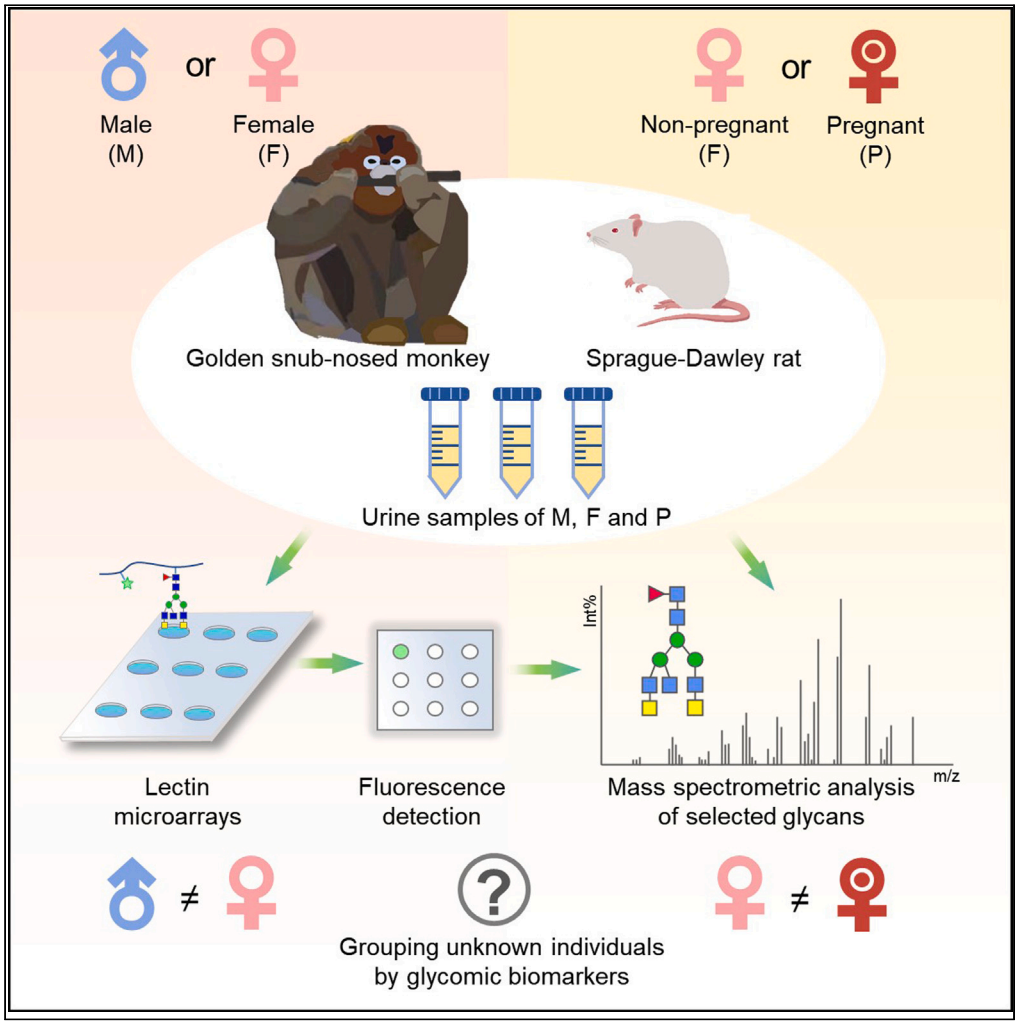
iScience

CellPress
OPEN ACCESS

Article

Pregnancy diagnosis and sex identification with urinary glycopatterns of two mammal species



Jiayi Yang, Shiyu Jin, Simeng Zhang, ..., Changhong Shi, Baoguo Li, Songtao Guo

songtaoguo@nwu.edu.cn

Highlights

Glycomic-based urinary biomarkers can identify pregnancy and sex for mammals, especially for rare species noninvasively

Lectin microarray combined with mass spectrometry could be applied on pregnancy diagnosis and sex identification in non-human primates

There are differences in N-glycans between pregnant & non-pregnant females, and between females & males

Several lectins could discern individual at different status as indicators, which are species-specific

Yang et al., iScience *26*, 108439
December 15, 2023
https://doi.org/10.1016/j.isci.2023.108439

iScience

CellPress
OPEN ACCESS

Article

Pregnancy diagnosis and sex identification with urinary glycopatterns of two mammal species

Jiayi Yang,[1] Shiyu Jin,[1] Simeng Zhang,[1] Shujun He,[2] Rui Chen,[1] Yinghu Lei,[3] Ruliang Pan,[1,4] He Zhang,[1] Zhongfu Wang,[5] Zheng Li,[6] Hanjie Yu,[6] Guixin Dong,[7] Changhong Shi,[8] Baoguo Li,[1] and Songtao Guo[1,9,10,*]

SUMMARY

Glycome in urine could be promising biomarkers for detecting pregnancy diagnosis and sex noninvasively for animals, especially for rare species. We explore the applicability of grouping golden snub-nosed monkeys by sex or diagnosing pregnancy based on their urinary glycopatterns, which are determined via lectin microarray combining mass spectrometry analysis. Sprague-Dawley rats are used to verify whether this approach and whether the glycomic biomarkers can be generalized to other mammalian species. The results show that, for both species, lectin microarray combining mass spectrometry can distinguish individuals' pregnancy status and sex; significant differences are found in the types, amounts, and terminal modification of glycans between pregnant and non-pregnant females and between females and males. This indicates the approach could be generalized to other mammalian species to group sex and detect pregnancy, yet the glycopatterns appear to be species-specific and markers developed from one species may not be directly applicable to another.

INTRODUCTION

Pregnancy and sex-related metabolites or genetic substances in the animal urine can be measured to provide a reference for indication. Hormonal test based on estrogen, progesterone, and gonadotropin is a common pregnancy detection method which has been applied in primates (e.g., bonobo, *Pan paniscus*;[1] pileated gibbon, *Hylobates pileatus*;[2] hanuman langur, *Presbytis entellus*)[3] and non-primate mammalian species (e.g., cows,[4] cats;[5] and *Lynx pardinus*).[6] However, the accuracy of hormonal test may rely on long-term monitoring and be unsensitive toward early-stage pregnancy such as the use of estradiol-3-glucuronide and pregnanediol-glucuronide in golden snub-nosed monkeys.[7] Specific antigen-immunoassay could also be used for the urine pregnancy test, but it requires development for species-specific antigens.[8–11] The time and expenses during this process is costly and even unpredictable for endangered and rare species.[11,12] As to genetic sex determination via urine, the limited amount of DNA would cause a high ratio of false positive results[13] and are of high cost.

A potential method to detect pregnancy of wild animals is using glycoprotein as biomarkers. Protein glycosylation is one of the most common post-translation modifications and the major types in mammals are N-glycosylation and O-glycosylation.[14] Glycans contribute in various biological processes, and the types of glycans and their expression level can reflect certain physiological and pathological status.[15] Studies have found pregnancy associated glycoproteins (PAGs) produced in placenta as a specific marker to indicate pregnancy.[16] Some serum PAG testing can already be highly accurate in providing a true positive reading for cattle.[17] PAG level in non-invasive sample such as milk is found to be positively related with blood PAG level in goats.[18] Researchers also study the pregnancy associated changes of serum and IgG N-glycosylation in camel (*Camelus dromedarius*) as an auxiliary diagnostic tool.[19] Even though the present application of glycan-related patterns such as PAGs or glycosylation as pregnancy markers is mostly invasive, comparing urinary glycopatterns differences among various reproductive status is a useful endeavor. Similarly, glycan-based urine analysis could be a promising way to discern sex, for studies on human found difference of urinary glycopatterns composition between two sexes.[20]

[1]Shaanxi Key Laboratory for Animal Conservation, College of Life Sciences, Northwest University, Xi'an 710069, China
[2]Shaanxi Institute of Zoology, Xi'an 710032, China
[3]Qinling Giant Panda Research Center (Shaanxi rare wildlife rescue base), Shaanxi Academy of Forestry, Xi'an 710402, China
[4]School of Human Sciences, The University of Western Australia, Australia and Centre for Evolutionary Biology, College of Biological Sciences, The University of Western Australia, Perth, WA 6009, Australia
[5]College of Food Science and Technology, Northwest University, Xi'an 710069, China
[6]Laboratory for Functional Glycomics, College of Life Sciences, Northwest University, Xi'an 710069, China
[7]Guangdong South China Rare Wild Animal Species Conservation Center, Zhuhai 511430, China
[8]Laboratory Animal Center, The Air Force Medical University, Xi'an 710032, China
[9]Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810008, China
[10]Lead contact
*Correspondence: songtaoguo@nwu.edu.cn
https://doi.org/10.1016/j.isci.2023.108439

Table 1. Comparison of p values and fold changes of lectins' NFIs in analysis of GSMs' urinary glycopatterns

	Fold change[a]		p values in Mann-Whitney U test	
Lectins	P/NP	F/M	P vs. NP	F vs. M
GSL-II	2.0813	–	0.0389	–
BS-I	2.4770	–	0.0201	–
PSA	3.0850	–	0.0201	–
PHA-E+L	0.3001	–	0.0201	–
DBA	0.3739	2.9517	0.0201	0.0495
LCA	–	0.3964	–	0.0495
PNA	–	0.2084	–	0.0463
VVA	–	0.2523	–	0.0495
MAL-I	–	2.6454	–	0.0495
BPL	–	2.9415	–	0.0495
PTL-I	–	4.7944	–	0.0463
RCA120	–	3.2916	–	0.0495
STL	–	3.1469	–	0.0495
SBA	–	2.1656	–	0.0495

[a]The fold changes of average normalized fluorescence intensities (NFIs) of lectins are compared between the pregnant (P) and non-pregnant (NP) groups, and between female (F) and male (M) groups. Lectins with NFIs fold changes greater than 2 or less than 0.5 times are selected as it presents in this table.

Lectin microarray is a tool to analyze glycoprotein glycan profiles with high flux and high sensitivity. Further analysis of glycan structures can be accessed by mass spectrometry, which also has high sensitivity, low sample consumption and can analyze complex compounds. Therefore, the approach of combining lectin microarrays and mass spectrometry will reduce the amounts of sample needed to get the informative results, which is particularly handy for wild and rare species that are difficult to get enough samples. Also, if specific biomarkers can be found, economic-effective test strips or kits can be produced and used widely. In this study, we want to answer whether this approach can find identifiers or biomarkers for pregnancy detection and sex identification on a wild non-human primate. Also, we hope to explore whether the identifiers and biomarkers found are consistent among species so that the markers found in one species can be generalized to other mammalian species.

This study utilized lectin microarrays combined with mass spectrometry (MALDI-TOF/TOF-MS) to discover urinary glycopatterns that can accurately group and characterize individuals based on their pregnancy and sex on golden snub-nosed monkeys (GSM). The golden snub-nosed monkey (*Rhinopithecus roxellana*), an endangered Old World monkey of *Colobine* subfamily, is endemic to China. The species needs an effective method for pregnancy diagnosis and sex determination for better conservation and therefore it meets the requirements for the study. The wild and captive groups of GSM are well habituated with clear adult age-sex characters.[21–23] GSMs are naturally social animals with one-male-units (OMUs) as the basic breeding units, while female and male individuals are sometimes paired together in captive settings, detection of pregnancy diagnosis earlier can practically help prevent mating damage, miscarriage or stillbirth[24] and provide the individual with nutrition and cares needed.[25] Meanwhile, although GSM is sexual dimorphic, there are occasions that observations are not feasible for the youth. The sex of juvenile (under around 3.5 years old) and infant GSM individuals could be hard to tell by eyes especially in the field study.[23,26] Their urine can be collected at a close distance whenever in the wild or captive.[27,28] This allows researchers to successfully obtain relatively clean and adequate urine samples easily.

This study investigates the different types of urinary glycans in GSM pregnant female (P), non-pregnant female (NP) for pregnancy diagnosis, and female (F) and male (M) for sex identification. Urine samples are first analyzed by lectin microarrays to find lectin combinations that reflect the types of glycans contained for each group. Next, the N-glycan profiles under one lectin for each group are investigated through mass spectrometry (MALDI-TOF-MS). Combining the two-step results, we aim to develop a novel glycomic-based biomarker system to detect pregnancy and identify sex of GSMs. Furthermore, to confirm that this approach can be generalized across different mammalian species to develop glycomic biomarkers for the same purposes, we use Sprague-Dawley (SD) rats as a verifying model as they are a common experimental animal model in mammals, which is readily accessible to individuals of different status. The SD rats are likewise divided into two groups for each purpose (P&NP/F&M) with identical measures carried out. The results of SD rats also reveal the variation and specification of glycan-patterns between different mammalian species, which answer the question of glycomic biomarkers' generalization across species. In addition to selecting positive results of lectin microarray, we selected a lectin with negative result in the lectin microarray results in the F and M group of GSM for mass spectrometry experiment. This was used to verify whether mass spectrometry analysis can identify the differences of N-glycans between groups? This research will establish an extra tool in urine analysis method, which could benefit the conservation and management for GSM and other endangered wild mammals.

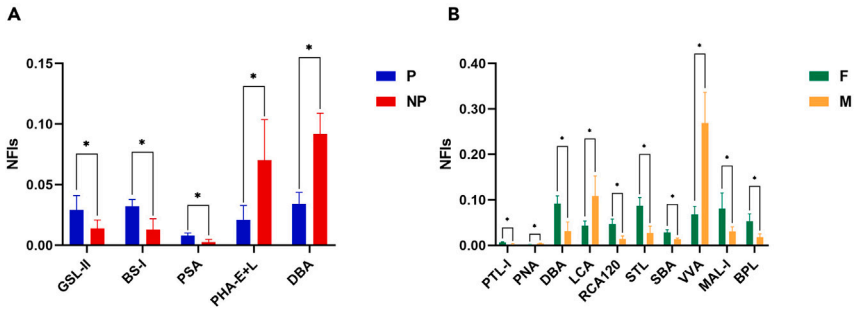


Figure 1. Lectins screened by microarray analysis of GSM

(A) Five lectins' NFIs that show significantly different level by Mann-Whitney U test, and fold changes between pregnant (P) and non-pregnant (NP) females are more than 2 or less than 0.5 times.

(B) Ten lectins' NFIs that show significance difference by Mann-Whitney U test, and fold changes between females (F) and males (M) are more than 2 or less than 0.5 times. *: $p < 0.05$.

RESULTS

Lectin microarray analysis of golden snub-nosed monkeys urine

Through the scanning the lectin microarrays of processed GSM urine samples, we obtained the fluorescent intensity of each spot and calculated the normalized fluorescent intensities (NFIs) for each urine sample. The names of lectins icons are listed in Table S1 [The information of 37 lectins] and the NFIs values are in Table S2 [The specific NFIs of every GSM sample], Related to Table 1 and Figure 1. There are two lectins, LEL and VVA, have average NFIs higher than 0.05 in all groups, indicating that the binding glycan types (see Table S1 [The information of 37 lectins]) of each lectin are generally abundant in GSM's urine. When it comes to each group with lectins that have NFIs higher than 0.05, there are 6 types of glycans for the P group, 7 types for the NP group, and 5 types for the M group. Regarding lectins that show weak signals with NFIs less than 0.01, there are 9 types of lectins for the P group, 14 types for the NP group and 15 types for the M group (Table S2 [The specific NFIs of every GSM sample], Related to Table 1 and Figure 1). The corresponding glycans for these lectins are considered rare in respective groups.

We compared the differences of lectins' NFIs between the P and NP group and between the NP and M group separately through Mann-Whitney U test. The results found that 12 lectins show significant difference ($p < 0.05$) in terms of NFIs between the P and NP groups, and 11 lectins between the F and M groups. These significantly different lectins were then selected according to their NFIs fold changes between groups. Lectins with NFIs fold changes greater than 2 or less than 0.5 times were selected. A total of 5 lectins were selected for the P and NP comparison, including PHA-E+L and DBA with higher NFIs in the NP group and GSL-II, BS-I and PSA higher in the P group. As for F & M comparison, 10 lectins were selected with TL-I, DBA, RCA120, STL, SBA, MAL-I, and BPL exhibiting higher NFIs in the F group and PNA, LCA and VVA being higher in the M group (Table 1; Figures 1A and 1B).

Lectin microarray analysis of Sprague-Dawley rat urine

The NFIs of rat urine sample are listed in Table S3 [The specific NFIs of every SD rat sample], Related to Tables 2, 3 and Figure 2. The results show that SD rats present a clearly different pattern than GSMs. Overall, AAL and SNA show strong signals (average NFIs >0.05) in all groups. There are 4 types of lectins in the rat NP group, 7 types totally in the P group, and 6 types totally in the M group that have NFIs higher than 0.05. The number of lectins with weak signals (NFIs <0.01) in each group is 11 for NP, 14 for P, and 19 for the M group.

Through statistical analysis and comparison of NFIs fold changes between groups, we found 10 lectins for P & NP and 8 lectins for F & M that meet the requirements of selection (i.e., showing a significant difference, $p < 0.05$; fold changes greater than 2 or less than 0.5 times). Between the P and NP groups, Jacalin and VVA have higher values in P and WFA, HHL, PTL-I, EEL, MPL, LEL, LCA, and PWM have higher values in NP; between the F and M groups, Jacalin has higher value in F while HHL, PTL-I, SJA, EEL, MPL, LCA, and PWM higher in M ($p < 0.05$, Table 2; Figures 2A and 2B).

Method verification analysis with Sprague-Dawley rats data

To verify whether the results of lectin microarrary analysis can be used practically to differ sex of individuals or to detect pregnancy, we further tested 37 lectin NFI data of 4 unknown SD rat individuals and compared the results with the group data obtained in the previous steps to classify them. The tested SD rats were numbered as t1 to t4 respectively (Table S3 [The specific NFIs of every SD rat sample], Related to Tables 2, 3 and Figure 2). Since in practice we should know what the unknown individuals are tested and classified for (whether it is a female or male or whether it is pregnant or not pregnant), we preserved partial group information when sampling the four tested SD rats. Therefore, before the classification analysis was conducted, the information we already had is: t1 and t2 are either male or non-pregnant female and they are tested to identify the sex; t2, t3 and t4 are either pregnant or non-pregnant female and they are tested to detect pregnancy. We would like

Table 2. Comparison of p values and fold changes of lectins' NFIs in analysis of SD rats' urinary glycopatterns

	Fold change[a]		p values between the groups	
Lectins	P/NP	F/M	P vs. NP	F vs. M
LEL	0.1885	–	0.0032	–
VVA	4.0939	–	0.0272	–
WFA	0.0206	–	0.0171	–
Jacalin	3.7324	0.4234	0.0032	0.0209
HHL	0.2527	11.3438	0.0122	0.0180
PTL-I	0.2527	4.1246	0.0032	0.0209
EEL	0.4015	3.2141	0.0092	0.0209
MPL	0.3474	3.4855	0.0092	0.0433
LCA	0.1654	3.1420	0.0032	0.0209
PWM	0.1321	11.3278	0.0032	0.0209
SJA	–	4.2946	–	0.0209

[a]The fold changes of average normalized fluorescence intensities (NFIs) of lectins are compared between the pregnant (P) and non-pregnant (NP) groups, and between female (F) and male (M) groups. Lectins with NFIs fold changes greater than 2 or less than 0.5 times are selected as it presents in this table.

to mention that although we used t2 twice here and may infer its group, the classification analyses themselves are independent and the group of each subject should be unknown previously.

Based on results of SD rats' lectin microarray analysis, the Euclidean distances (ED) between the tested individual and the group value were calculated and compared (Table 3). Smaller ED refers more similarity between the individual and the group. The results show that in the sex identification group, t1 has a smaller ED to the group M and t2 has a smaller ED to the F group, which suggests that t1 can be assigned to the group M and t2 to the group F; in the pregnancy detection group, t2 has smaller ED to the group NP and t3, t4 have smaller ED to P, which suggests that t2 can be assigned to the group NP and t3, t4 to P. Compared with each tested individuals' identity, we found grouping of every individual is consistent with their real status, which indicates the method is valid and applicable.

Mass spectrometry analysis of golden snub-nosed monkeys' urinary N-glycan profile

We chose BS-I to conduct mass spectrometry analysis of their binding N-glycans. Also, BS-I showed large difference in the group P and NP in GSM (Fold change = 2.4770, p = 0.0201), lectin PSA showing a bigger difference in the group P and NP than BS-I, but this lectin usually recognizes core-fucosylated, the relationship between fucosylation and pregnancy has been reported in relevant articles,[29,30] therefore, we chose BS-I to enrich glycoproteins and further analyze about the structure of related N-glycan for GSM. The N-glycans were obtained by first isolating the glycoproteins from the animal urine and then purification. The mass spectrometry analysis of BS-I binding N-glycans shows that there are 35 N-glycans of different structures identified and annotated in the P group, and 26 identified and annotated in the NP group. All the N-glycans are specifically recognized by BS-I (Figures 3A and 3B) and their proposed structures are presented in Table S4 [The N-glycan information of isolated glycoproteins from GSM pregnant females and non-pregnant females], Related to Figures 3A and 3B. Among these N-glycans, 20 of them were observed in both groups, 15 were observed only in the P group, and 6 were observed only in the NP group. In the P group, 15 types of N-glycans were found modified by sialic acid at the end, accounting for 42.86% of the total glycans identified; 23 N-glycans had fucosylation modification, accounting for 65.71% of the total (Table S5 [Modification of N-glycan by mass spectrometry of GSMs]). In the NP group, the N-glycans with sialylation modification (8 types) accounted for 30.77% of the total glycans; the N-glycans with fucosylation modification (15 types) accounted for 57.69% (Table S5 [Modification of N-glycan by mass spectrometry of GSMs]).

Table 3. The Euclidean distances (ED) between tested individuals (t1, t2, t3 and t4) and the group (M&F; P&NP)

	ED			*ED*		
Subject	M	F	Grouping[a]	P	NP	Grouping[a]
t1	0.0192	0.0614	M	–	–	–
t2	0.0827	0.0571	F	0.1107	0.0920	NP
t3	–	–	–	0.0196	0.1240	P
t4	–	–	–	0.0605	0.1137	P

[a]Smaller ED indicates the individual is more similar to the group, which produces the results of grouping.

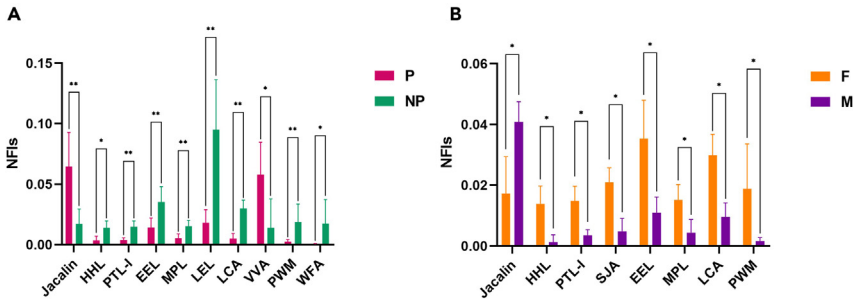


Figure 2. Lectins screened by microarray analysis of SD rats

(A) Ten lectins' NFIs showing significantly different level by Mann-Whitney U test, and fold changes between pregnant females and non-pregnant females are more than 2 or less than 0.5 times.

(B) Eight lectins show significantly different level by Mann-Whitney test, and fold changes between non-pregnant females and males are more than 2 or less than 0.5 times. *: $p < 0.05$, **: $p < 0.01$.

As for F & M comparison, 26 and 24 N-glycans were identifies and annotated in the F group and the M group respectively. All of these N-glycans are specifically recognized by BS-1 (Figures 3B and 3C) and their proposed structures are shown in Table S6 [The N-glycan information of isolated glycoproteins from GSM females and males], Related to Figures 3B and 3C. Nineteen N-glycans were observed in both F and M groups; 7 N-glycans were observed only in the F group; 5 N-glycans were observed only in the M group. The N-glycans in the M group have sialylation modification (9 types) accounting for 37.50% and fucosylation modification (14 type) 58.33% of the total glycans (Table S5 [Modification of N-glycan by mass spectrometry of GSMs]). The results of the F (NP) group have been reported above.

Mass spectrometry analysis of Sprague-Dawley rats' urinary N-glycan profile

As to the mass spectrometry analysis for SD rats, we found Jacalin shows the biggest difference in the P&NP groups (Fold change = 4.0703, $p = 0.0011$) and is also significantly different in the F& M group (Fold change = 0.3574, $p = 0.0090$), which is rarely reported in the literature. Therefore, Jacalin was chosen. There were 22 different types of N-glycans identified in the P group, 20 of which are specifically recognized by Jacalin; 11 N-glycans were identified in the NP group, 8 of which are specifically recognized by Jacalin. The identified N-glycans and their proposed structures are presented in Figures 4A, 4B and Table S7 [The N-glycan information of isolated glycoproteins from SD rat pregnant females and non-pregnant females], Related to Figures 4A and 4B. Three N-glycans were observed in both groups, whereas 19 N-glycans were observed only in P and 8 N-glycans were observed only in the NP group. In the P group, 40.91% (9 types) of N-glycans identified were found modified by sialic acid at the ends; 68.18% (15 types) of N-glycans had fucosylation modification. In the NP group, 18.18% of (2 types) N-glycans had sialylation modification; 72.73% (8 types) of N-glycans had fucosylation modification (Table S8 [Modification of N-glycan by mass spectrometry of SD rats]). In the M group, 28 N-glycans with different structure were identified, 26 of which are N-glycans specifically recognized by Jacalin. The identified N-glycans and their proposed structures are presented in Figures 4B, 4C and Table S9 [The N-glycan information of isolated glycoproteins from SD rat males and females], Related to Figures 4B and 4C. Comparing the N-glycans identified in the F and M group, 3 types were observed in both groups, 8 were only in the F group, and 25 were only in the M group. A proportion of 39.29% (11types) of total glycans identified in the M group had sialylation modification; 71.43% (20 types) had fucosylation modification (Table S8 [Modification of N-glycan by mass spectrometry of SD rats]).

DISCUSSION

We tested the pregnancy and sex-specific glycopatterns in two mammalian species (GSMs and SD rats) to diagnosis pregnancy and identify sex through urine analysis of glycomes. We found that lectin analysis combining mass spectrometry analysis of urine present distinct glycopatterns between different groups of each species. This suggests that our approach could be a useful method to detect pregnancy and identify sex of GSM and that it can be expand to other mammalian species for we have replicated the result in SD rats. We found glycomic biomarkers are species-specific for pregnancy and sex identification in two species. Pregnancy status and sex can be effectively grouped by glycans structure analysis, even if the markers of negative results shown in lectins microarrays.

Characterization of urinary glycopatterns between pregnant and non-pregnant individuals

The results of lectin microarray analysis found that the types and the amount of urinary glycans vary a lot between pregnant and non-pregnant individuals of two species. Notably, although a lectin can bind with several types of glycans and contribute similarly to the fluorescent intensity, the significant differences found in lectins can still reflect that individuals of different reproductive status have various types of glycan metabolites in urine. It is possible to use the lectins that present distinctive NFIs as identifiers to indicate whether the individual is pregnant or not. To find out the identifier, we selected the lectins whose NFIs show greater than 2 or less than 0.5 times differences between the P and NP groups. Based on the test of GSMs, BS-I, GSL-II and PSA, the binding glycans of which become more abundant in the urine when the

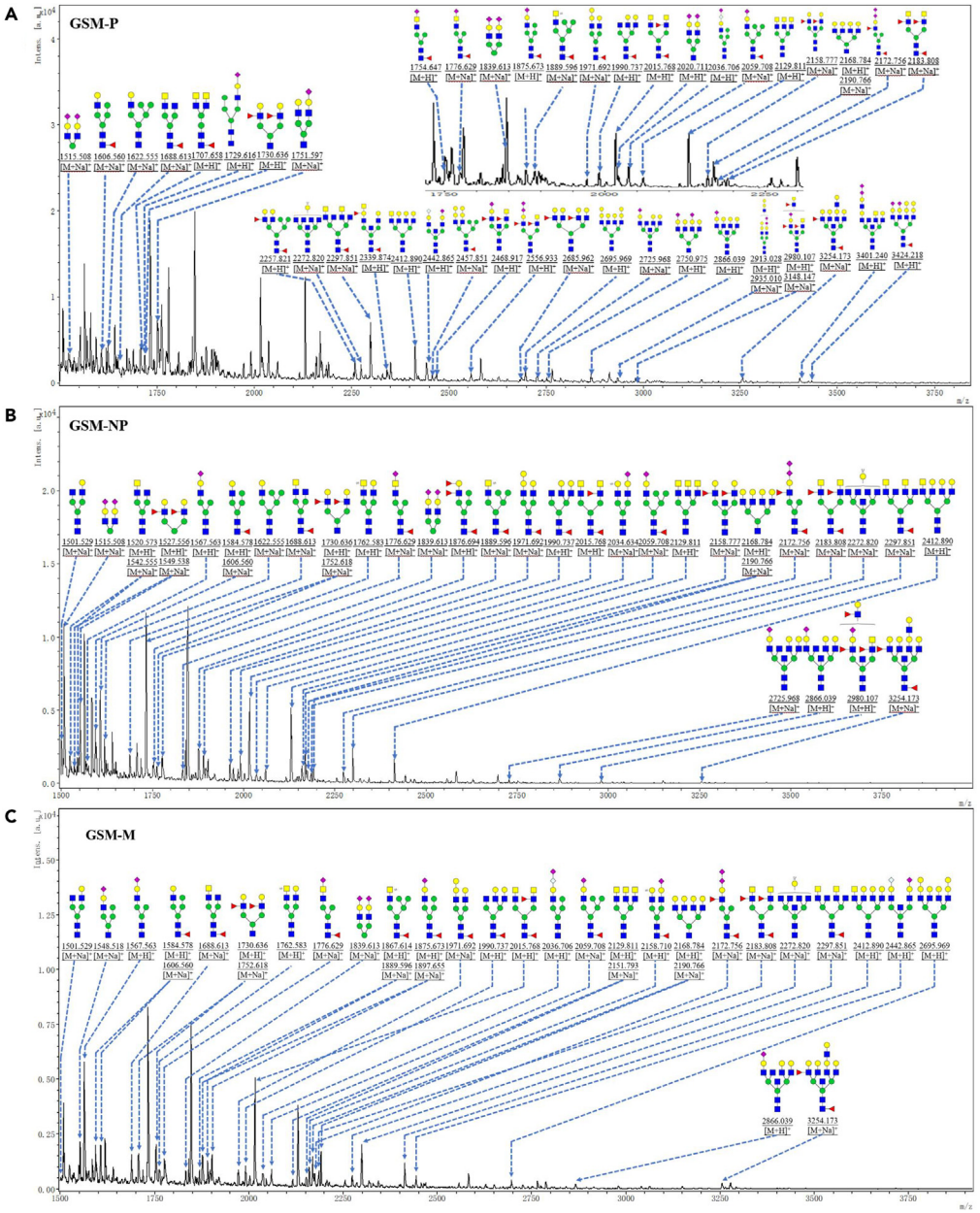


Figure 3. Mass spectra of BS-I recognized N-glycans of GSM

(A) Mass spectrum of BS-I recognized N-glycans in pregnant individuals (P) of GSM.

(B) Mass spectrum of BS-I recognized N-glycans in non-pregnant females (NP/F) of GSM.

(C) Mass spectrum of BS-I recognized N-glycans in males (M) of GSM.

female individuals are pregnant, as well as DBA and PHA-E+L, whose binding glycans are more abundant in non-pregnant females, may be potentially used in practice to differentiate pregnant and non-pregnant female GSMs. There are lectins can be selected for SD rats, but the types are different than that for GSMs, which is unsurprisingly owing to the species specification of urinary glycopatterns.

This study further investigated the characterization of N-glycans in urine. BS-I specifically binds with glycans of α-Gal, α-GalNAc, Galα-1,3-Gal, Galα-1,6Glc and this binder was chosen for GSM groups. Jacalin is the binder of Galβ1-3GalNAcα-Ser/Thr(T), GalNAcα-Ser/Thr(Tn), GlcNAcβ1-3-GalNAcα-Ser/Thr(Core3), sialyl-T(ST) and it was chosen for rat groups. The average NFIs of BS-I and Jaclin show significant differences, and the fold change are greater than 2 or less than 0.5 between the P and NP group of each species group. We classify N-glycans according to the terminal modification of the chosen lectin for each pair through MALDI-TOF/TOF-MS, and found that the proportion of glycans that are modified by sialic acid or had fucosylation are different. For GSMs, both types of glycosylation have higher proportion in the

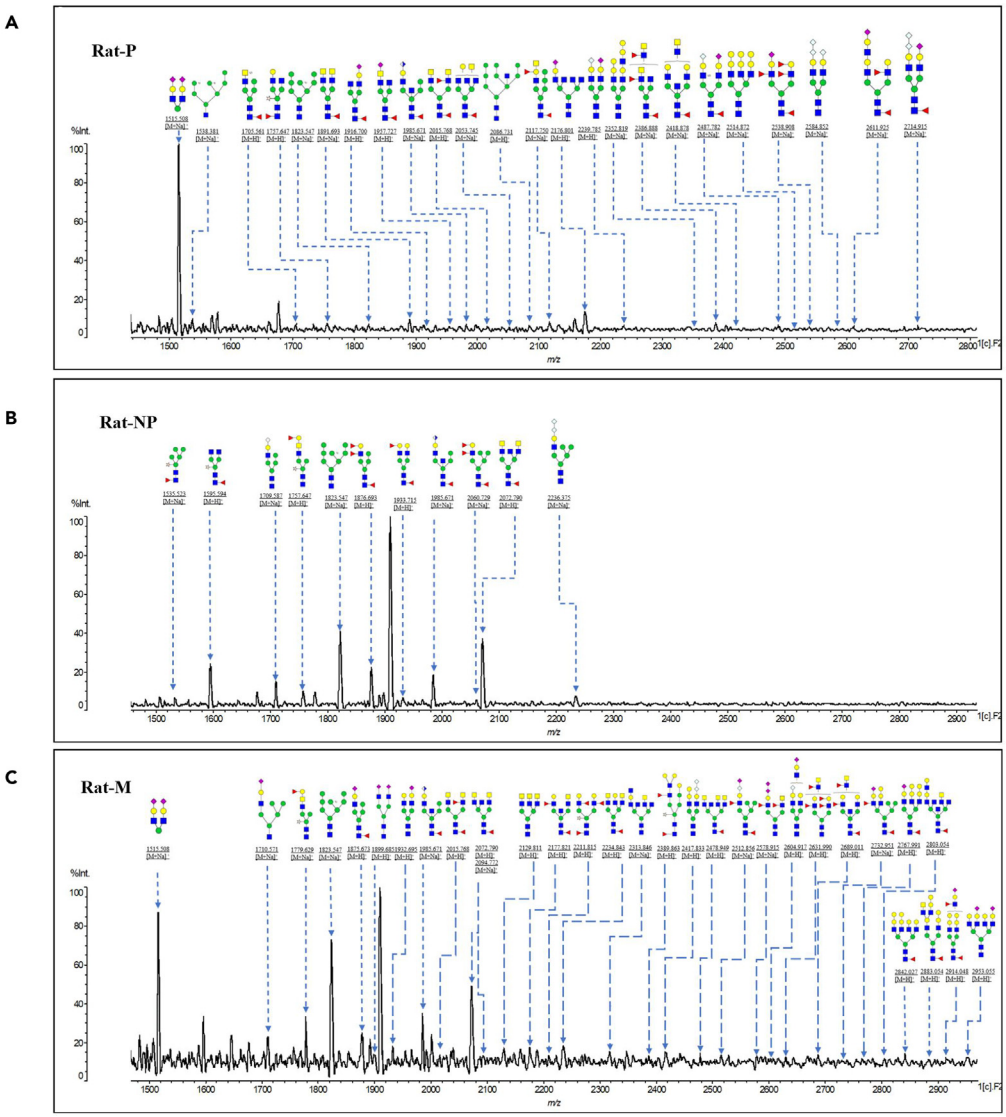


Figure 4. Mass spectra of Jacalin recognized N-glycans of SD rats

(A) Mass spectrum of Jacalin recognized N-glycans in pregnant (P) SD rats.

(B) Mass spectra of Jacalin recognized N-glycans in non-pregnant female (NP) SD rats.

(C) Mass spectrum of Jacalin recognized N-glycans in male (M) SD rats.

P group than NP; for SD rats, the proportion of glycans that had sialylation modification is higher in the P group while those had fucosylation modification are high and roughly the same for two groups.

We observed in GSMs and SD rats that glycans of pregnant individuals have more types and higher percentage of sialyation modification at the end; the percentage of fucosylation modification has slight differences between groups. Because of the genetic and antigenic difference between mother and offspring, maternal immune system is partially suppressed during pregnancy to protect the fetus, which is believed that glycosylation contributes a lot in the process.[31,32] This is because protein glycosylation, particularly the sialylation at the end, plays essential roles in regulating pro-inflammatory and anti-inflammatory immune responses.[33] The pattern of sialyation we found on the two species echoes the findings in human clinical studies using plasma, which found that α2, 3- and α2, 6- linked sialic acids on plasma glycans will increase during pregnancy, and decrease after gestation.[33–35] In the studies on other mammals such as camel (*Camelus fromedarius*), the proportion of structures with two galactosylated outer-arims, sialylation and core-fucosylation increase during pregnancy.[19]

On the other hand, studies on PAGs in pregnant mammal find that serum glycoprotein concentration may vary when the number, the sex and the birthweight of fetuses, whether pregnancy-related glycosylation undergoes similar influence needs further investigation.[36] Again, the glycosylation happens during pregnancy is specific to each species, and so should be the glycan markers discovered for pregnancy diagnosis.

Characterization of urinary glycopatterns between female and male individuals

In this study, we compared the male individuals only with females out of gestation period. Many physiological differences between male and female mainly come from the activity of steroid hormones produces by the gonads, and studies of human and other animals reveal that gene expression level are related to sex.[37,38] Apart from directly measuring sex hormones or extracting genetic information from samples for analysis, we hope to add a supplementary non-invasive approach to detect sex of GSMs.

Likewise, the lectin microarray analysis found there are 10 lectins can be used as identifiers to discern GSM's sex. The NFIs of BS-I did not show significant difference between the F and M groups of GSMs. However, the results of MALDI-TOF/TOF-MS that there are also different types of N-glycan and different modification between the F and M groups in GSM, reveal that mass spectrometry analysis is more sensitive to detect the different composition of glycans. A study on human serum found that the level of α-1-6-fucose glycans is higher in females, while the level of branching-α-1,3-fucose glycans is higher in males.[39] These results suggest that it is feasible to use animal body fluid samples for sex and gestation identification with glycomic approach. Nevertheless, as to GSMs, analysis of other lectins is needed to obtain the patterns of glycosylation for different sex as the differences may be linked to the lectin used.

Limitations of the study

Although we collected urine samples to reduce the problem caused by limited accessibility of the animal, the number of GSM individuals being sampled for this study is still small. In particular, GSMs have long gestation period and low reproductive success, it is hard to get a large sample size and reduce the impact of confounding factors or individual differences for a natural wild group. During the sampling process, we managed to control as many variables as possible, but factors such as exactly same age were still not considered here. In addition, this study has displayed the differences of urinary glycopatterns amid GSM individuals of different reproductive status or sex. The replicated test and verification analysis on SD rats suggested that our approach is valid and can be generalized to other species. However, we have not identified any specific glycan or glycoprotein markers that can be directly used. Our findings suggest that certain combinations of lectins that show significant NFI differences with more than 2 times fold changes between groups could be used in practice as possible pregnancy or sex indicators. This actually has some benefits, because lectin microarray is a high-throughput and commercially available method, which could be more feasible for animal conservation practitioners than the relatively more time-consuming and technically high-demanding mass spectrometry analysis only. Yet, lectin microarrays analysis just informs the range of many glycans and cannot provide the researchers detailed structures of glycan. Thus, this method may not be informative enough to enhance our comprehension of the structure, the role and the cause of glycans.

Conclusions

This study investigates the urinary glycopatterns of GSMs through lectin microarray and mass-spectrometric analysis to detect pregnancy and sex based on the results. SD rats were used as a verification. The results of two mammalian species show that the types, amounts and structure of N-glycans and the proportion of sialylation and fucoslation of N-glycans are different between pregnant and non-pregnant females, and between (non-pregnant) females and males. Several lectins could be used as indicators to discern individual at different status, but they are specific to each species. For mammals, especially for endangered species such as GSM, this method provides reference for pregnancy diagnosis and sex identification, which will benefit the management of the animals.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Urine sample collection
- METHOD DETAILS
  - Processing of urine samples
  - Lectin microarrays
  - Glycoprotein from urine by lectin-magnetic particle conjugates
  - Release the N-glycans by PNGase F
  - Purification of N-glycans
  - Characterization of the N-glycans by MALDI-TOF/TOF-MS
- DATA ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108439.

ACKNOWLEDGMENTS

This study was supported by the Natural Science Foundation of China (32220103002, 32370534); Strategic Priority Research Program of the Chinese Academy of Sciences, XDB31020302; Shaanxi Innovation Capability Supporting Project International Technology Cooperation Base; the Qinghai Province High-level Innovative "Thousand Talents" Program; Key program of Forestry Science Research of Shaanxi (SHLY-2018-07). We thank Colin A. Chapman for his editing on this article.

AUTHOR CONTRIBUTIONS

Conceptualization & methodology, S.G; formal analysis & investigation, J.Y and S.Z; writing-original draft S.G, J.Y, S.J, S.Z, S.H, R.C, R.P and H.Z; visualization J.Y, S.J, S.Z and R.C, writing-review & editing, J.Y, S.J S.Z and S.G; resource, Y.L Z.W, Z.L, H.Y, G.D and C.S; supervision, B.L and S.G.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 23, 2023
Revised: September 29, 2023
Accepted: October 16, 2023
Published: November 14, 2023

REFERENCES

1. Heistermann, M., Möhle, U., Vervaecke, H., van Elsacker, L., and Hodges, J.K. (1996). Application of Urinary and Fecal Steroid Measurements for Monitoring Ovarian Function and Pregnancy in the Bonobo (Pan paniscus) and Evaluation of Perineal Swelling Patterns in Relation to Endocrine Events1. Biol. Reprod. *55*, 844–853.
2. Aramaki, Y., Oae, H., Mouri, Y., Takizawa, K., Nakanishi, K., Hyoudou, Y., Omasa, M., and Kusunoki, H. (2010). Urinary estrogens, progesterone, and LH changes during normal menstrual cycles of a captive female pileated gibbon (Hylobates pileatus). J. Med. Primatol. *39*, 381–384.
3. Ziegler, T., Hodges, K., Winkler, P., and Heistermann, M. (2000). Hormonal correlates of reproductive seasonality in wild female hanuman langurs (Presbytis entellus). Am. J. Primatol. *51*, 119–134.
4. Dana, O.I., Ghaidan, M.T., Mukhtar, R.H., and Dyary, H.O. (2020). Comparison of a Barium Chloride Test with ELISA for Pregnancy Detection in Cows. J. Vet. Res. *64*, 567–571.
5. de Haas van Dorsser, F.J., Lasano, S., and Steinetz, B.G. (2007). Pregnancy diagnosis in cats using a rapid, bench-top kit to detect relaxin in urine. Reprod. Domest. Anim. *42*, 111–112.
6. Braun, B.C., Frank, A., Dehnhard, M., Voigt, C.C., Vargas, A., Göritz, F., and Jewgenow, K. (2009). Pregnancy diagnosis in urine of Iberian lynx (Lynx pardinus). Theriogenology *71*, 754–761.
7. Muren, Kusuda, S., Kusuda, S., Doi, O., Naito, H., and Hashikawa, H. (2017). Puberty, ovarian cycle, pregnancy, and postpartum ovulation in captive Sichuan golden monkeys (Rhinopithecus roxellana) based on changes in urinary and fecal gonadal steroid metabolites. Theriogenology *87*, 179–186.
8. González, F., Sulon, J., Garbayo, J.M., Batista, M., Cabrera, F., Calero, P., Gracia, A., and Beckers, J.F. (1999). Early pregnancy diagnosis in goats by determination of pregnancy-associated glycoprotein concentrations in plasma samples. Theriogenology *52*, 717–725.
9. Osborn, D.A., Beckers, J.-F., Sulon, J., Gassett, J.W., Muller, L.I., Murphy, B.P., Miller, K.V., and Marchinton, R.L. (1996). Use of Glycoprotein Assays for Pregnancy Diagnosis in White-Tailed Deer. J. Wildl. Manag. *60*, 388–393.
10. Karen, A., Beckers, J.F., Sulon, J., de Sousa, N.M., Szabados, K., Reczigel, J., and Szenci, O. (2003). Early pregnancy diagnosis in sheep by progesterone and pregnancy-associated glycoprotein tests. Theriogenology *59*, 1941–1948.
11. Zhao, W., Li, C., Hong, J., Lin, P., and Lei, Y. (2021). Preparation and Test of Enzyme Labeled Progesterone for Giant Panda. J. Anim. Sci. Vet. Med. *40*, 245–248. 252.
12. Tang, J., Li, J., Wang, B., Suo, L., Wang, Y., and Yan, X. (2018). Effect of Air Exposure Time on Progesterone Level of Feces in Forest Musk Deers. Prog. Vet. Med. *39*, 44–49.
13. Hrovatin, K., and Kunej, T. (2018). Genetic sex determination assays in 53 mammalian species: Literature analysis and guidelines for reporting standardization. Ecol. Evol. *8*, 1009–1018.
14. Raman, R., Raguram, S., Venkataraman, G., Paulson, J.C., and Sasisekharan, R. (2005). Glycomics: an integrated systems approach to structure-function relationships of glycans. Nat. Methods *2*, 817–824.
15. Haltiwanger, R.S., and Lowe, J.B. (2004). Role of glycosylation in development. Annu. Rev. Biochem. *73*, 491–537.
16. Wallace, R.M., Pohler, K.G., Smith, M.F., and Green, J.A. (2015). Placental PAGs: gene origins, expression patterns, and use as markers of pregnancy. Reproduction *149*, R115–R126.
17. Pohler, K., Araujo Franco, G., Reese, S., Guirado Dantas, F., Ellis, M., and Payton, R. (2016). Past, present and future of pregnancy detection methods. Appl. Reprod. Strat. Beef Cattle, 251–259.
18. Singh, S.P., Natesan, R., Sharma, N., Goel, A.K., Singh, M.K., and Kharche, S.D. (2019). Pregnancy-associated glycoprotein profile in milk and its relationship with the circulating level during early pregnancy in goats. Small Rumin. Res. *173*, 81–87.
19. Adamczyk, B., Albrecht, S., Stöckmann, H., Ghoneim, I.M., Al-Eknah, M., Al-Busadah, K.A.S., Karlsson, N.G., Carrington, S.D., and Rudd, P.M. (2016). Pregnancy-Associated Changes of IgG and Serum N-Glycosylation in Camel (Camelus dromedarius). J. Proteome Res. *15*, 3255–3265.
20. Li, H., Patel, V., DiMartino, S.E., Froehlich, J.W., and Lee, R.S. (2020). An in-depth Comparison of the Pediatric and Adult Urinary N-glycomes. Mol. Cell. Proteomics *19*, 1767–1776.
21. Guo, S.-T., He, S.-J., Zhang, H., Bai, R.-F., Zhang, S.-M., Hou, R., Grueter, C.C., Chapman, C.A., Dunn, D.W., and Li, B.-G. (2020). Male social rank and food competition in a primate multi-level society. Am. J. Phys. Anthropol. *173*, 630–642.
22. Zhang, P., Li, B., Wada, K., Tan, J., and Watanabe, K. (2003). Social structure of a group of Sichuan snub-nosed monkeys (Rhinopithecus roxellana) in the Qinling Mountains of China. Acta Zool. Sin. *49*, 727–735.
23. Liang, B., Qi, H., Zhang, S., and Ren, B. (2001). Developmental traits of captive sichuan snub-nosed monkeys (rhinopithecus roxellana) at different age stages. Curr. Zool. *47*, 381–387.
24. Yan, C.E., Jiang, Z., Li, C., Zeng, Y., Tan, N.N., and Xia, S.Z. (2003). Monitoring the menstrual cycle and pregnancy in the Sichuan golden monkey (Rhinopithecus roxellana) by measuring urinary estradiol and progesterone. Acta Zool. Sin. *49*, 693–697.
25. Lv, W. (2006). Progress on Sex Determination Using Mammalian Early Embryo. Prog. Vet. Med. *27*, 5–7.
26. Ren, R., Yan, K., Su, Y., Zhou, Y., Li, J.J., Zhu, Z., and Hu, Y. (2000). The Society of Golden Monkey (Rhinopithecus Roxellanae) [in Chinese] (Beijing University Press).
27. Fan, P., Chen, H., Yao, H., Wang, Z., Yang, J., and Xiang, Z. (2013). Measurement of urinary and fecal steroid metabolites in a provisioned group of golden snub-nosed monkeys (Rhinopithecus roxellana) at Shennongjia

Reserve, Hubei, China. Acta Theriol. Sin. *33*, 286–292.
28. Ren, B. (2002). Reproductive Behaviors of the Sichuan Snub-Nosed Monkeys, Rhinopithecus roxellana. PhD dissertation (Beijing Normal University).
29. Yu, M., Cui, X., Wang, H., Liu, J., Qin, H., Liu, S., and Yan, Q. (2019). FUT8 drives the proliferation and invasion of trophoblastic cells via IGF-1/IGF-1R signaling pathway. Placenta *75*, 45–53.
30. Yu, M., Wang, J., Liu, S., Wang, X., and Yan, Q. (2017). Novel function of pregnancy-associated plasma protein A: promotes endometrium receptivity by up-regulating N-fucosylation. Sci. Rep. *7*, 5315.
31. Szekeres-Bartho, J. (2002). Immunological relationship between the mother and the fetus. Int. Rev. Immunol. *21*, 471–495.
32. Trowsdale, J., and Betz, A.G. (2006). Mother's little helpers: mechanisms of maternal-fetal tolerance. Nat. Immunol. *7*, 241–246.
33. Jansen, B.C., Bondt, A., Reiding, K.R., Lonardi, E., de Jong, C.J., Falck, D., Kammeijer, G.S.M., Dolhain, R.J.E.M., Rombouts, Y., and Wuhrer, M. (2016). Pregnancy-associated serum N-glycome changes studied by high-throughput MALDI-TOF-MS. Sci. Rep. *6*, 23296.
34. Ruhaak, L.R., Uh, H.-W., Deelder, A.M., Dolhain, R.E.J.M., and Wuhrer, M. (2014). Total Plasma N-Glycome Changes during Pregnancy. J. Proteome Res. *13*, 1657–1668.
35. Bondt, A., Rombouts, Y., Selman, M.H.J., Hensbergen, P.J., Reiding, K.R., Hazes, J.M.W., Dolhain, R.J.E.M., and Wuhrer, M. (2014). Immunoglobulin G (IgG) Fab glycosylation analysis using a new mass spectrometric high-throughput profiling method reveals pregnancy-associated changes. Mol. Cell. Proteomics *13*, 3029–3039.
36. Lobago, F., Bekana, M., Gustafsson, H., Beckers, J.-F., Yohannes, G., Aster, Y., and Kindahl, H. (2009). Serum profiles of pregnancy-associated glycoprotein, oestrone sulphate and progesterone during gestation and some factors influencing the profiles in Ethiopian Borana and crossbred cattle. Reprod. Domest. Anim. *44*, 685–692.
37. Srivastava, A., Wang, J., Zhou, H., Melvin, J.E., and Wong, D.T. (2008). Age and gender related differences in human parotid gland gene expression. Arch. Oral Biol. *53*, 1058–1070.
38. Lukacs, J.R., and Largaespada, L.L. (2006). Explaining sex differences in dental caries prevalence: Saliva, hormones, and "life-history" etiologies. Am. J. Hum. Biol. *18*, 540–555.
39. Ding, N., Nie, H., Sun, X., Sun, W., Qu, Y., Liu, X., Yao, Y., Liang, X., Chen, C.C., and Li, Y. (2011). Human serum N-glycan profiles are age and sex dependent. Age Ageing *40*, 568–575.
40. Liu, X., Yu, H., Qiao, Y., Yang, J., Shu, J., Zhang, J., Zhang, Z., He, J., and Li, Z. (2018). Salivary Glycopatterns as Potential Biomarkers for Screening of Early-Stage Breast Cancer. EBioMedicine *28*, 70–79.
41. Yu, H., Li, X., Chen, M., Zhang, F., Liu, X., Yu, J., Zhong, Y., Shu, J., Chen, W., Du, H., et al. (2019). Integrated Glycome Strategy for Characterization of Aberrant LacNAc Contained N-Glycans Associated With Gastric Carcinoma. Front. Oncol. *9*, 636.
42. Yang, G., Cui, T., Wang, Y., Sun, S., Ma, T., Wang, T., Chen, Q., and Li, Z. (2013). Selective isolation and analysis of glycoprotein fractions and their glycomes from hepatocellular carcinoma sera. Proteomics *13*, 1481–1498.
43. Zhang, J., Zhong, Y., Zhang, P., Du, H., Shu, J., Liu, X., Zhang, H., Guo, Y., Jia, Z., Niu, L., et al. (2019). Identification of abnormal fucosylated-glycans recognized by LTL in saliva of HBV-induced chronic hepatitis, cirrhosis, and hepatocellular carcinoma. Glycobiology *29*, 242–259.
44. Kuno, A., Ikehara, Y., Tanaka, Y., Angata, T., Unno, S., Sogabe, M., Ozaki, H., Ito, K., Hirabayashi, J., Mizokami, M., and Narimatsu, H. (2011). Multilectin Assay for Detecting Fibrosis-Specific Glyco-Alteration by Means of Lectin Microarray. Clin. Chem. *57*, 48–56.
45. Qin, Y., Zhong, Y., Zhu, M., Dang, L., Yu, H., Chen, Z., Chen, W., Wang, X., Zhang, H., and Li, Z. (2013). Age-and sex-associated differences in the glycopatterns of human salivary glycoproteins and their roles against influenza A virus. J. Proteome Res. *12*, 2742–2754.
46. Smyth, G.K., and Speed, T. (2003). Normalization of cDNA microarray data. Methods *31*, 265–273.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Golden snub-nosed monkey	This study, see Table S2	N/A
Sprague-Dawley rats	This study, see Table S3	N/A
Chemicals, peptides, and recombinant proteins		
Albumin bovine serum (BSA)	Sigma-Aldrich	Cat#V900933
BCA Protein Assay Kit	Beyotime	Cat#P0011
Cy3 Fluorescent Dye	GE Healthcare	Cat#25801083
phosphate buffered saline (PBS)	Sigma-Aldrich	Cat#P3813
Tween 20	Sigma-Aldrich Corp	Cat#P1379
Glycine	Amresco	Cat#0167
NaB_4O_7	Sigma-Aldrich	Cat#71997
H_3BO_3	Sigma-Aldrich	Cat#B6768
NaOH	Sigma-Aldrich	Cat#71687
epoxy-coated magnetic particles	Scrbio	Cat#SRSH-EM-05-10
ethanol amine	Sigma-Aldrich	Cat#1.37044
Tris-HCl	Sigma-Aldrich	Cat#1.08219
NaCl	Sigma-Aldrich	Cat#793566
$CaCl_2$	Sigma-Aldrich	Cat#1.02379
$MgCl_2$	Sigma-Aldrich	Cat#208337
$MnCl_2$	Sigma-Aldrich	Cat#328146
Urea	Sigma-Aldrich	Cat#1706698
NH_4HCO_3	Sigma-Aldrich	Cat#09830
Dithiothreitol (DTT)	Sigma-Aldrich	Cat#DTT-RO
Iodoacetamide (IAM)	Sigma-Aldrich	Cat#8.04744
PNGase-F	NEB Biolabs	Cat#P0704S
Trifluoroacetic acid (TFA)	Sigma-Aldrich	Cat#88.08260
Acetonitrile (CAN)	Sigma-Aldrich	Cat#360457
Methanol	Sigma-Aldrich	Cat#179337
2,5-dihydroxybenzoic acid(DHB)	Sigma-Aldrich	Cat#149357
Jacalin, Unconjugated	Vector	Cat#L-1150-25
Erythrina Cristagalli Lectin (ECA)	Vector	Cat#L-1140-10
Hippeastrum Hybrid Lectin (HHL)	Vector	Cat#L-1385-2
Wisteria Floribunda Lectin (WFA)	Vector	Cat#L-1350-5
Griffonia Simplicifolia Lectin II (GSL-II)	Vector	Cat#L-1210-5
Maackia Amurensis Lectin II (MAL-II)	Vector	Cat#L-1260-2
Phaseolus Vulgaris Erythroagglutinin(PHA-E)	Vector	Cat#L-1120-5
Psophocarpus Tetragonolobus Lectin I (PTL-I)	Vector	Cat# L-1360
Sophora Japonica Agglutinin (SJA)	Vector	Cat#L-1135
Peanut Agglutinin (PNA)	Vector	Cat#L-1070
Euonymus Europaeus Lectin (EEL)	Vector	Cat#L-1331
Aleuria Aurantia Lectin (AAL)	Vector	Cat#L-1390-2
Lutos Tetragonolobus Lectin (LTL)	Vector	Cat#L-1320-5

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Maclura Pomifera Lectin (MPL)	Vector	Cat#L-1340-5
Lycopersicon Esculentum Lectin (LEL)	Vector	Cat#L-1170-2
Griffonia Simplicifolia Lectin I (GSL-I)	Vector	Cat#L-1100-5
Dolichos Biflorus Agglutinin (DBA)	Vector	Cat#L-1035-5
Lens Culinaris Agglutinin (LCA)	Vector	Cat#L-1040
Solanum Tuberosum (Potato) Lectin (STL)	Vector	Cat#L-1160-5
Psophocarpus Tetragonolobus Lectin II (PTL-II)	Sigma-Aldrich	Cat#L3014
Datura Stramonium Lectin (DSA)	Vector	Cat#L-1180-5
Vicia Villosa Lectin (VVA)	Vector	Cat#L-1230-5
Maackia Amurensis Lectin I (MAL-I)	Vector	Cat#L-1310-5
Galanthus Nivalis Lectin (GNA)	Vector	Cat#L-1240-5
Narcissus Pseudonarcis sus Lectin (NPA)	Vector	Cat#L-1370-5
Amaranthus caudatus lectin (ACA)	Vector	Cat#L-1250
Bauhinia Purpurea Lectin (BPL)	Vector	Cat#L-1280-5
Phaseolus Vulgaris Agglutinin (E + L) (PHA-E+L)	Vector	Cat#L-1120
Sambucus Nigra Lectin (SNA)	Vector	Cat#L-1300-5
Ricinus Communis Agglutinin I (RCA120)	Vector	Cat#L-1080-10
Bandeiraea simplicifolia Agglutinin (BS-I)	Vector	Cat#L-1100-5
Pisum Sativum Agglutinin (PSA)	Vector	Cat#L-1050-10
Soybean Agglutinin (SBA)	Sigma-Aldrich	Cat#L1395
Wheat Germ Agglutinin (WGA)	Sigma-Aldrich	Cat#L9640
Ulex Europaeus Agglutinin I (UEA-I)	Vector	Cat#L-1060
Phytolacca americana (PWM)	Sigma-Aldrich	Cat#L9379
Canavalia ensiformis (ConA)	Sigma-Aldrich	Cat#L7647
Software and algorithms		
GenePix Pro 6.0	Molecular Devices	RRID:SCR_010969
SPSS 20	IBM	RRID:SCR_016479
GraphPad Prism 8	GraphPad	RRID:SCR_002798
RStudio	Posit	RRID:SCR_00432
R version 4.2.1	R project	RRID:SCR_001905

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Songtao Guo (songtaoguo@nwu.edu.cn).

Materials availability

All materials used and generated in this study will be made available on request from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

All data reported in this paper will be shared by the "key resources table" and "STAR methods".

Any additional information required to reanalyze the data reporter in this paper is available from the lead contact upon request.

This paper does not report original code.

Raw data from Figures 3, 4, Table 1, 2, S2, S3, S4, S6, S7, S9, S10 and S11 were deposited on [Mendeley]: [https://doi.org/10.17632/2mpdxk8442.2].

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Urine sample collection

The studies were reviewed and approved by the Northwest University (NWU) Animal Welfare Committee. The resolution number is NWU-AWC-20171102R.

GSM urine samples were taken from Shaanxi Wild Animal Rescue Center and were collected from 14 captive adult individuals, including 6 pregnant females, 3 non-pregnant females and 5 males. Once a GSM urinates when the floor is clean, fresh urine would be collected with a sterile syringe and temporarily kept in 5 mL cryotubes at −20°C. The urine sample of three male individuals (HH, JJ, TT) were mixed them together to make sure that the minimum protein concentration is available, and this results in a sample size of 3 for males (see Table S10 [GSM urine sample information and protein concentration]). Urine samples of the SD rats were collected from animal lab center at The Fourth Military Medical University. The samples were divided into three groups: 15 pregnant females, 5 non-pregnant females, and 5 males (see Table S11 [SD rat urine sample information and its protein concentration]). Among these individuals, 4 SD rats across three groups were randomly selected as the testing set (numbered as t1-t4), and the other 21 individuals (13 pregnant females, 4 non-pregnant females, and 4 males) were used for the initial test and analysis. All the samples were collected for three consecutive days, then were stored at −80°C for following processing.

The samples from the same individuals of one day were mixed, and each individual in the same group were mixed before protein-extraction to avoid the variation due to individual differences. All the samples were collected under the same environment and feeding conditions and at the same time period of the day, avoiding the differences due to the variation of physiological reaction and other conditions.

METHOD DETAILS

Processing of urine samples

Samples were kept with 15 mL sterile centrifuge tubes and centrifuged at 5,000 rpm for 10 min at 4°C to remove insoluble substances. Protein was then enriched in urine with 3KDa ultrafiltration tube. The supernatant was then added to centrifuging at the speed of 8,000 g for 15 min at 4°C, and the liquid at the pipe bottom was discarded. This stage was repeated three times until 100 μL samples were gotten. The remaining liquid was centrifuged with an inverted filter for matching centrifuge tube at 5,000 rpm at 4°C for 5 min. The solution at the tube bottom was then collected and divided into a 1 mL centrifuge tube with the marks. Protein concentration is determined by Bicinchonimic Acid (BCA) Protein Assay kit. Concentration for each individual sample is listed in Table S10 [GSM urine sample information and protein concentration] and Table S11 [SD rat urine sample information and its protein concentration]. The samples did not reach the required concentration were discarded. Proteins were labeled with Cy3 fluorescent reagent, and free fluorescent dye was removed with Sephadex G-25 desalting column.

Lectin microarrays

Lectin microarray used is produced with 37 plant lectins that are combined with different N- and O-linked glycans. Each lectin has a specifically recognized range of glycan structures and monosaccharides are added to maintain this specificity (the details for the specificity of each lectin and other information are listed in Table S1 [The information of 37 lectins]).

Thirty-seven types of lectins dissolved in 1 mg/mL spot buffer containing 1 mmol/L monosaccharide, two units of negative control with 1 mg/mL Bovine serum albumin (BSA) and 1 unit of 1 mg/mL Cy3 dye labeled positive control were prepared. They were spotted on epoxysilane-coated slides into four 10 × 12 matrices with micro spotting pins by a smart microarrayer for 3 times, and then incubated at 50% humidity at 37°C for 2 h in the vacuum dryer. Later, they were blocked with buffer containing 1% (w/v) BSA with a pH 7.4 10 mM phosphate buffered saline (PBS) at 25°C for 1 h, then rinsed with 1 × PBST (10 mM PBS, 0.2% Tween 20, pH 7.4) twice, followed by another rinse in 10 mM PBS. Four μg of Cy3-labeled protein, and mixed with 0.5 mL incubation buffer (2% (w/v) BSA, 500 mM glycine, 0.1% Tween 20, 10 mM PBS) to incubate with the lectin microarrays at 25°C for 3 h in the hybridization oven at 4 rpm. They were then washed twice for 8 min with 1 × PBST, and twice for 5 min with 10 mM PBS. Subsequently, slides were dried by centrifugation at 600 rpm for 5 min at ambient temperature. They were then scanned with 70% of photomultiplier tube and 100% of laser power settings with a Genepix 4100A microarray scanner. Their images were analyzed at 532 nm for Cy3 detection by Genepix 3.0 software.[40,41]

Glycoprotein from urine by lectin-magnetic particle conjugates

The epoxy-coated magnetic particles (2 mg) were rinsed three times with coupling buffer (5 Mm NaB_4O_7, 180 mM H_3BO_3, 150 mM Na^+, pH 7.4) and reacted with 600 μL of 0.5 mg/mL lectin solution (lectins dissolved in coupling buffer) at room temperature (RT) for 3 h under gentle shaking according to the protocol. The unbound lectins were removed from the conjugates by washing six times with the coupling buffer.

The prepared lectin-magnetic particle conjugates were blocked by 600 μL blocking buffer (2% ethanol amine, 0.1% BSA, pH 9.0) at RT for 1 h under gentle shaking, and then rinsed 3 times with the binding buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM $MnCl_2$, pH 7.4). 2 mg proteins were dissolved in binding buffer, 600 μL total. Then added to lectin-magnetic particle conjugates at RT for 3 h under gentle shaking. The unbound proteins were removed from the conjugates by washing six times with the washing buffer (0.1% Tween 20, 100 mM Tris-HCl, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM $MnCl_2$, pH 7.2). Next, the selective isolated glycoproteins were eluted with 300 μL eluting buffer (8 M Urea, 100 mM NH_4HCO_3) for 1 h. Repeat the previous step once and mix the supermatant collected. The lectin-isolated glycoproteins were quantitated by BCA.[42,43]

Release the N-glycans by PNGase F

The lectin-isolated glycoproteins were concentrated by 10 kDa molecular weight cutoff (MWCO) ultrafiltration unit (Millipore), and rinsed three times with 300 μL 8 M urea solution. Add 200 μL 8 M urea at RT for 1 h under gentle shaking, centrifuge and remove the filtrate. Then added 200 μL DTT solution (10 mM) at RT for 1 h under gentle shaking, centrifuge and remove the filtrate. 20 μL IAM solution (20 mM) was added and reacted at RT for 1 h in the dark, centrifuge and remove the filtrate. The mixture is rinsed 3 times with 50 mM NH_4HCO_3 in the 10 k Da MWCO ultrafiltration unit, add 50 mM NH_4HCO_3 to total volume of 400 μL. 2 μL PNGase F (New England BioLabs) is added and incubated at 37°C overnight to release the N-glycans from the glycopeptides. Finally, centrifuge and collect the filtrate, added 200 μL ultrapure water centrifuge and collect the filtrate again, then mixed the filtrate and lyophilized it by CentriVap (Labconco, America).

Purification of N-glycans

Pre-clean the Hypercarb SPE column successively with 3 mL 1 M NaOH solution, 3 mL ultrapure water, 3 mL 30% trifluoroacetic acid solution, and 3 mL ultrapure water. Added 3 mL 50% ACN/0.1%TFA and 3 mL 5%ACN/0.1% TFA solutions to the column successively to condition column. Added 1 mL 0.1% TFA. N-glycan dissolved with 500 μL 0.1%TFA was added to the column, the filtrate was collected, and the sample was repeated for 3 times. Added 3 mL ultrapure water,3 mL 5% ACN/0.1%TFA solutions to the column successively to wash the column. Finally, 400 μL 50% ACN/0.1%TFA was added in the column, eluted twice, collected and lyophilized the filtrate.

Characterization of the N-glycans by MALDI-TOF/TOF-MS

Added 20 μL N-glycan solution (Methanol/aqueous solution, the volume is 1:1) to freeze-dried N-glycan. spotted 2 μL the dissolved N-glycan sample directly on an MTP AnchorChip sample target and vacuumized, repeat this step twice. 1 μL 20 mg/mL DHB matrix solution was added to the crystallized N-glycan sample and vacuumized. The target plate was placed on the mass spectrometer, ionization was operated by irradiation with a nitrogen laser (337 nm) at 1 kHz, the mass calibration was operated with peptide calibration standards (250 calibration points; Bruker). N-glycan chains identification was taken in positive and reflectron mode, we selected the intense ions from MS spectra to MS/MS with a higher energy (80%–90%) detection. The data were generated using FlexAnalysis software (Bruker Daltonics), analyzed and annotated using the GlycoWorkbench.[42,43]

DATA ANALYSIS

Signal intensity of the fluorescence for each spot and background is obtained with GenePix Pro 6.0. To avoid the influence of non-specific adsorption, we only studied the intensity of each spot, provided it is more than average level of the background. After having been screened, the median values of the valid signals were calculated. Background of the chip is the fluorescent value at the spotless area around the lectin point read by the chip scanner that is the background value of the lectin point. If fluorescence intensity value of the lectin point is less than the average background $\pm$standard deviation, it is considered to be an invalid value, and no subsequent calculations are performed. Raw fluorescent data were processed with max-normalization among 37 lectins to eliminate the biased fluorescence among different datasets.[44]

According to the principle of normalization, the sum of Normalized fluorescent intensities (NFIs) of 37 lectins is 1, and the average NFIs of each lectin is 0.027. If NFIs is more than twice of the average value – higher than 0.050, the abundance is considered to be abundant. On the other hand, if it is less than 1/2 of the average value – less than 0.013 (in our case, 0.01), the abundance is considered to be rare.[45] Microarray data were then tested with Mann-Whitney U test from SPSS 20. The processed parallel datasets were then analyzed based on fold changes according to the criteria of >2 or <0.5.[46] A value of $p < 0.05$ indicates a significantly increased or decreased for a certain kind of glycopatterns, respectively. GraphPad Prism 8 were used to explore the variation among the samples analyzed. In the verification analysis based on SD rats' data, R 4.2.1 was used to calculate the Euclidean distances (ED) between tested individuals and the group vale. Smaller ED refers more similarity between the individual and the group, which can then be used to classify the individual.