Diagnostic and Prognostic performance of blood plasma glycan features in Women Epidemiology Lung Cancer (WELCA) study

Yueming Hu and Chad R. Borges*

School of Molecular Sciences and The Biodesign Institute at Arizona State University, Tempe AZ 85287

Introduction

- Our lab's bottom-up "glycan node analysis" approach captures interesting glycan features such as "a2-6 sialylation", "B1-4 branching" and "outer-arm fucosylation" as single analytical signals.
- From the WELCA study, plasma samples of 208 female lung cancer patients in stage I-IV and 207 age-matched healthy women were obtained.

Blood plasma and serum glycomics represents a promising source of new generation cancer biomarkers. Glycan node analysis, is a molecularly bottom-up approach to P/S glycomics developed by Borges et al. in 2013, focusing on monosaccharides and linkage specific "glycan nodes" rather than the intact glycan structures .1,2 This approach captures all P/S glycans including N-, O-, and lipid-linked glycans and breaks them down into monosaccharides that maintains linkage information, by applying glycan linkage (methylation) analysis to whole biofluids (Fig. 1).

Furthermore, the important glycan features are captured and quantified as single analytical signals. In addition, many glycan nodes serve as direct surrogates for the activities of glycosyltransferases (GTs), enzymes that facilitate the construction of glycans and each of which is most commonly responsible for producing a unique glycan monosaccharide linkage pattern.

Recently, we have applied glycan node analysis to several cancer studies, including pancreatic cancer, ovarian cancer, prostate cancer, and lung cancer case-control study. For this study, 208 female patients with newly diagnosed stage I-IV lung cancer and 207 age-matched health women were obtained from Women Epidemiology Lung Cancer (WELCA) study. The purpose of this study was to further validate glycan node analysis as a means of detecting and predicting patient outcomes in lung cancer specifically in women. Interestingly, there are several important gender differences between men and women in lung cancer, including the facts that 1) after adjusting for the number of cigarettes smoked women have a three-fold greater risk of lung cancer than men; 2) never-smoker women are at significantly greater risk for lung cancer than men, and 3) women tend to have better survival rates than men. As such, we felt that for any differences observed in this study relative to our previously reported results in lung cancer, it would also be important to look for any existing gender-based differences in glycan nodes as they may occur in the context of lung cancer.

Concept/Method

- Does Not Require Pre-Isolation of Proteins or Glycans
- Covers N-, O-, and Lipid-Linked Glycans
- Peak area normalized to internal standards (heavy glucose & heavy GlcNAc)
- 10µL of plasma are hydrolyzed for 4 h with 1 M NaOH. Partially permethylated Glycan Analysis (PAPA) was performed by GC/MS.

Results

Diagnostic capacity of glycan features in lung cancer

- n values of each cohort:
  - Control: n = 207; Stage I: n = 16; Stage II: n = 13; Stage III: n = 45; Stage IV: n = 99.
  - The results of Kruskal-Wallis test and ROC curve agree with our observations in a prior study.

Dependence on smoking-status, age and histological type

- Glycan nodes/features were found to be independent of gender from two previous dual gender lung cancer sets. Therefore, a possible explanation for the better diagnostic performance of glycan features in early stages is the non-smoking matched control group involved in the WELCA set. Thus, dependence of glycan nodes/features on smoking-status, as well as age and histological type, was evaluated.

Figure 4. The minor dependence on smoking status and age of the top performing glycan node 3,4-linked GlcNAc in the WELCA study. (a) The univariate distributions of outer-arm fucosylation within the control group are shown, subdivided by smoking status. Different letters above data points indicate statistically significant differences between groups. ROC curves for stage I-IV lung cancer cases vs controls are provided in panels a-h. Areas under the ROC curves are provided in parenthesis next to the specified stages. "NS" next to the AUC values indicates that the ROC curve is not statistically significant.

Ability to predict survival in lung cancer

- The top quartiles of all four glycan node markers predicted all-cause mortality with hazard ratios range from 2.3 and p < 0.01, by Cox proportional hazards regression model with adjustment for age, smoking status, and cancer stage. These results agree with our observations in a prior study.

Conclusions/Summary

- Four glycan node-based features were able to separate lung cancer patients in nearly every stage from age-matched controls.
- Glycan features have minor or negligible dependence on smoking-status, age and histological type.
- The top quartiles of all four glycan node markers predicted all-cause mortality of lung cancer relative to all other quartiles combined.
- The above 3 conclusions validate findings observed in our previously published work.
- Marked early-stage detection was observed in WELCA set compared to other two lung cancer sets.
- Gender did not appear to account for the improved early-stage diagnostic performance of glycan nodes in this study compared to our previously published work.

References


---

Figure 1: Conceptual overview: An upregulated GT (e.g., GnT-V) causes an increase in the quantity of a specific, uniquely linked glycan monosaccharide residue (a 2,6-linked mannose "node" in this example)—which, through the subsequent action of other GTs, can lead to formation of a mixture of heterogeneous whole-glycan structures at low copy number each. Analytically pooling together the glycan "nodes" from amongst all the aberrant glycan structures in a given biomatrix provides a more direct surrogate measurement of GnT-V activity than any single intact glycan. Actual extracted ion chromatograms from 10µL of plasma samples shown.