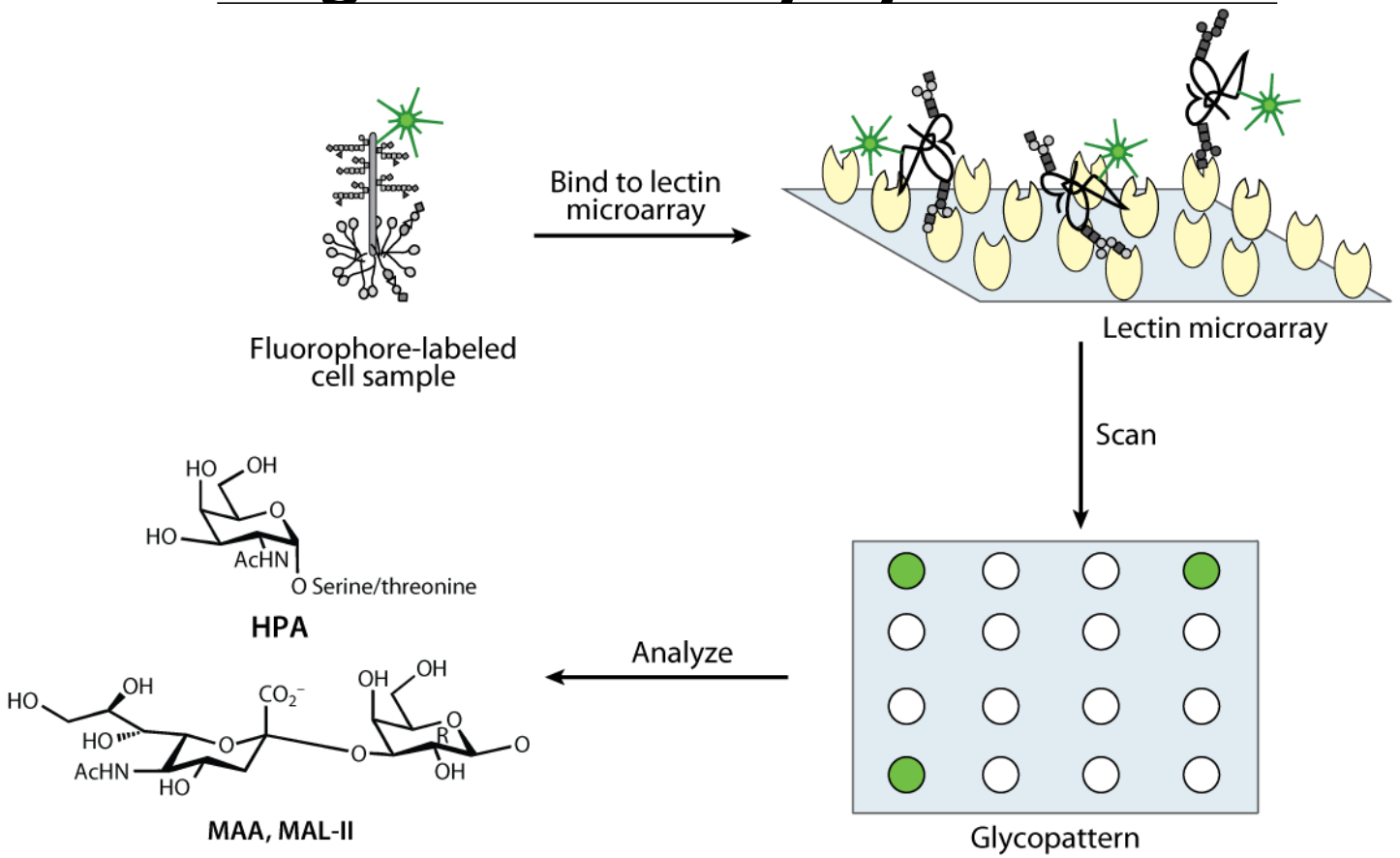
- **Lectins: non-enzymatic, non-immunological carbohydrate-binding proteins**
- **Found throughout the tree of life and demonstrate a wide range of structural design and biological function**
- **Lectins are often modified post-translationally (particularly in higher organisms)**
- **Many are easily commercially available**
- **Some have very broad specificities (WGA) some are extremely restricted (PHA-E)**

Generation of Lectin Microarrays



- **Lectins are printed on NHS-ester coated glass slides in high spatial density at 10°C and ambient humidity**
- **Protein lysine residues react with esters to form amide-bound conjugates**
- **Unreacted esters are blocked with ethanolamine**
- **Slides can be stored for up to two months**

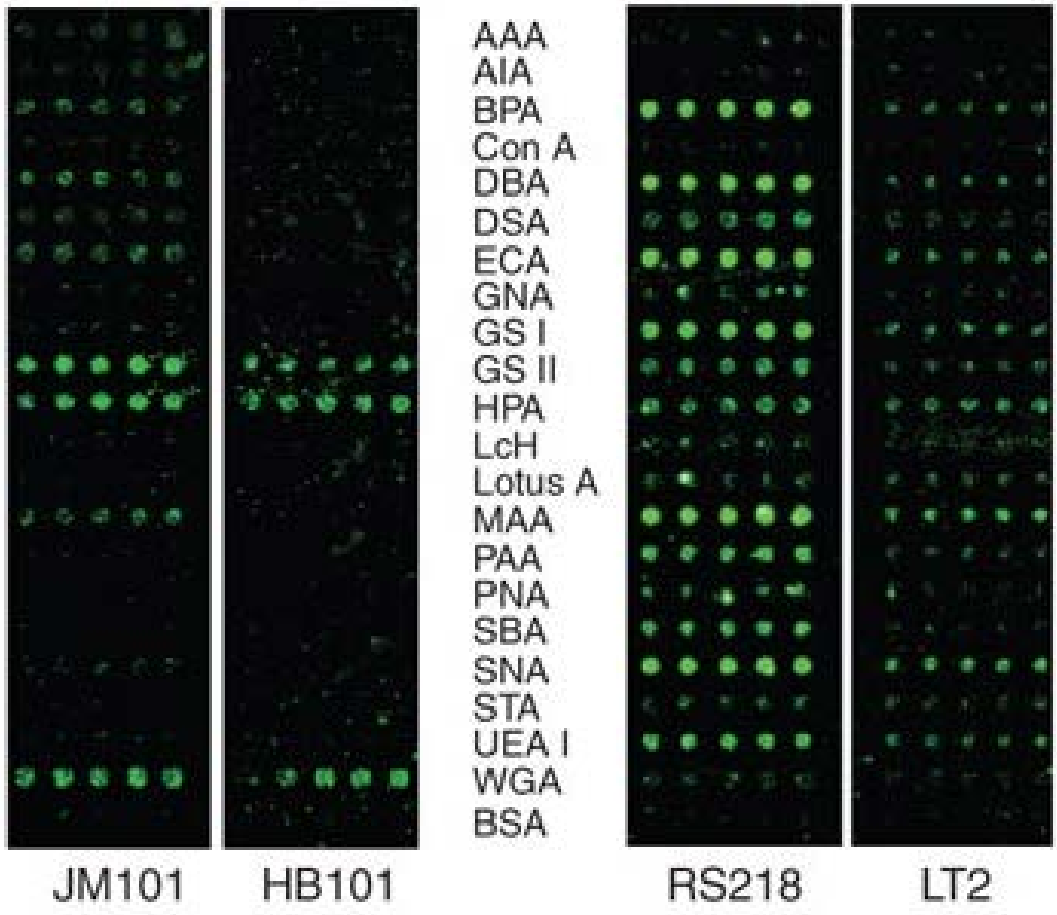
Single color array hybridization



1. **Glycome of sample of interest is isolated as membrane liposome**
2. **Membrane protein content fluorescently labeled**
3. **Labeled sample hybridized to lectin microarray**
4. **"Glycopattern" of lectin specificities reveals expressed carbohydrate cohort**

Pilobello et al, *PNAS*, 2007
Rakus and Mahal, *Ann Rev Anal Chem*, 2011

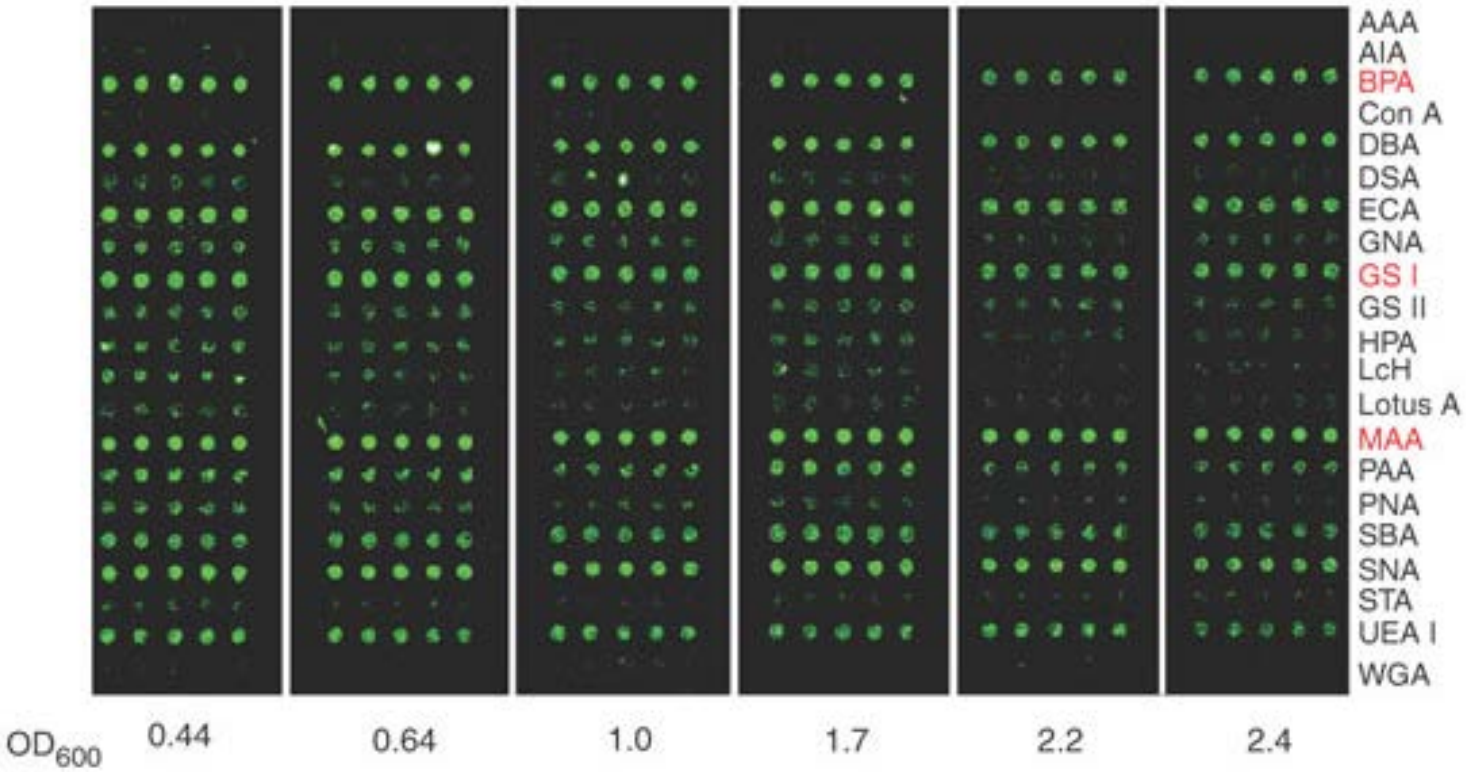
Single color lectin array approach differentiates glycomes of *E. coli* strains



- **Surface glycome of pathogenic (RS218) and non-pathogenic (JM101, HB101) *E. coli* show distinct differences**
- **Pathogenic *S. typhimurium* (LT2) glycome distinct**

Hsu et al, *Nat Chem Biol*, 2006

Single color lectin array approach differentiates glycomes of *E. coli* strains

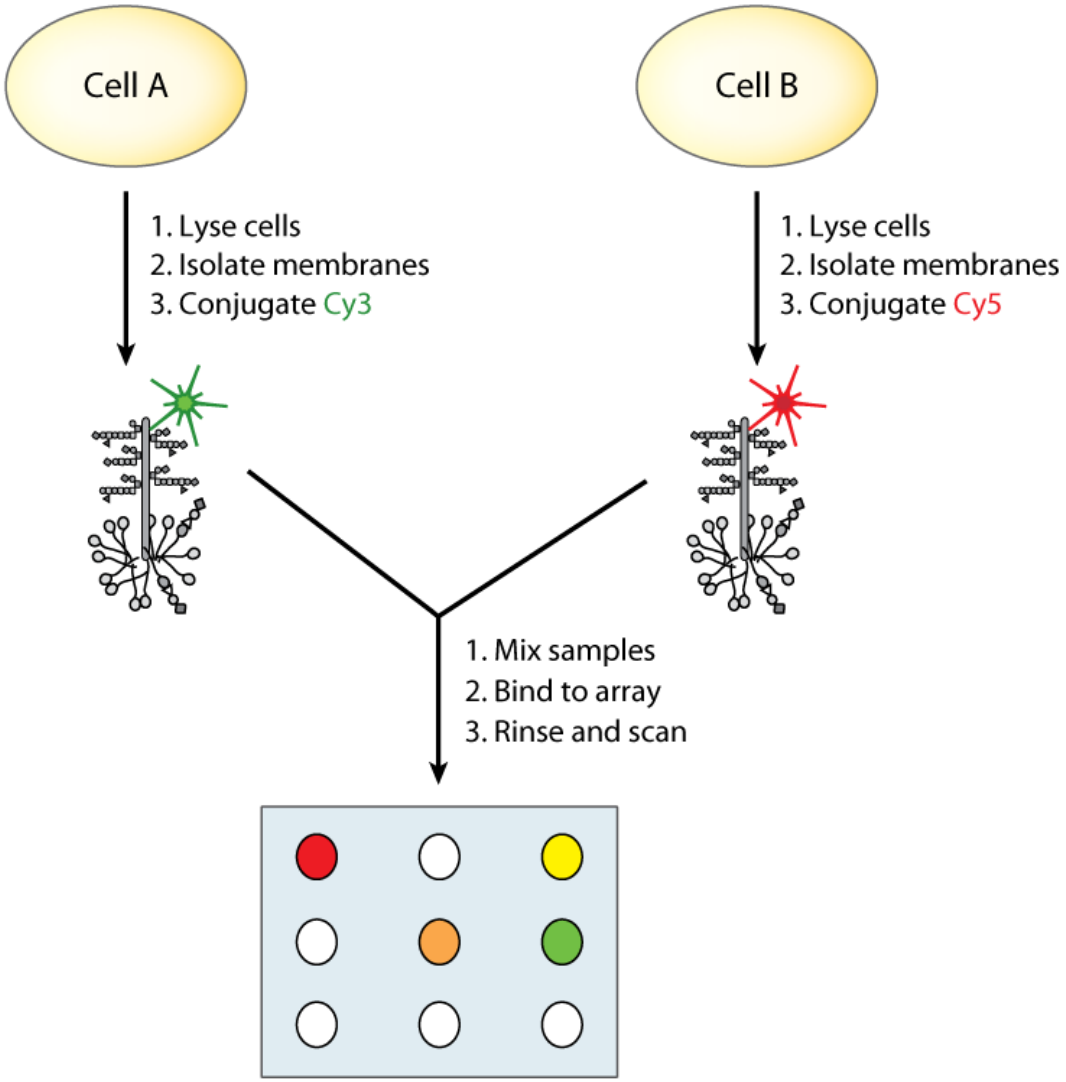


- **The bacterial glycome is dynamic**

- ***E. coli* RS218 glycome varies as it proceeds from log to stationary growth phase**

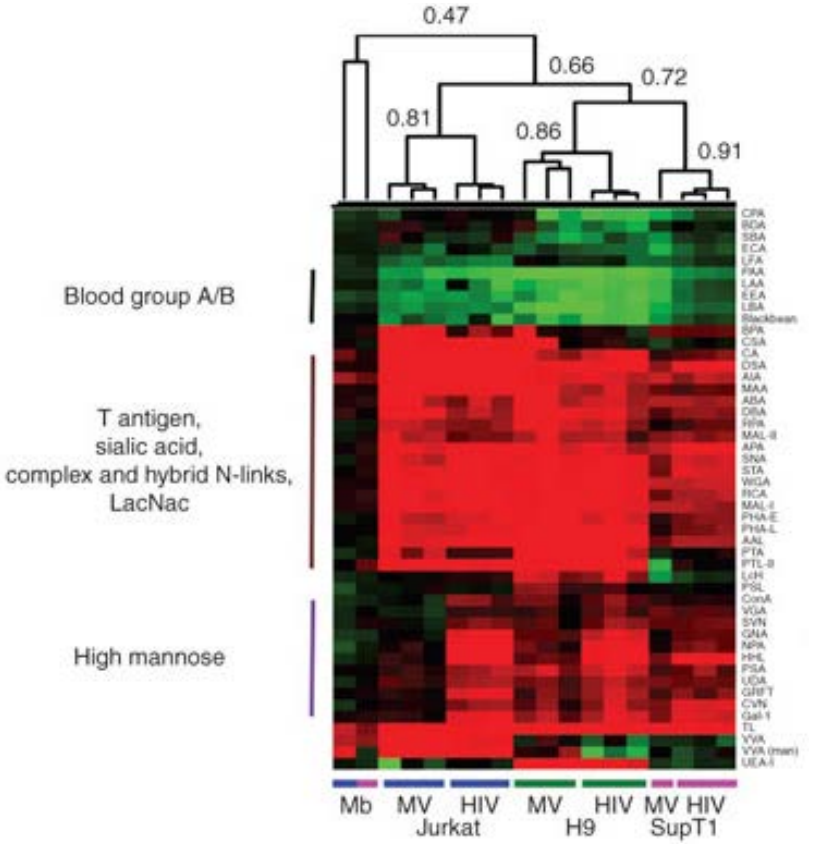
Hsu et al, *Nat Chem Biol*, 2006

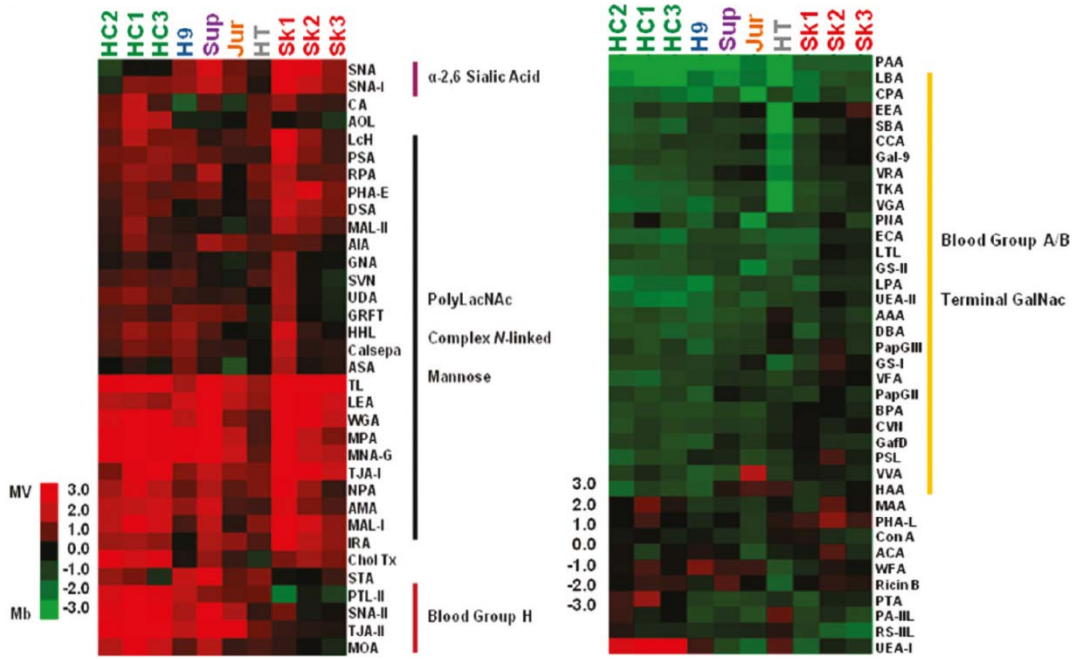
Ratiometric lectin microarray analysis for semi-quantitative analysis of the dynamic glycome



Pilobello et al, *PNAS*, 2007
Rakus and Mahal, *Ann Rev Anal Chem*, 2011

Microvesicles have a distinct glycomic signature which is hijacked by HIV





Glycome of HIV virions are more closely related to microvesicles than membranes

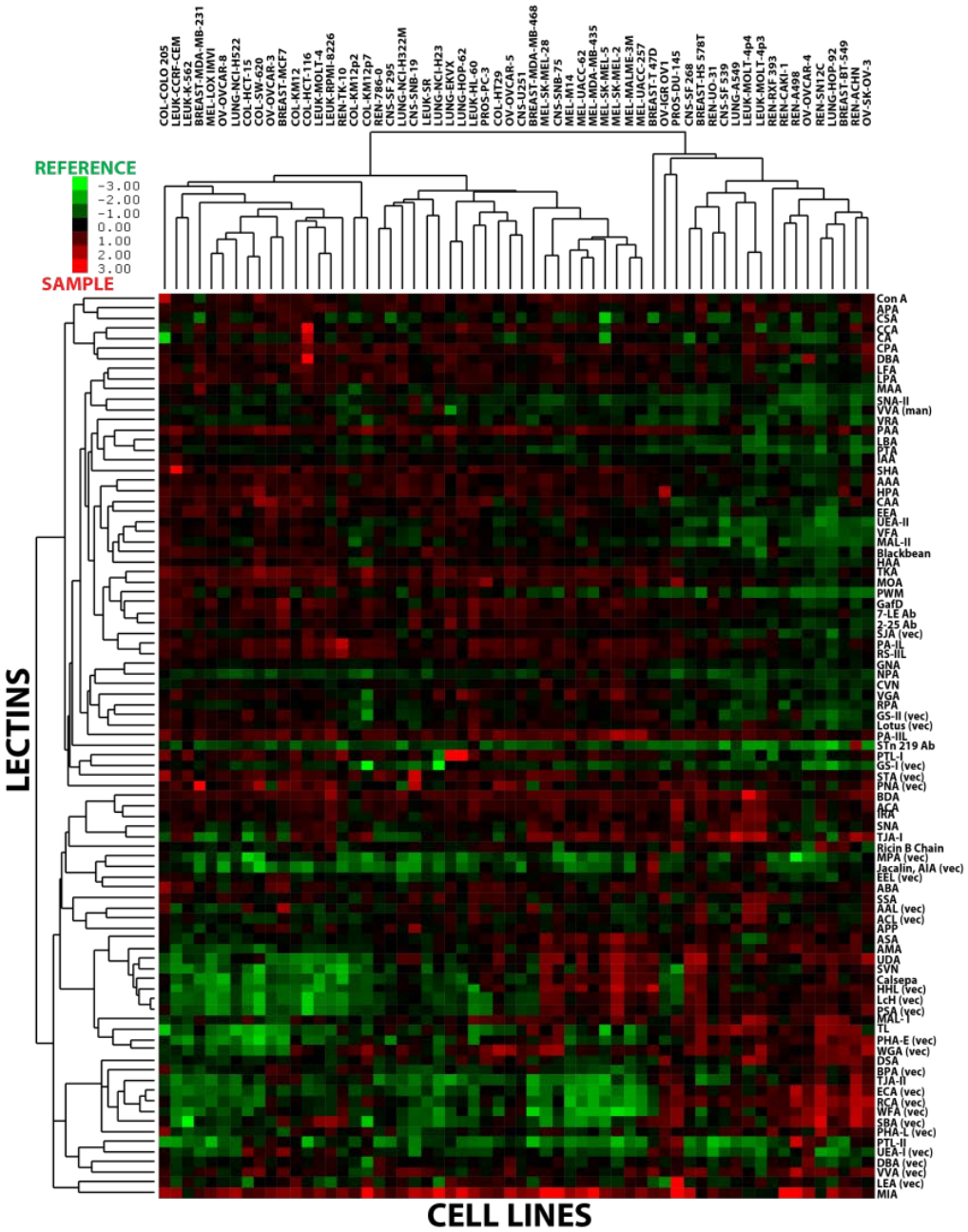
Microvesicular glycome is distinct from membrane and independent of cellular origin

Krishnamoorthy et al, *Nat Chem Biol*, 2009
Batista et al, *J Proteome Res,* 2011

Glycomic Analysis of the NCI-60

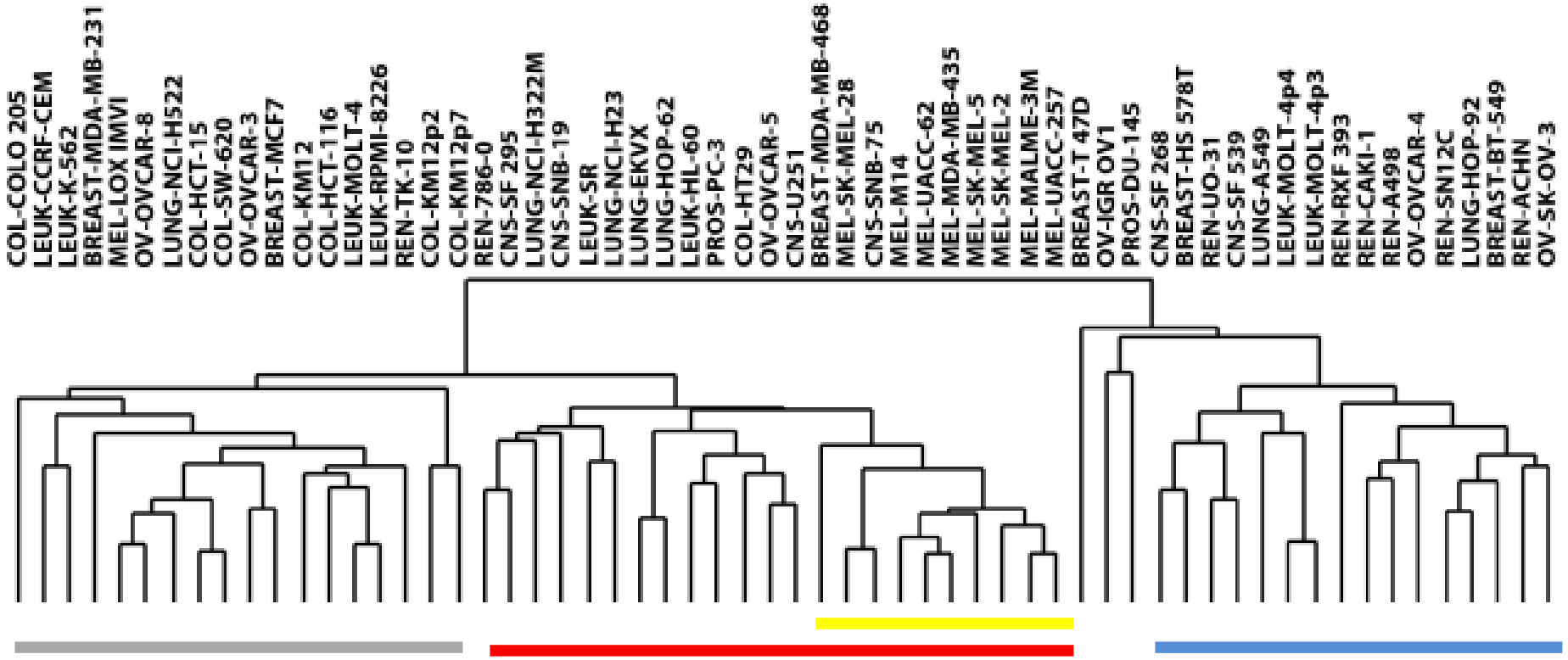
- **Sample Preparation: All NCI-60 cell lines were grown to high density for membrane collection and labeling for hybridization**
- **Array Preparation: 88 lectins and 3 antibodies were printed in triplicate on NHS-ester hydragel slides**
- **Experimental Design: Hybrid each cell line in 2-color experiment with a common, mixed cell-line reference (Reference = OVCAR-3, OVCAR-4, HS 578T, LOX IMVI, PC-3, SNB-19)**
- **Data Manipulation: Threshold positive signals and normalize each array experiment by centering the fluorescence data to the median-ratio for the fluorescence signals of that array**
- **Data Interpretation: Perform cluster analysis of all array experiments to identify expression patterns and trends**
- **Confirmation: Select representative cell lines and lectins to perform live cell, lectin-histology**

Lectin Microarray Analysis: NCI-60



- Two lectins were discarded as misprints
- Two lung cell lines were discarded due to technical error (NCI-H226, NCI-H460)
- One colon line was not included due to inability to culture (HCC 2998)
- Biological replicates of KM12 and MOLT-4 were included

Cell lines cluster based on their glycosylation



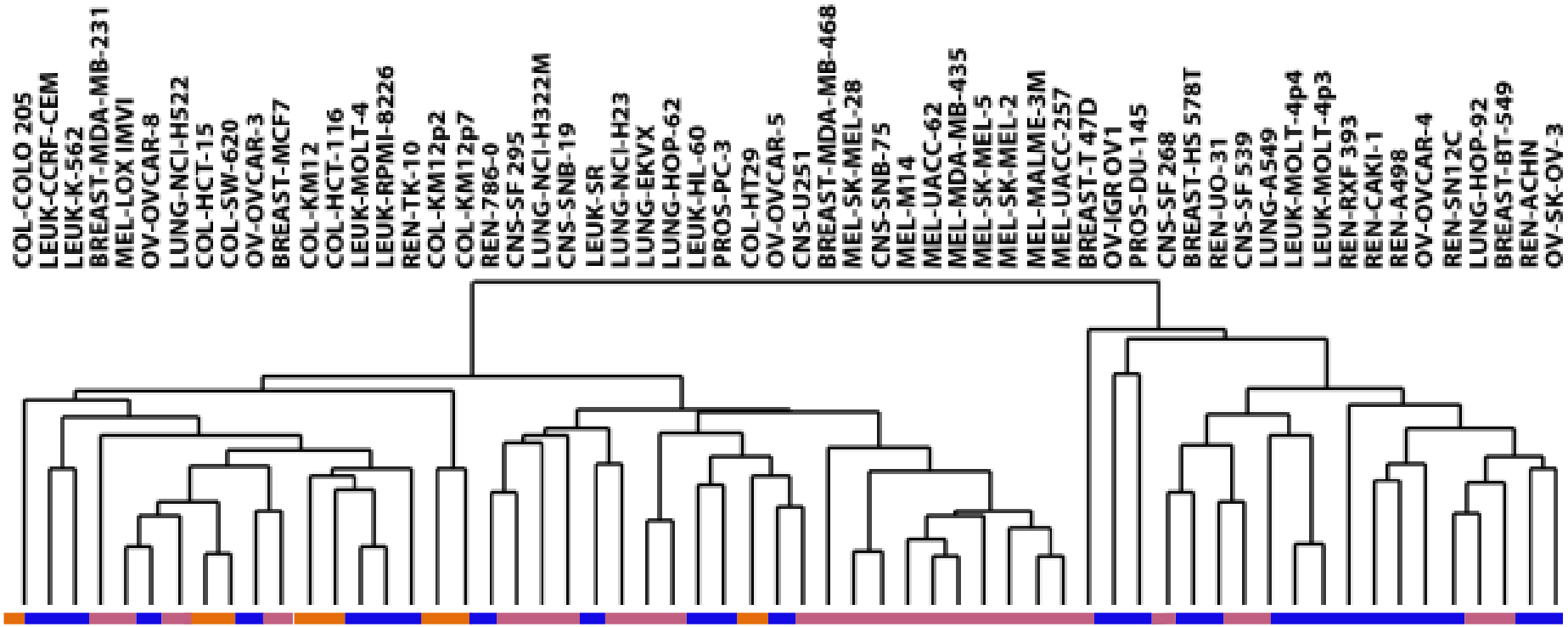
COLON (5/6, p=0.39)
LEUKEMIA (4/6, p=0.39)

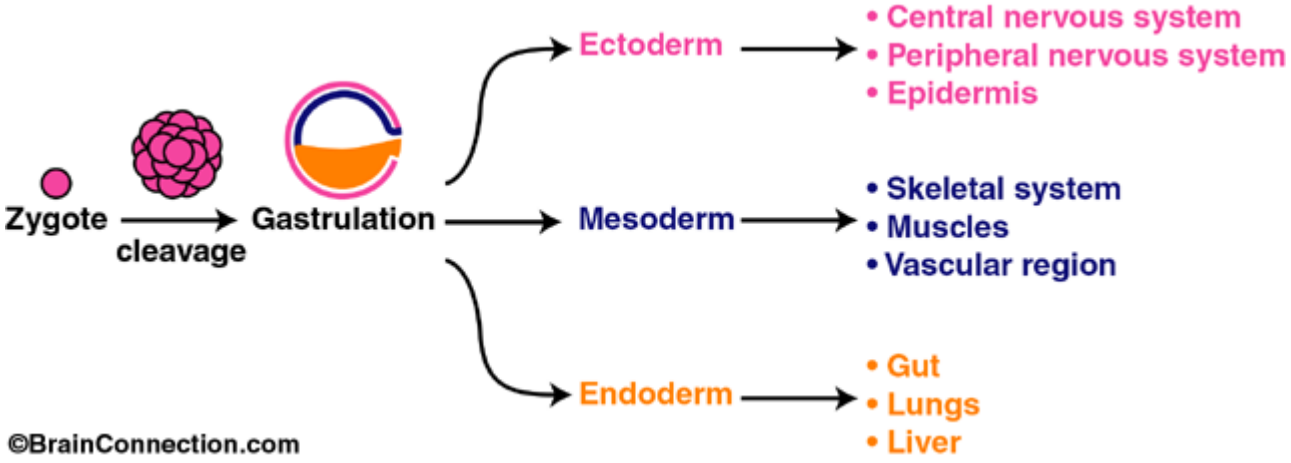
CNS (4/5, p=0.44)
LUNG (4/7, p=0.44)

MELANOMA (8/9, p=0.60)

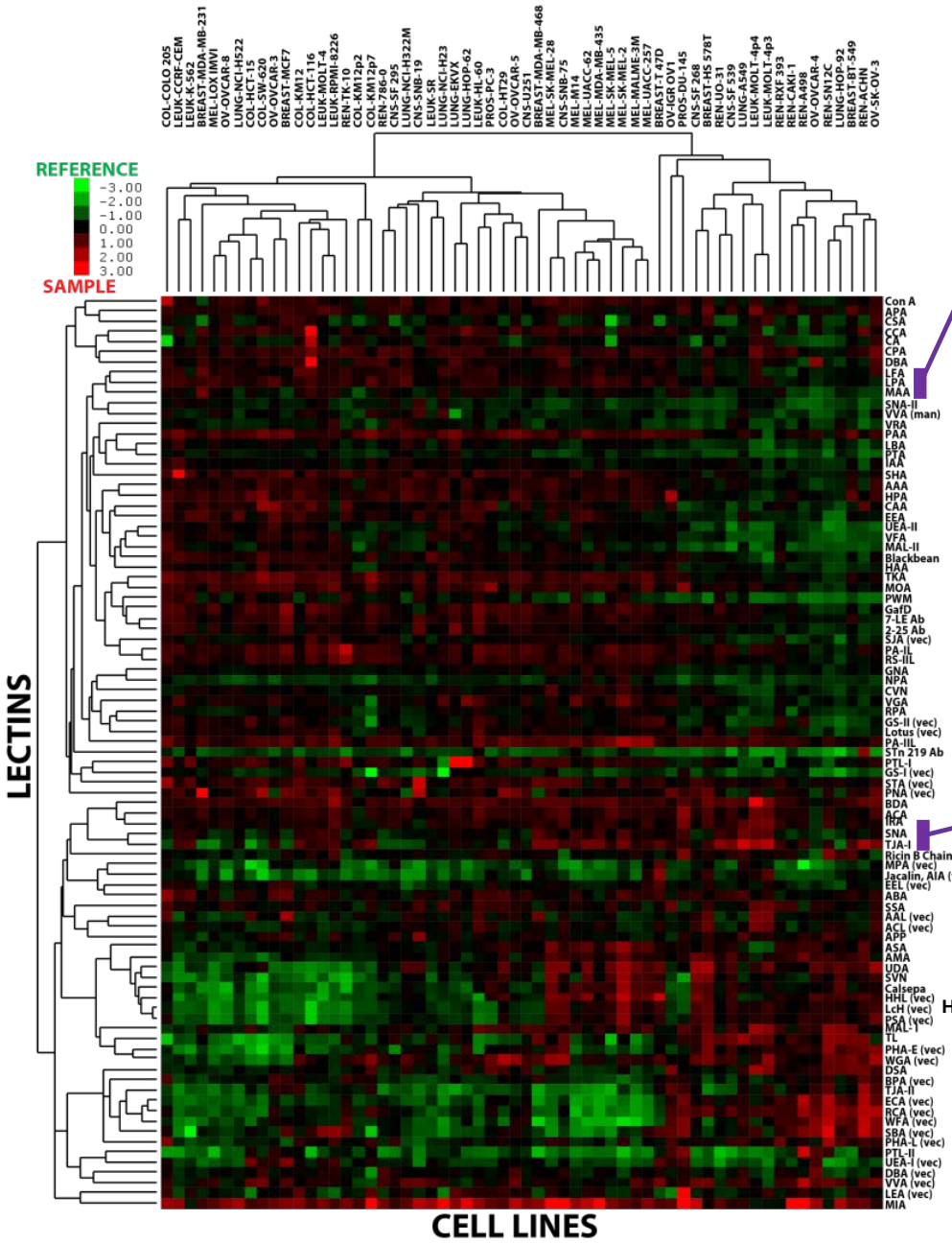
RENAL (6/8, p=0.39)

Histological origin does not explain cluster pattern





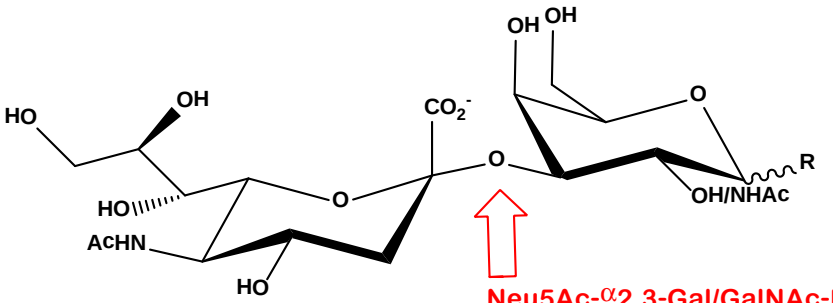
Glycosylation signatures: Sialic acid



α2,3 sialic acid recognizing lectins

LFA, LPA, MAA

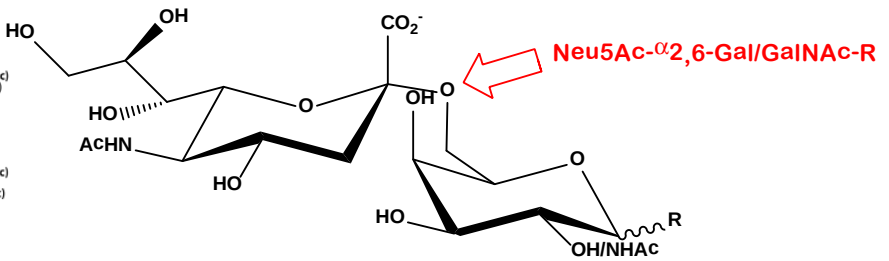
in RENAL cell lines



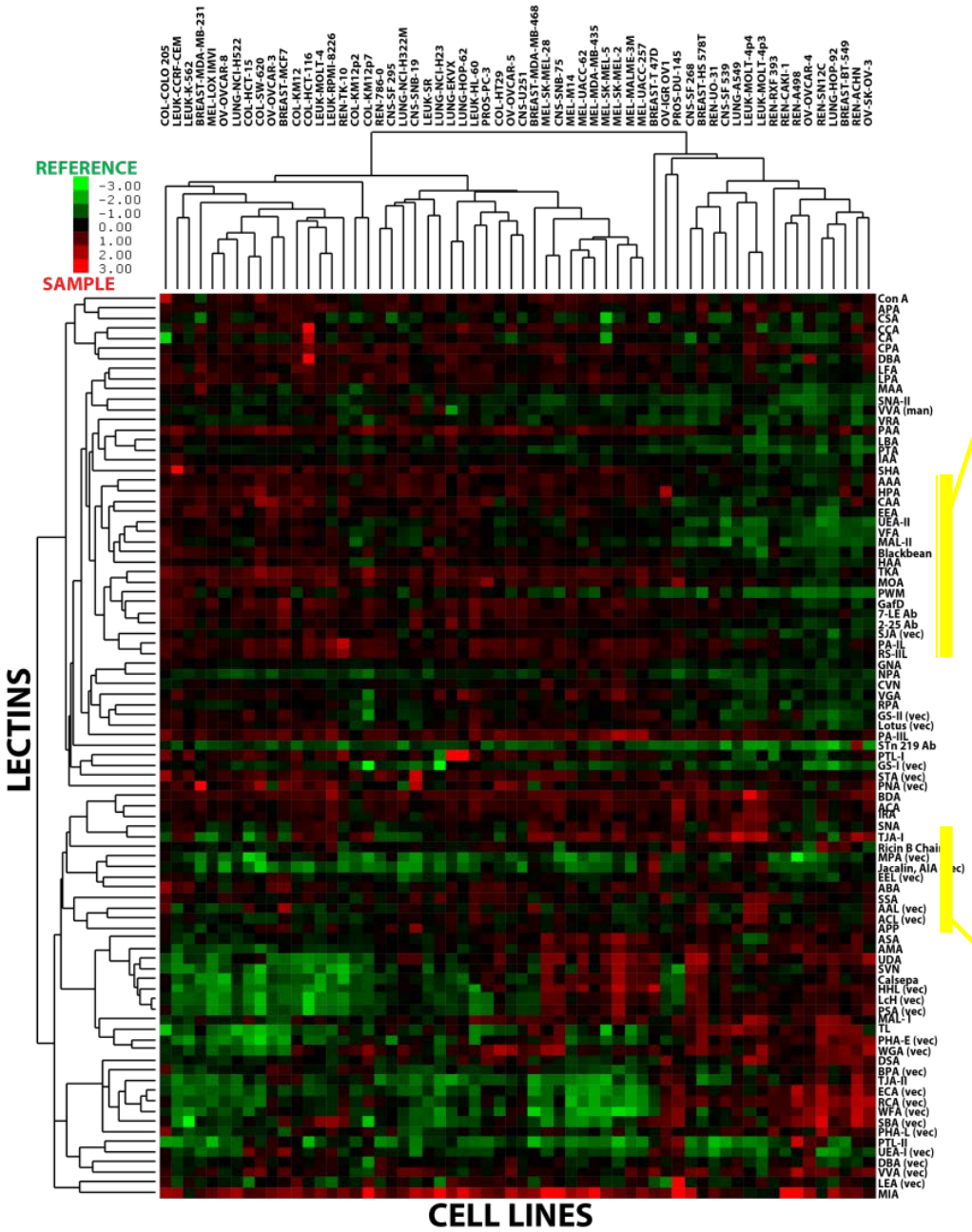
α2,6 sialic acid recognizing lectins

SNA, TJA-I

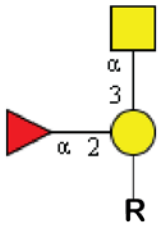
in MELANOMA and LUNG cell lines

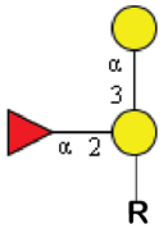


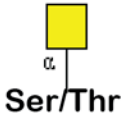
Glycosylation signatures: Gal/GalNAc



Blood groups/Gal/GalNAc
MOA, EEA, LBA, VRA, IAA, PTA, HPA

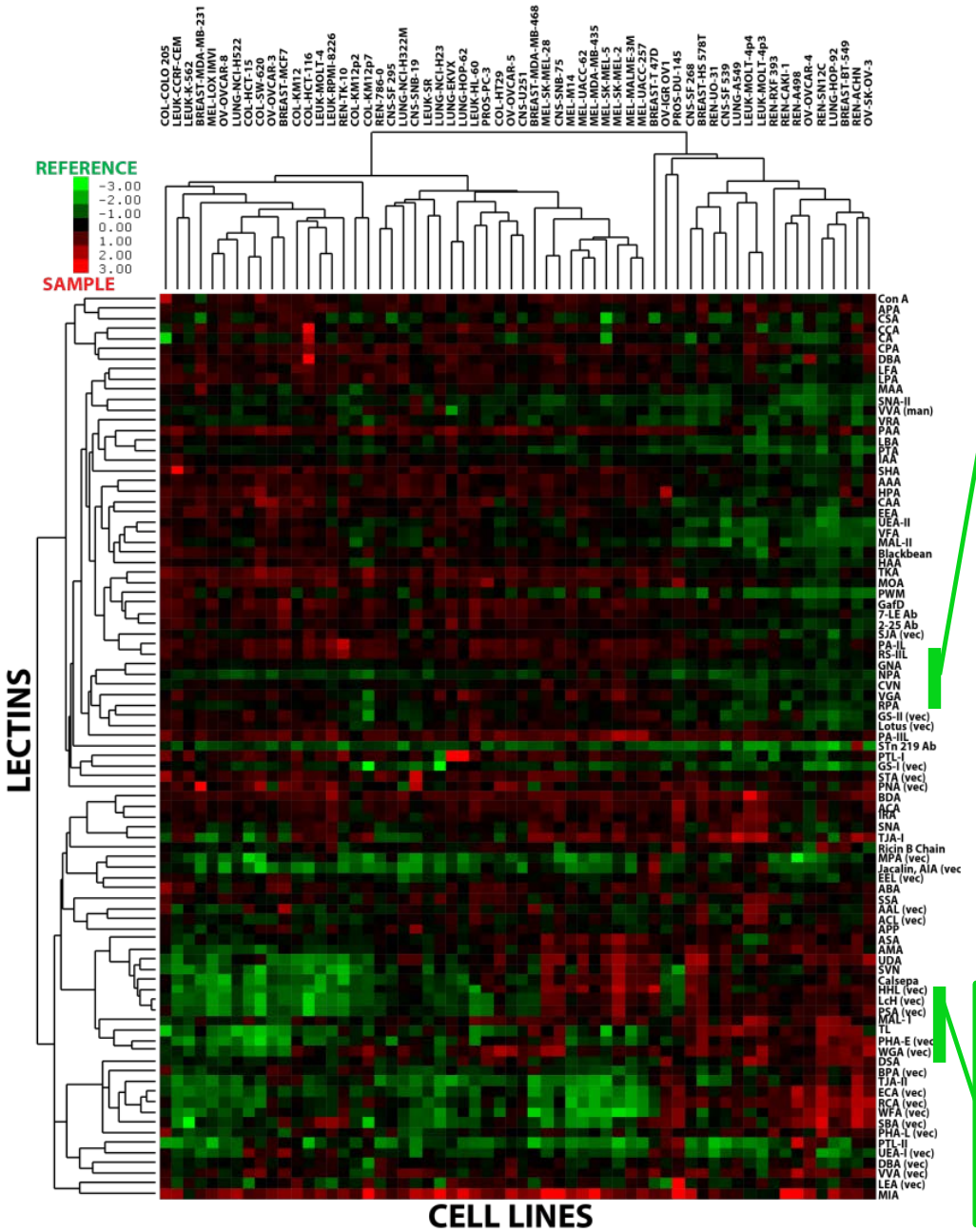


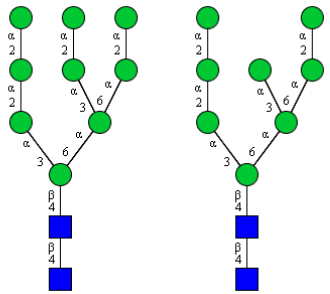




Blood groups/Gal/GalNAc
BDA, RICIN, ABA, AIA, AAL

Glycosylation signatures: Mannose

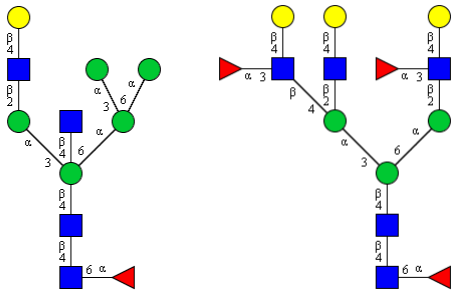




RENAL

MELANOMA
CNS, LUNG
COLON

Mannose (high)
GNA, NPA, CVN, RPA



COLON, LUNG

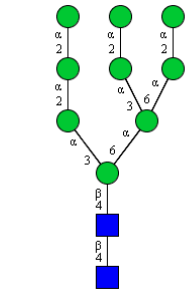
MELANOMA, RENAL

Mannose (complex)
UDA, SVN, Calsepa, HHL, LcH, PHA-E, MAL-I, PSA

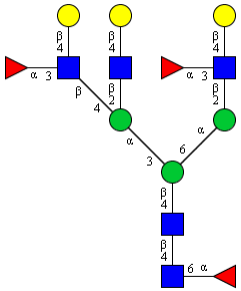
Lectin histology confirms glycan presence

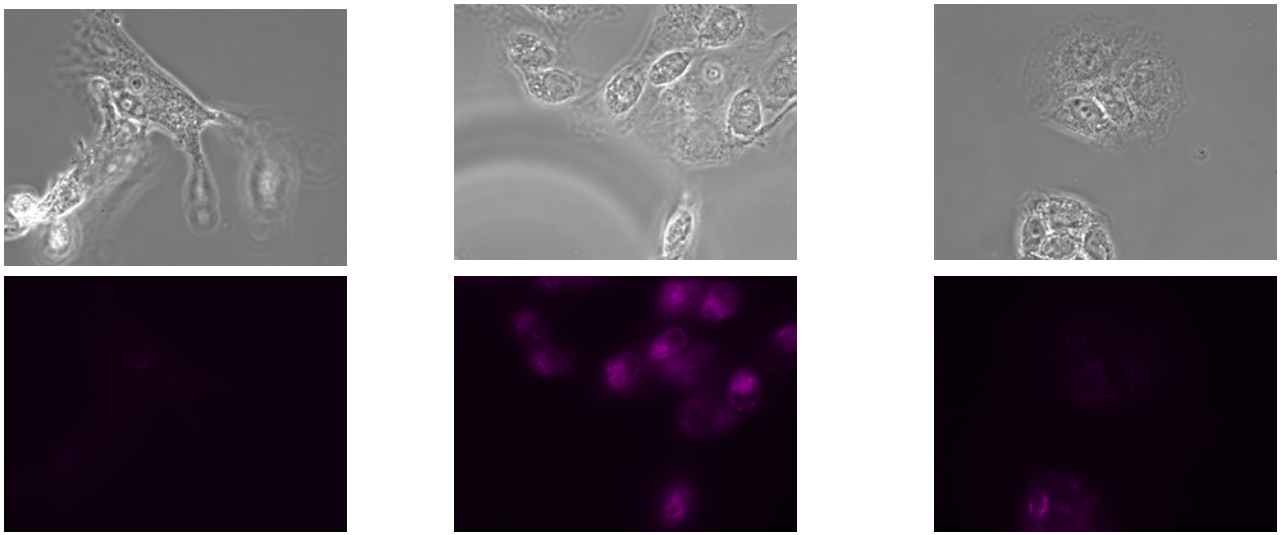
SK-MEL-2 ACHN MCF7

GNA



UDA

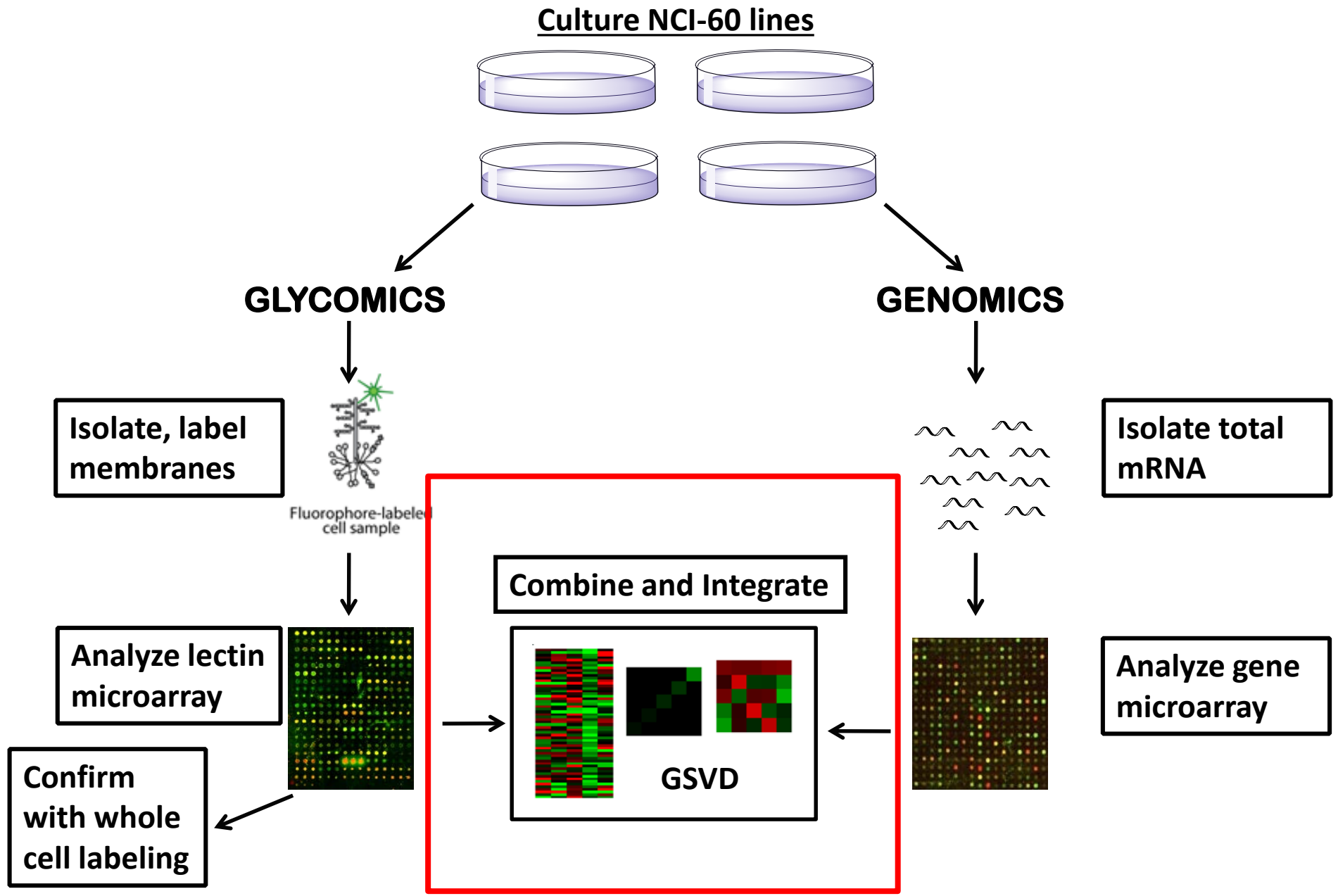




Agenda

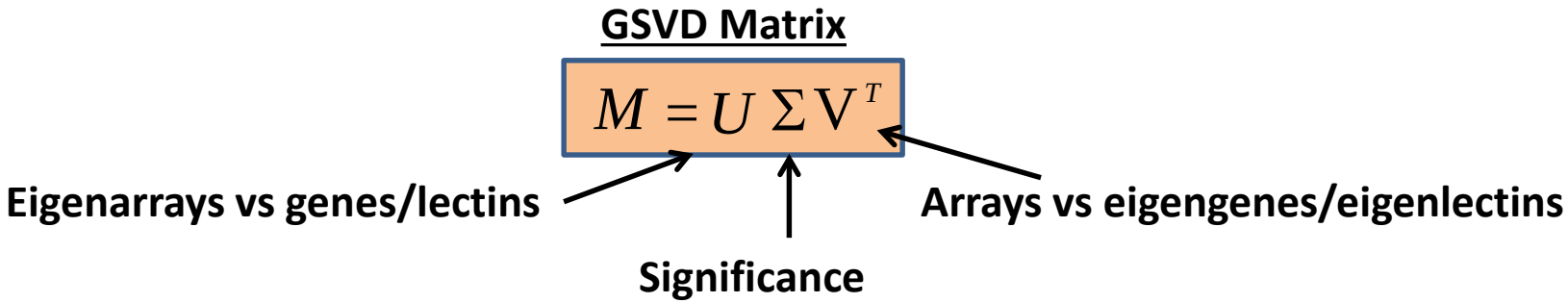
- Carbohydrates are a diverse and critically important class of biological macromolecules
- Microarray analysis of the glycome
- **Methods to evaluate and integrate microarray data sets**
- Mechanisms for genetic regulation of the glycome

Experimental Outline



Generalized Singular Value Decomposition (gSVD)

- **SVD and gSVD are methods to decompose a rectangular matrix into constrained, component matrices**
- **This method can be used to identify significant, co-varying patterns in large data sets**
- **We apply it to the lectin or mRNA array data individually or both data sets combined for SYSTEMS BIOLOGY examination of glycosylation pathways**



Do genetic expression patterns co-vary with the glycome?

- **Work of Kanoelani Pilobello**
- **Performed SVD analysis of mRNA (and miRNA) data with NCI-60 lectin microarray results**
- **Analyzed all melanoma, renal, colon and leukemia cell lines**

Regulation of gene expression across the NCI-60

- **mRNA and miRNA expression patterns investigated across entire NCI-60 (Liu et al, *Mol Biol Cell,* 2010)**

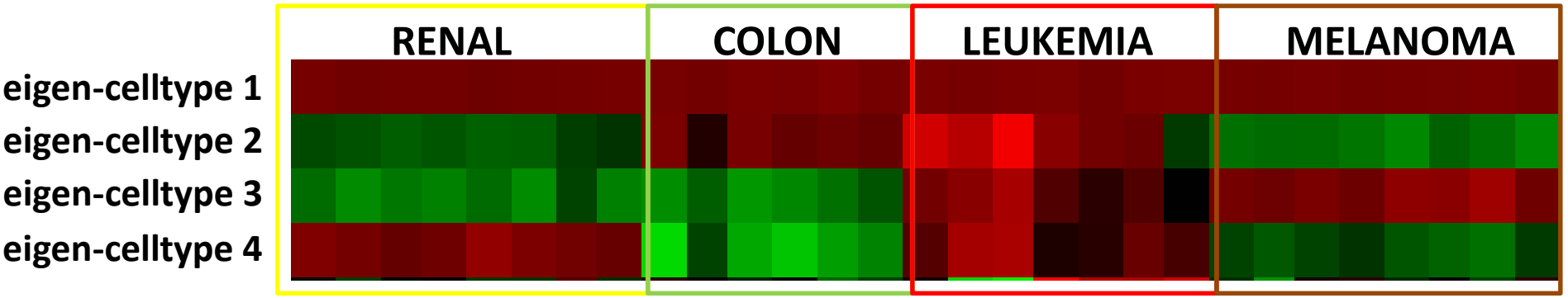
OBSERVATION OF HIGHLY EXPRESSED AND DIVERSE PROBES

-Melanoma (9/10 mRNA, 10/10 miRNA) and leukemia (6/6 mRNA and miRNA) lines cluster together based on expression patterns

-Renal (8/8 mRNA, 6/8 miRNA) and colon (6/7 mRNA and miRNA) also show strong correlation

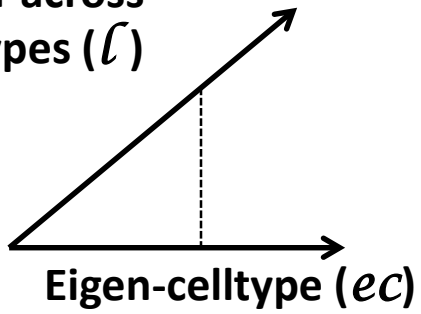
-CNS and lung cancer show little correlation

GSVD Analysis of the glycome



Projection correlation = $\mathcal{C} \,||\, ec$

Gene/lectin vector across cell types ($\mathcal{C}$)



Glycomic Signatures

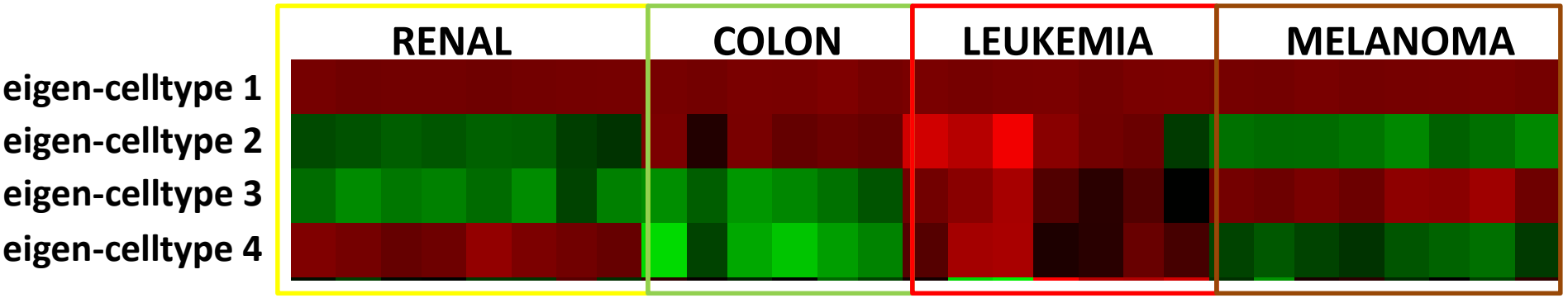
eigen-celltype 2	APA, Con A, DSA , GNA, LcH, LPA, NPA, PAA ,SNA (vector), UDA, HHL, MAL-II, CVN, SVN, GRFT, ACA, PSL, TJA-I, AMA, Calsepa, IRA, RS-IIL	High mannose, α2,6 sialic acid (*N*-linked)
eigen-celltype 3	VRA, PTL-II, HAA	Terminal GalNAc (*O*-linked)
eigen-celltype 4	TKA, PapGII	Glycolipid

GSVD Matrix

$$M = U \Sigma V^T$$

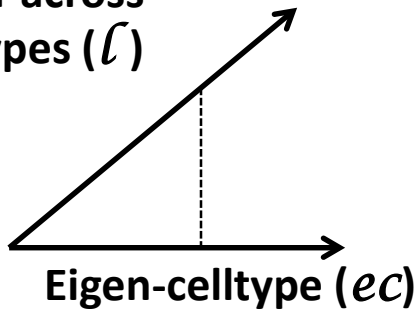
Lani Pilobello

GSVD Analysis of the genome



Projection correlation = $\ell \,||\, ec$

Gene/lectin vector across cell types (ℓ)



GSVD Matrix

$$M = U \Sigma V^T$$

Genomic Signatures

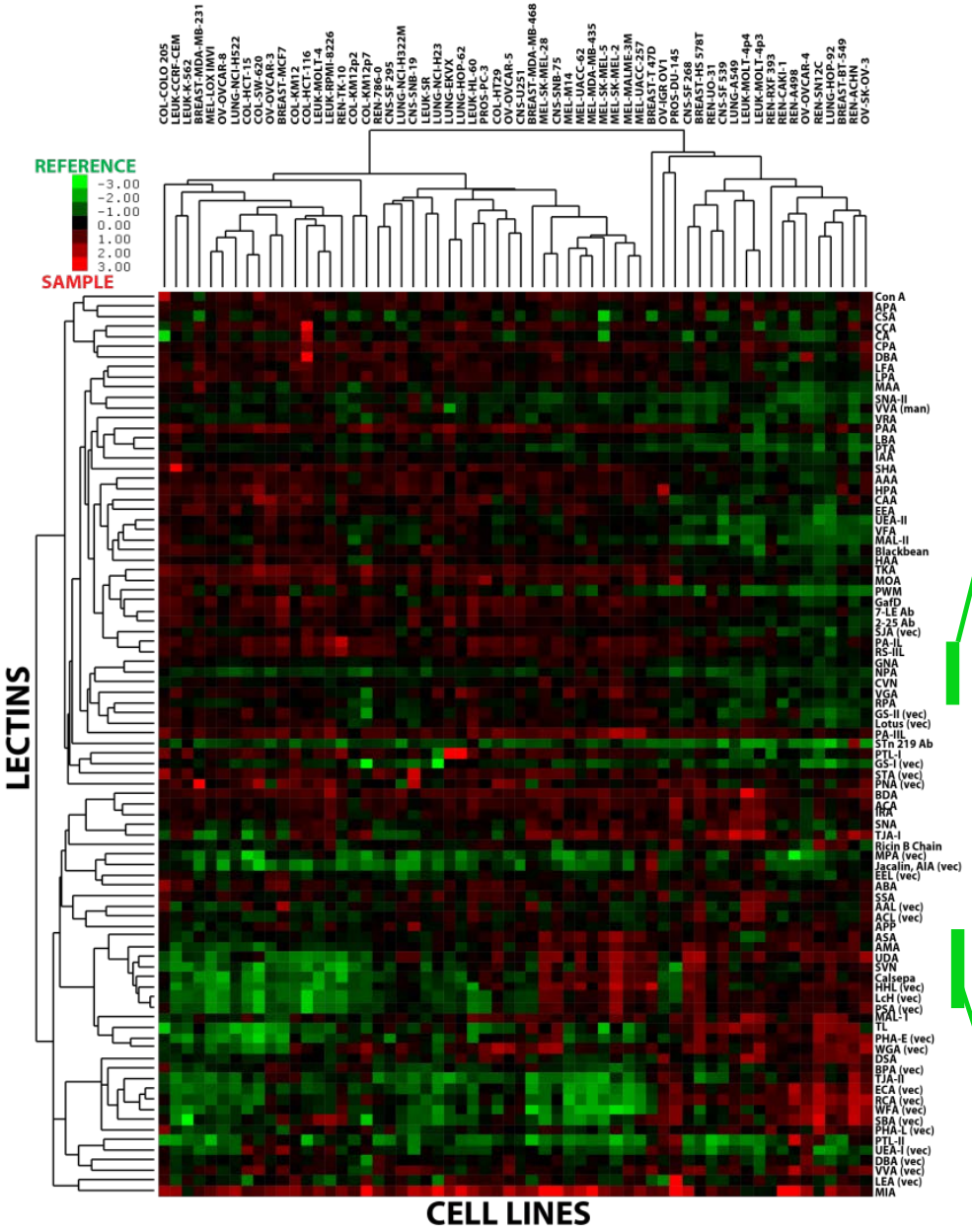
eigen-celltype 2	RNA-binding	Cytoskeletal	Organelle Lumenal
eigen-celltype 3	Ribosomal	Organelle lumenal	mRNA splicing/ processing
eigen-celltype 4	Mitochondrial	Organelle lumenal	

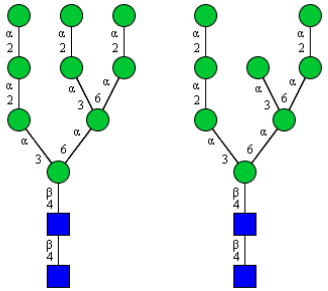
Lani Pilobello

Agenda

- Carbohydrates are a diverse and critically important class of biological macromolecules
- Microarray analysis of the glycome
- Methods to evaluate and integrate microarray data sets
- **Mechanisms for genetic regulation of the glycome**

Glycosylation signatures: Mannose

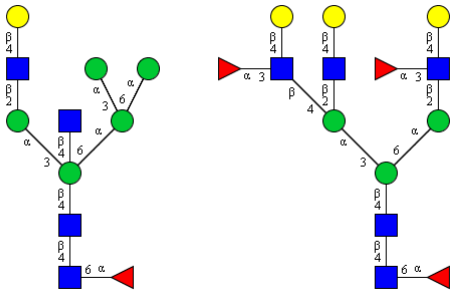




RENAL

MELANOMA
CNS, LUNG
COLON

Mannose (high)
GNA, NPA, CVN, RPA

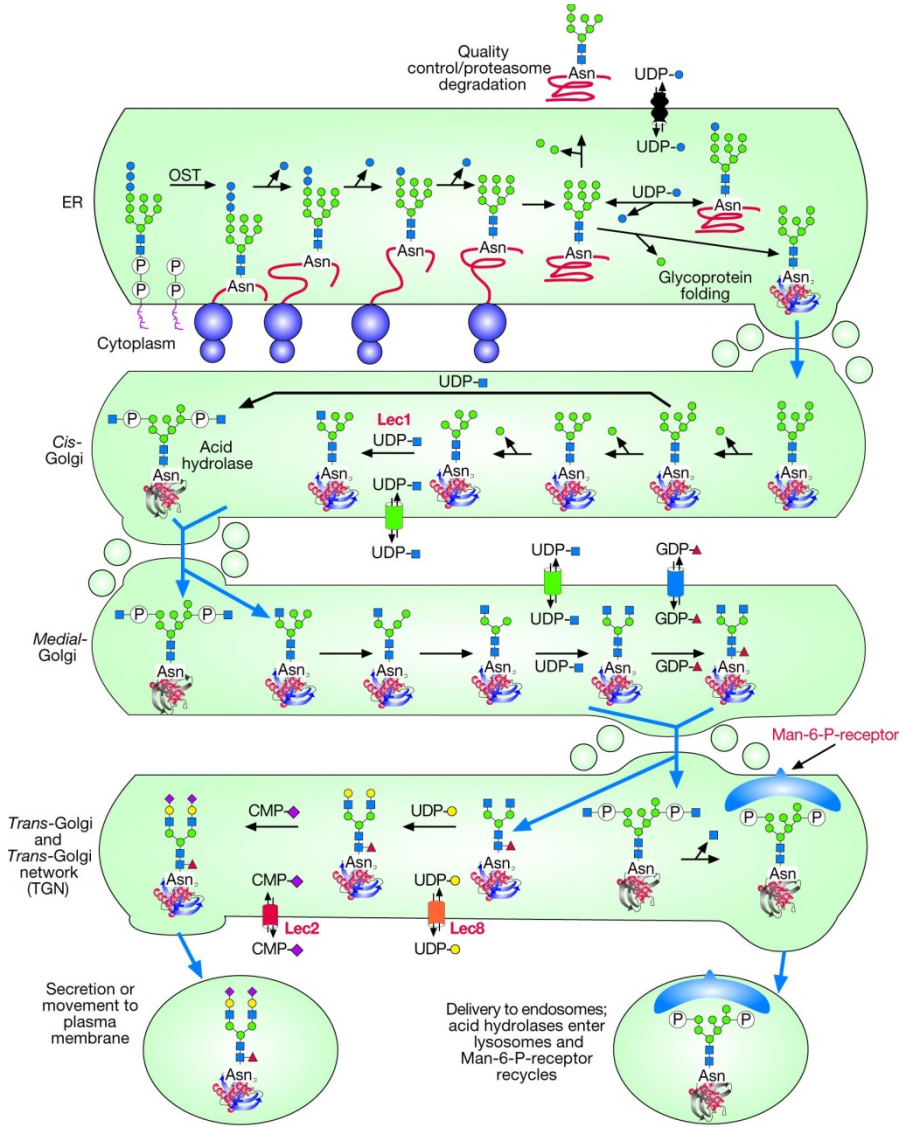


Mannose (complex)
UDA, SVN, Calsepa, HHL, LcH, PHA-E, MAL-I, PSA

COLON, LUNG

MELANOMA, RENAL

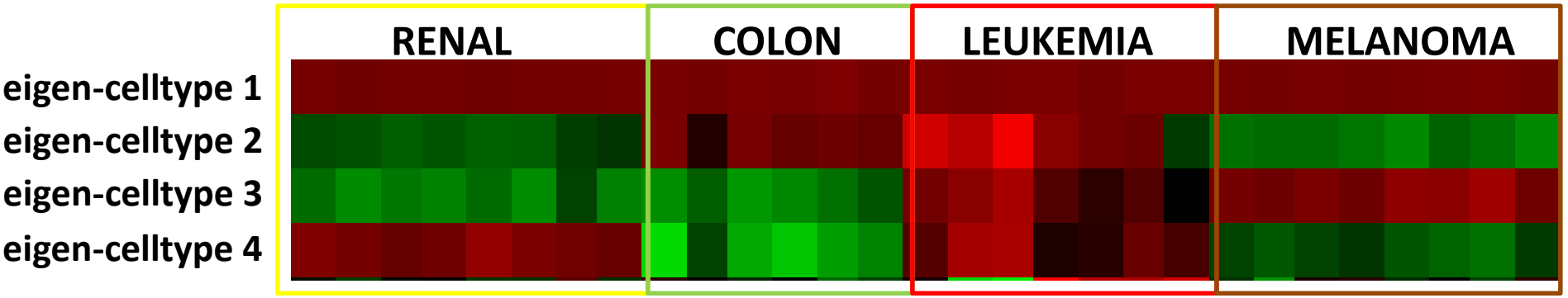
Maturation of *N*-linked glycans



- $Glc_3Man_9GlcNAc_2$ precursor is transferred to growing polypeptide in ER
- Upon proper folding, glycan is trimmed to high-mannose and, potentially further modified to hybrid/complex
- Maturation occurs in Golgi
- Our data suggests this pathway is involved in controlling cellular identity

Essentials of Glycobiology

GSVD Analysis of the genome

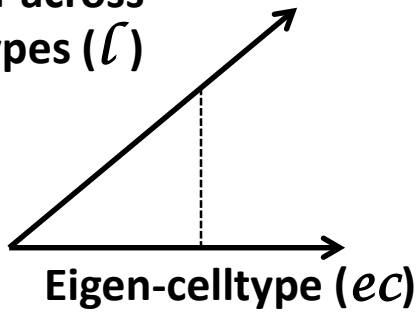


Genomic Signatures

eigen-celltype 2	RNA-binding	Cytoskeletal	Organelle Lumenal

Projection correlation = $\ell \,||\, ec$

Gene/lectin vector across cell types (ℓ)



Eigen-Celltype 2 lectin profile is *N*-link maturation related.

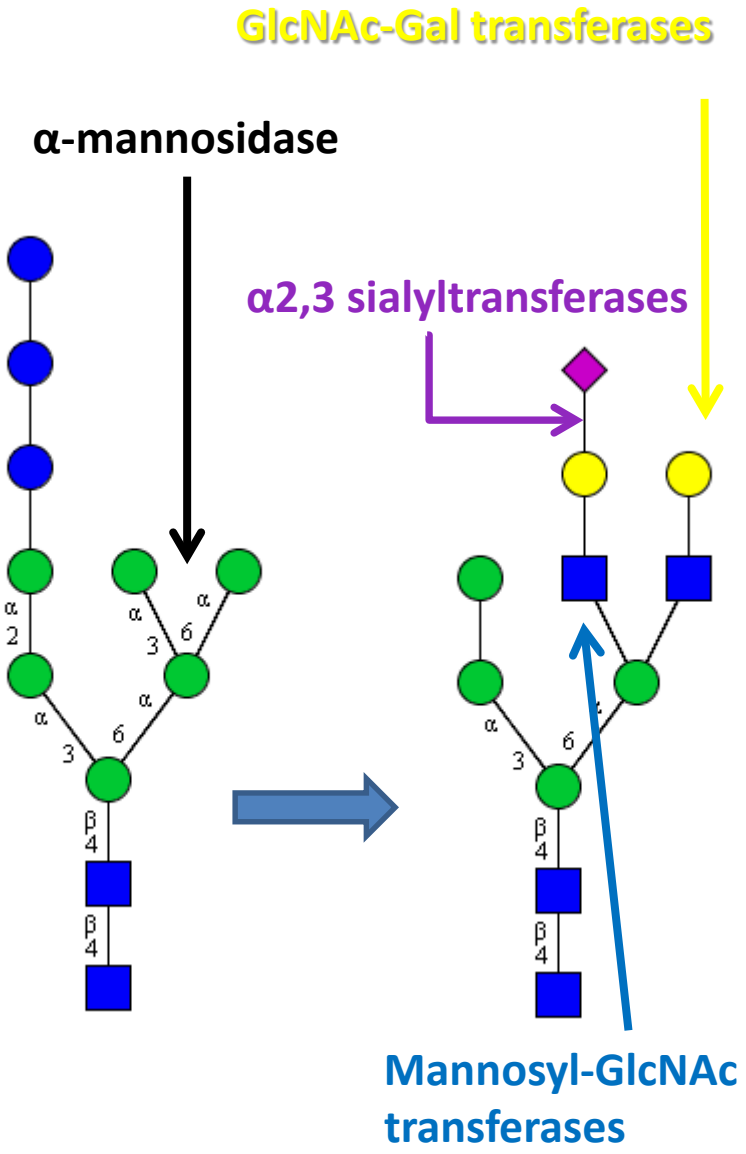
Do gene associates suggest pathway for *N*-linked maturation?

GSVD Matrix

$$M = U \Sigma V^T$$

Lani Pilobello

Eigen-celltype 2 genes: Transferases



α-mannosidase – formation of hybrid and complex glycans

-cleaves mannose residues to $Man_3GlcNAc_2$

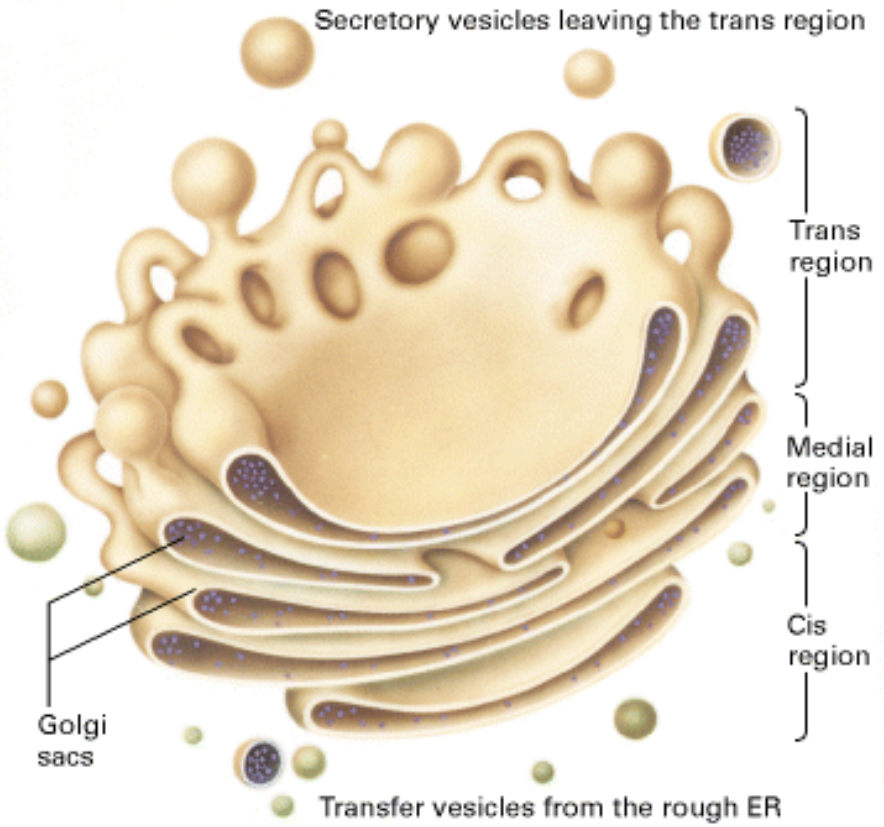
mannosyl-GlcNAc transferase – formation of hybrid and complex glycans

-adds GlcNAc's to cleaved mannose core

GlcNAC-Gal transferases – formation of LacNAc structures on hybrid, complex, *O*-linked

α2,3- sialyl transferases – hybrid, complex, O-linked, glycolipid

Eigen-celltype 2 genes: Golgi Structural Genes



http://php.med.unsw.edu.au/

Golgins

-GOLGA1, GOLGA3, GOLGA4, GOLGA5

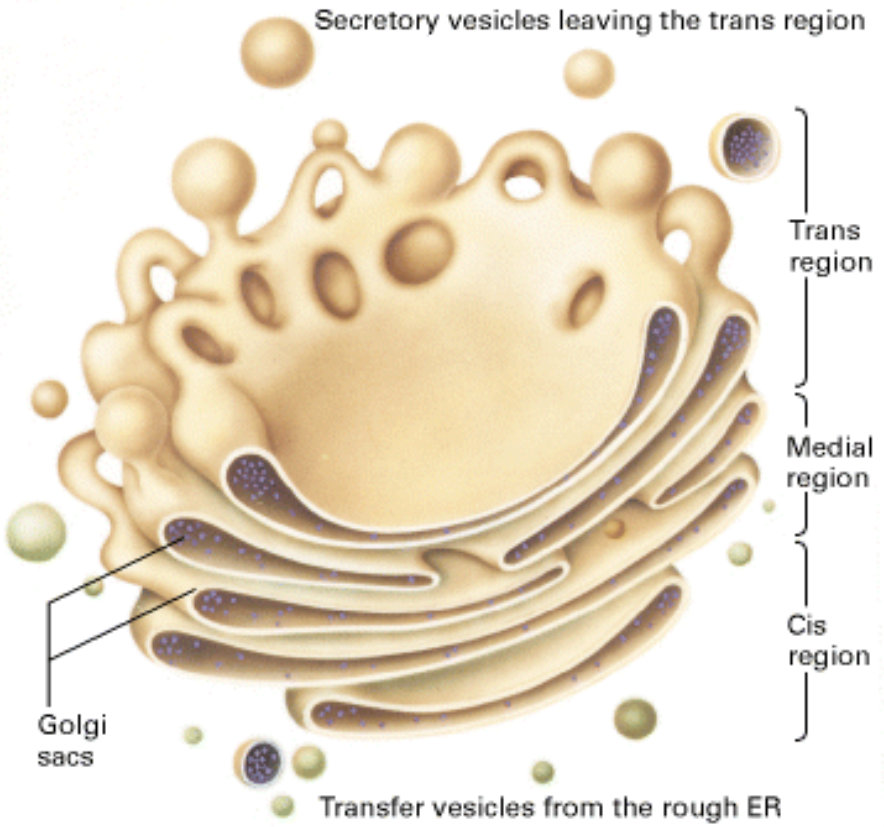
-*trans* Golgi associated
-recruit cytoskeletal proteins to form Golgi structure
-Recruit RAB family GTPases

RAB GTPases

-Rab1a, Rab1b, Rab2a, Rab7a, Rab9, Rab10, Rab11, Rab21, Rab27, Rab34

-Organize Golgi transporters and associated proteins via golgins

Eigen-celltype 2 genes: Golgi Vesicular Genes



http://php.med.unsw.edu.au/

Coatomer Protein Complex

-COPα, COPβ1,2, COPγ, COPζ, Arf1

-Forms complex that coats vesicles

-anterograde and retrograde transport through Golgi

-Transports proteins, lipids and other cargo between organelles

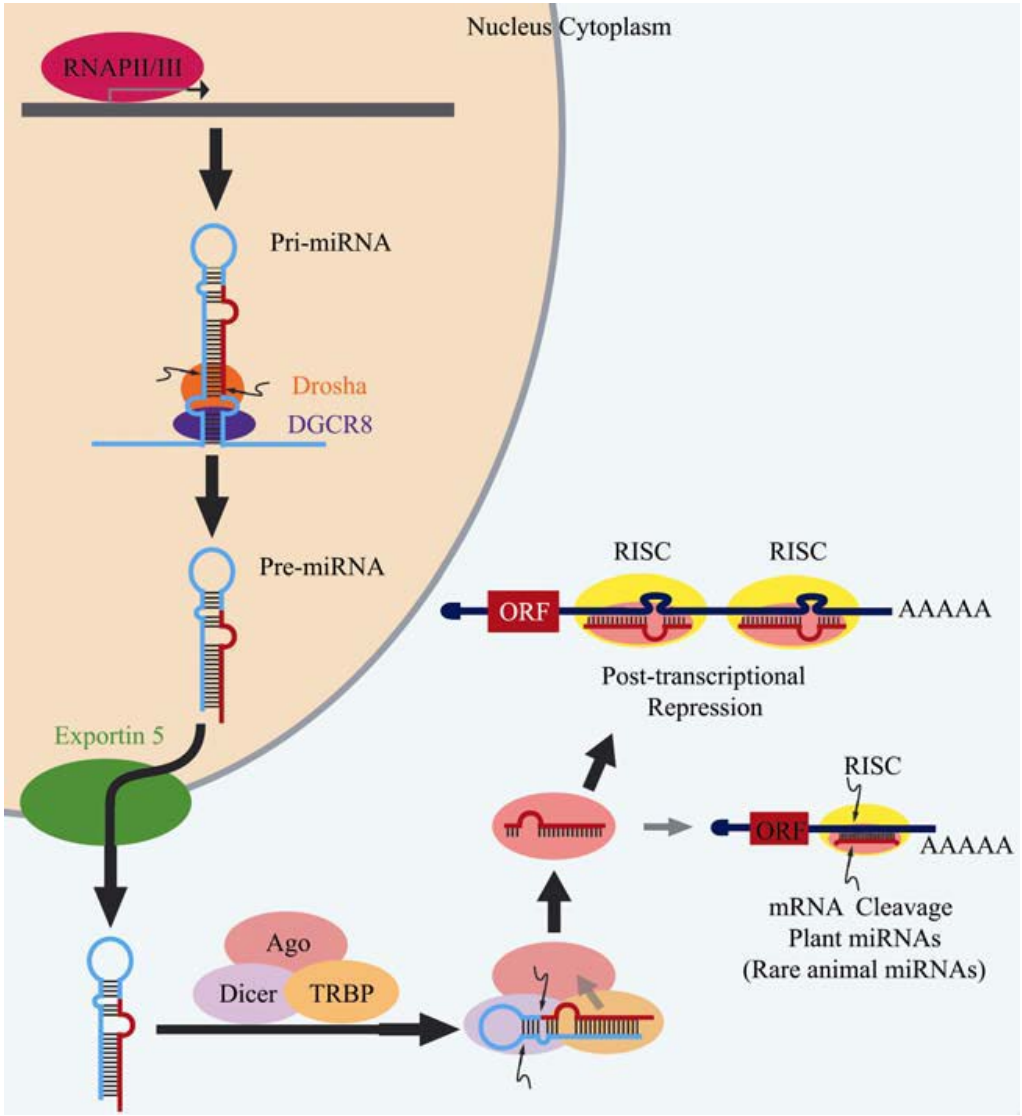
Observations

- ***N*-glycan maturation through the Golgi is a distinctive glycomic determinant in the NCI-60**
- **Glycosylation genes co-vary with associated lectins but…**
- **Golgi structural and transport genes also co-vary**
- **Implication that Golgi structure and localization mediates a direct effect on the glycome**

Currently Ongoing

- **We have analyzed the transcriptome of the NCI-60 panel with a custom-made gene array**
- **We are investigating whether Golgi localization and structure affects the glycome in a non-glycosyltransferase dependent manner**
- **We are interested in miRNAs which may affect expression patterns that control the glycome**

microRNAs (miRNA)

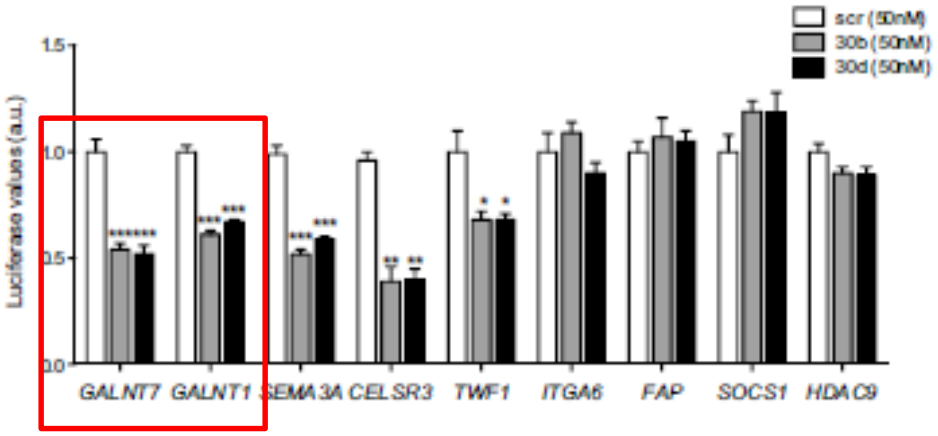


- **miRNA are genomically encoded short (~22 nt) molecules involved in repressing expression of mRNA**
- **They are processed and recruit a silencing complex which binds to the 3' untranslated region of mRNA and inhibit translation of the message to protein**
- **They affect many cellular processes, including glycosylation**

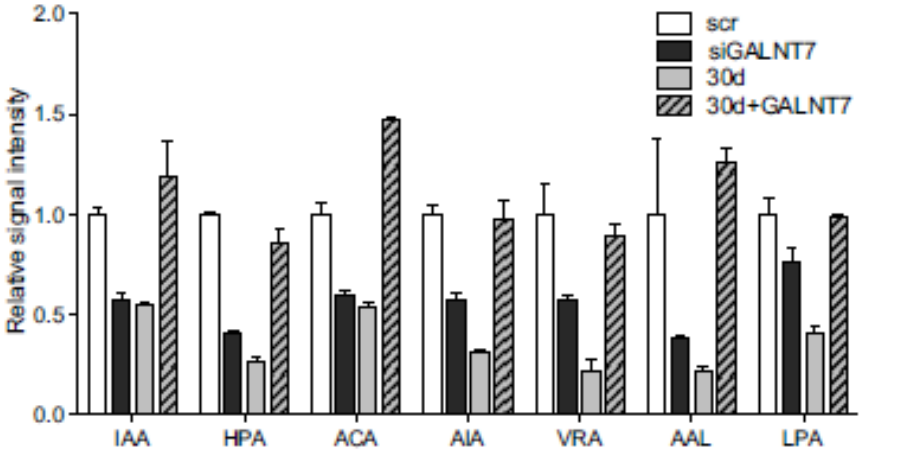
Cong et al, Springer Press, 2009

miR-30b/d repress *O*-linked transferases, promoting melanoma invasion and immunosupression

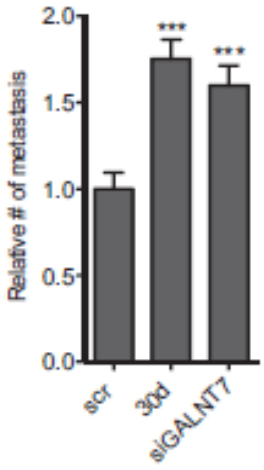
miRNA regulation GalNT expression

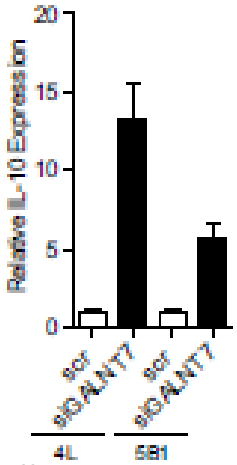


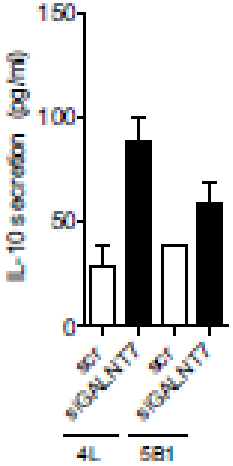
GalNT7 repression affects *O*-linked glycome



Metastasis and Immunosuppresion Increase







Gaziel-Sovran et al, *Cancer Cell*, 2011

-We have demonstrated miRNA regulation of a glycosyltransferase

-This regulation affects the glycome, promotes invasion and immunorepression

-How else may miRNAs be controlling the glycome?

Conclusions

- **We have applied a microarray strategy to the analysis of the glycomes of nearly 60 cell lines**
- **We have integrated glycomic and genomic array methods**
  - **This analysis has identified gene expression patterns which co-vary with glycomic expression**
- **We are investigating previously un-considered Golgi structural pathways and miRNA regulation of the glycome**

Future Directions

- **Analysis of the complete NCI-60 glycome/genome array data set is ongoing**
- **We are pursuing miRNA regulation of the melanoma glycome (collaboration with NYU Medical School)**
- **We are investigating differential regulation of *N*- and *O*-linked pathways**
- **We are investigating the role of the Golgi structure in determining the glycome**

Acknowledgements

Prof. Lara Mahal

Current group members

-Dr. Praveen Agrawal*

-Dr. Bianca Batista*

-Dr. Kanoelani Pilobello*

-Dr. João Ribeiro

-Linlin Wang

-Yaxuan Liang*

-Tomasz Kurcon

-Zhongyin Liu

-Will Eng*

-Sarah Abbassi (undergraduate)

Former group members

-Dr. Ku-Lung Hsu*

-Dr. Lakshmi Krishnamoorthy*

-Prof. Eva Hernando (NYU Med)

-Avital Gaziel-Sovran (NYU Med)

NATIONAL INSTITUTES OF HEALTH

NIH DIRECTOR'S NEW INNOVATOR AWARD