

Expanding the Molecular Glycophenotype Ontology to include model organisms and acquired diseases

Jean-Philippe Gourdine¹, Nicole Vasilevsky¹, Lilly Winfree¹, Matthew Brush¹,

¹Oregon Clinical and Translational Research Institutes
Oregon Health and Science University
Translational and Integrative Sciences Lab (TISLAB)
Portland, Oregon
gourdine@ohsu.edu

Melissa Haendel*¹

*Oregon State University
TISLAB, Linus Pauling Institute
Corvallis, Oregon USA

Abstract— Glycans are an underappreciated class of molecules despite the fact that they are implicated in more than 100 known diseases. We have developed an ontology model that captures glycan abnormalities at the molecular level (glycophenotypes) called the molecular glycophenotype ontology (MGPO). Only 30% of known glycosyltransferases have been implicated in human genetic disorders of glycosylation. Ortholog glycosyltransferases from model organism can cover relevant biological information on potential human diseases. Thus, extending MGPO to represent additional phenotypes and support annotation of model organism data will help cross-species comparison. Expansion of MGPO will also include annotation of glycophenotypes from acquired diseases.

Keywords—glycans- glycophenotypes- diseases- model organisms.

I. INTRODUCTION

Descriptions of phenotypic abnormalities in ontologies have been useful for computational comparisons of genotype-phenotype associations across species (for example, via the Human Phenotype Ontology, HPO [1,2] and Mammalian Phenotype Ontology, MP[3]). Such ontologies have supported the creation of new tools that help gene variant prioritization based on phenotypes [2] (e.g. Exomiser [4]). We have developed a prototype of a molecular glycophenotype ontology (MGPO) that captures abnormality of glycosylation at a molecular level (glycophenotypes) that may occur in human genetic diseases [5]. Molecular phenotyping can enhance precision medicine, for instance, fucosidosis ranks on Monarch Phenotypic Similarity tools [6] were increased with glycophenotypes terms [7]. We are currently using MGPO to annotate glycophenotypes to hundreds of genetic glycanopathies (e.g. diseases related to N-glycosylation pathways). Many glycanopathies and glycosylation pathways have been discovered using biological models (e.g. yeast [8]). Hence, model organisms can inform human glycanopathies. To improve cross-species comparison of genetic diseases, enhancements to the current MGPO model are needed. In

addition, many acquired diseases can show glycophenotypes. For instance, hyposialylated transferrin is a glycophenotype for chronic alcohol exposure [9]. Annotating glycophenotypes from acquired diseases in MGPO in addition to anatomical phenotypes could help enhance the disease description.

II. GLYCOBIOLOGY OF MODEL ORGANISMS CAN BE INFORMATIVE FOR HUMAN DISEASES

Only 128 genes are known to be involved in human glycosylation disorders (Congenital Disorder of Glycosylation, CDG) [10]. While 21% of them are related to non-glycan processes such as transport of small molecules and ions in the Golgi/ER, the remaining 79% is related to glycan synthesis or degradation. These account for only about 30% of known human glycosyltransferases reported in the database for Carbohydrate-Active Enzymes Database (CAZY, <http://www.cazy.org/>) [11]. Mutations in ortholog of glycosyltransferases in model organisms can be useful for disease comparison, because in patients with undiagnosed diseases, only the known clinical variants are reported from genome/exome analyses, and thus, other relevant mutations can be missed that may have been identified in model organisms. Exomiser [4] can help variant prioritization using cross-species genotypes-phenotypes association. In order to expand disease comparison between species, curation and representation of glycophenotype related to genes can help. For instance, Baker's yeast (*Saccharomyces cerevisiae*) has been used as a model to understand N-glycan synthesis in eukaryotes and led to understanding of many CDGs [8]. In yeast and fly models, a defect in O-mannosylation was discovered and led to understanding of many dystroglycanopathies (e.g. POMT1) [8]. Mouse knockout models can be useful as many knock-out genes are associated with glycophenotypes and are available in public databases like the Consortium for Functional Glycomics (CFG, <http://www.functionalglycomics.org/>) [12]. However, these data from CFG are not annotated with MGPO terms nor represented in HPO yet. Finally, knock-out of glycan related genes in zebrafish [13] have showed many anatomical

abnormalities that could be informative for comparison of many human glycanopathies.

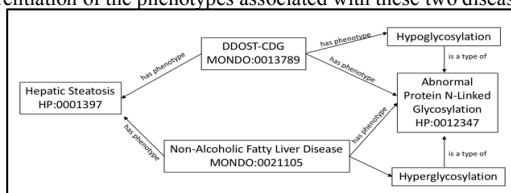
III. MANY ACQUIRED DISEASES HAVE GLYCOPHENOTYPES BUT ARE NOT REPRESENTED YET IN HPO

Glycophenotypes can provide more precise disease descriptions for acquired diseases and genetic diseases. For instance, the phenotype ‘hepatic steatosis’ (HP:0001397) is common in 188 diseases in the cross-species phenotype-genotype data integration platform, Monarch Initiative [6] (www.monarchinitiative.org) including non-liver fatty disease (NLFD, MONDO:0021105) and DDOST-CDG (MONDO:0013789). NLFD and DDOST-CDG share the phenotype ‘Abnormal Protein N-Linked Glycosylation’ (HP:0012347). The main difference is the type of abnormal glycoprotein, hypoglycosylation for DDOST-CDG and hyperglycosylation for NLFD but these hyper/hypoglycosylation are not yet integrated in HPO (Fig. 1). Thus capturing glycophenotypes for acquired disease can help differentiate these diseases. Many cancers also show specific glycosylation changes that can be measured in glycan binding assays [14]. In the HPO, 5 types of cancers are referenced: Cervix cancer (HP:00300790), Colon cancer (HP:0003003), Merkel cell skin cancer (HP:0030447), Prostate cancer (HP:0012125), Stomach cancer (HP:0012126). All of them have glycophenotypes that are not represented yet in MGPO : GalNAc α -1-3Gal (cervix cancer) [15], reduced expression of core 3 and 4 glycans and increased Tn/sTn antigen (Colon cancer)[16], increased sialylation and fucosylation (prostate cancer) [17], decreased levels of high-mannose-type glycans (Stomach cancer)[18]. On the other side of the spectrum of glycan and cancer resistance, some rare animal models like naked mole rats can help find molecular strategy to fight cancer (as they present with high hyaluronan mass and cancer resistance) [19].

IV. CONCLUSION

We believe that an integration of glycomics data in HPO and MP will allow for a better understanding the interconnectivity of molecules, and thus, better diseases comparison and diagnoses. Indeed, to this day, 25% of patients have been diagnosed with traditional methods like exome analysis (Undiagnosed Diseases Network, personal communication); adding molecular phenotyping could help those who remain undiagnosed. Finally, molecular phenotyping with MGPO could provide better insights toward possible treatments, e.g. dietary supplementation of glycans [20].

Figure 1: Example of a genetic and acquired disease that share two phenotypes. Increasing the granularity of glycophenotype can lead to a better differentiation of the phenotypes associated with these two diseases.



- [1] Robinson PN, Köhler S, Bauer S, Seelow D, Horn D, Mundlos S. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am J Hum Genet.* 2008;83: 610–615.
- [2] Köhler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Aymé S, et al. The Human Phenotype Ontology in 2017. *Nucleic Acids Res.* 2017;45: D865–D876.
- [3] Smith CL, Eppig JT. The Mammalian Phenotype Ontology as a unifying standard for experimental and high-throughput phenotyping data. *Mamm Genome.* 2012;23: 653–668.
- [4] Smedley D, Jacobsen JOB, Jäger M, Köhler S, Holtgrewe M, Schubach M, et al. Next-generation diagnostics and disease-gene discovery with the Exomiser. *Nat Protoc.* 2015;10: 2004–2015.
- [5] Gourdiine J-P, Metz T, Koeller D, Brush MH, Haendel M. Building a Molecular Glyco-phenotype Ontology to Decipher Undiagnosed Diseases. *ICBO/BioCreative.* 2016. Available: http://ceur-ws.org/Vol-1747/IP06_ICBO2016.pdf
- [6] McMurry JA, Köhler S, Washington NL, Balhoff JP, Borromeo C, Brush M, et al. Navigating the Phenotype Frontier: The Monarch Initiative. *Genetics.* 2016;203: 1491–1495.
- [7] Gourdiine JP, Brush M.H, Vasilevsky N.A, Shefchek K., McMurry J, Haendel M.A. Toward deep molecular phenotyping with glycomics data to decipher rare and undiagnosed diseases. Unpublished. 2018;
- [8] Lehle L, Strahl S, Tanner W. Protein glycosylation, conserved from yeast to man: a model organism helps elucidate congenital human diseases. *Angew Chem Int Ed Engl.* 2006;45: 6802–6818.
- [9] Flahaut C, Michalski JC, Danel T, Humbert MH, Klein A. The effects of ethanol on the glycosylation of human transferrin. *Glycobiology.* 2003;13: 191–198.
- [10] Freeze HH, Schachter H, Kinoshita T. *Known Human Glycosylation Disorders.* Cold Spring Harbor Laboratory Press; 2017.
- [11] Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2014;42: D490–5.
- [12] Comelli EM, Head SR, Gilmartin T, Whisenant T, Haslam SM, North SJ, et al. A focused microarray approach to functional glycomics: transcriptional regulation of the glycome. *Glycobiology.* 2006;16: 117–131.
- [13] Ruzicka L, Bradford YM, Frazer K, Howe DG, Paddock H, Ramachandran S, et al. ZFIN, The zebrafish model organism database: Updates and new directions. *Genesis.* 2015;53: 498–509.
- [14] Hashim OH, Jayapalan JJ, Lee C-S. Lectins: an effective tool for screening of potential cancer biomarkers. *PeerJ.* 2017;5: e3784.
- [15] Li Q, Anver MR, Li Z, Butcher DO, Gildersleeve JC. GalNAc α 1-3Gal, a new prognostic marker for cervical cancer. *Int J Cancer.* 2010;126: 459–468.
- [16] Holst S, Wuhler M, Rombouts Y. Glycosylation characteristics of colorectal cancer. *Adv Cancer Res.* 2015;126: 203–256.
- [17] Munkley J, Mills IG, Elliott DJ. The role of glycans in the development and progression of prostate cancer. *Nat Rev Urol.* 2016;13: 324–333.
- [18] Ozcan S, Barkauskas DA, Renee Ruhaak L, Torres J, Cooke CL, An HJ, et al. Serum glycan signatures of gastric cancer. *Cancer Prev Res.* 2014;7: 226–235.
- [19] Tian X, Azpurua J, Hine C, Vaidya A, Myakishev-Rempel M, Ablueva J, et al. High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. *Nature.* Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2013;499: 346.
- [20] Dalziel M, Crispin M, Scanlan CN, Zitzmann N, Dwek RA. Emerging principles for the therapeutic exploitation of glycosylation. *Science.* 2014;343: 1235681.