Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Chemical Biology

Glycomics meets genomics, epigenomics and other high throughput omics for system biology studies

Vlatka Zoldoš[1], Tomislav Horvat[1] and Gordan Lauc[2,3]

Majority of eukaryotic proteins are glycosylated and their glycan moieties have numerous important structural, functional and regulatory roles. Because of structural complexity of glycans and technological limitations glycomics, and particularly glycoproteomics was not able to follow rapid progress in genomics and proteomics over last 30 years. However, the field of glycan has been progressing rapidly and first large-scale studies of the glycome have been completed recently. These studies have revealed significant differences in glycome composition between individuals, which may contribute to the human phenotypic variability. The current state-of-the-art in high-throughput glycomics and its integration with genomics, epigenomics and lipidomics is reviewed in this article.

Addresses
[1] University of Zagreb, Faculty of Science, Zagreb, Croatia
[2] University of Zagreb, Faculty of Pharmacy and Biochemistry, Zagreb, Croatia
[3] Genos Ltd, Glycobiology Laboratory, Zagreb, Croatia

Corresponding author: Lauc, Gordan (glauc@pharma.hr)

Current Opinion in Chemical Biology 2012, **17**:xx–yy

This review comes from a themed issue on **Omics**

Edited by **Pauline Rudd** and **Matthew Bogyo**

1367-5931/$ – see front matter, © 2012 Elsevier Ltd. All rights reserved.

http://dx.doi.org/10.1016/j.cbpa.2012.12.007

Introduction

Nearly all membrane and secreted proteins as well as numerous intracellular proteins are modified by covalent addition of complex oligosaccharides (glycans), which play roles in almost every biological process and are involved in every major disease [1••,2•]. A peculiarity of glycan moieties of glycoproteins is that they are not synthesized using a direct genetic template. Instead, they result from the activity of several hundreds of enzymes organized in complex pathways. Changes in the activity and/or localization of any of the enzymes involved in glycan biosynthesis will affect the final structure of a glycan. Therefore, in addition to being defined by an individual's genetic background, the glycome (defined as a set of all glycans in an analyzed tissue or organism) is shaped by environmentally induced changes in gene expression (Figure 1).

Both inherited (genetic) and acquired (environmental) factors that modulate glycosylation affect numerous molecular processes, including interactions with specific receptors or half-life of numerous membrane proteins [3•]. Both quantitative and qualitative changes in the repertoire of glycan structures have been found in many complex diseases and cancer [4]. However, due to the structural complexity of glycans and technological limitations the knowledge of functional importance of glycans is significantly lagging behind the knowledge of DNA and proteins.

The development of high-throughput quantitative glycomic analysis

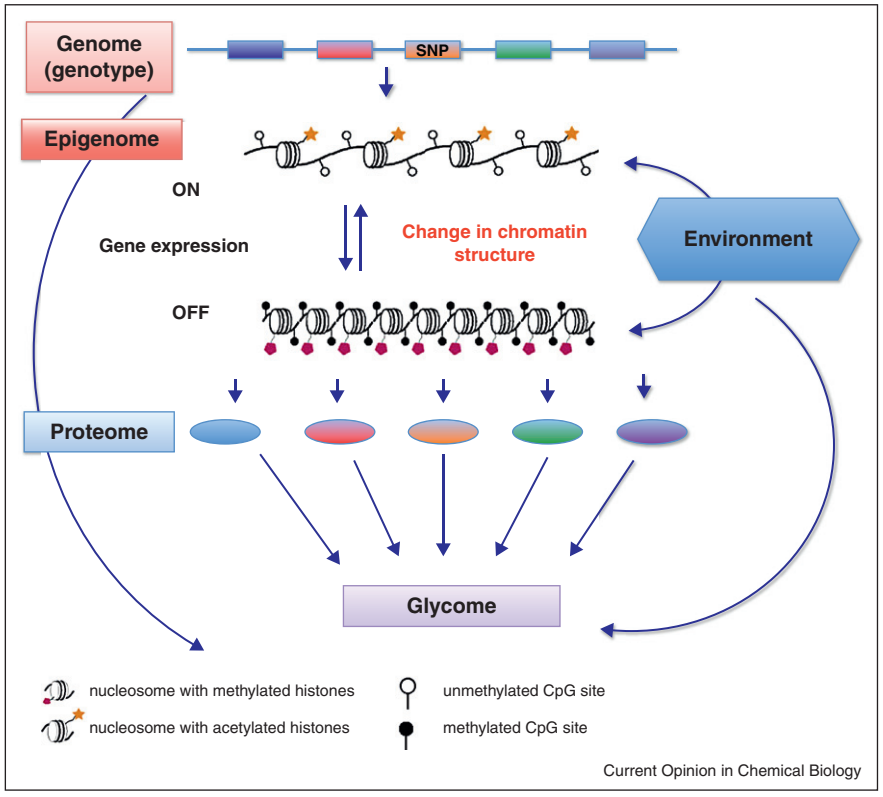
Until only a few years ago glycan analysis was extremely laborious and complex, hampering large-scale studies of the glycome. However, significant progress has been made in the past few years and several methods for high-throughput glycan analysis now exist [5–9,10•]. In principle, three main analytical methods are being used for high-throughput analysis of glycans: liquid chromatography (high performance — HPLC and ultra performance — UPLC, Figure 2), mass spectrometry (MS) and capillary electrophoresis (CE). An overview of some basic characteristics of these three methods is presented in Table 1, while more detailed description can be found in a recent review by Rakus and Mahal [11].

In general, UPLC and CE enable reliable glycan quantification. However, the development of analytical procedures for analysis of a particular biological system is demanding since glycan structures present in each chromatographic/electrophoretic peak need to be individually determined by MS or egzoglycosidase digestion. In addition, separation of glycans is rarely complete, thus each peak is a mixture of variable relative signal contributions originating from different glycans. MS analysis provides more structural details, but is generally less quantitative. An additional problem lies in the fact that signal response factors are different for different glycans, which complicates calculation of the derived traits, like galactosylation, fucosylation, sialylation, *etc.*

It is estimated that many thousands of different glycans are attached to human proteins [11,12•] and none of the currently available analytical methods can simultaneously quantify all of them. Therefore, each analysis provides only a limited insight into the complexity of the glycome. On top of this, majority of glycoproteins have multiple glycosylation sites, which can be occupied by different

Please cite this article in press as: Zoldoš V, et al.: Glycomics meets genomics, epigenomics and other high throughput omics for system biology studies, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.12.007

Figure 1



Complex genetics and epigenetics of protein glycosylation. Glycans are synthetized by dynamic interaction between many enzymes and other proteins. In addition to sequence variation in DNA, glycosylation is strongly influenced by environmental effects. Environmental factors (like nutrition, stress, smoking, *etc.*) can act on the level of enzymatic reactions, but they can also modulate gene expression by epigenetic mechanisms like affecting DNA and histone modifications (such as histone acetylation and methylation). Cellular memory stored in epigenetic modifications can affect glycan synthesis long after the disappearance of the environmental factor itself. Rectangles represent individual genes; ovals in different colors represent different enzymes involved in biosynthetic pathway of glycan structures.

glycan species with highly specialized functions. We are very far from understanding the role of glycans on majority of glycoproteins, but some prominent examples, like IgG (which has been studied in more detail), clearly demonstrate the importance of glycans for structural stability and biological function of glycoproteins [3•].

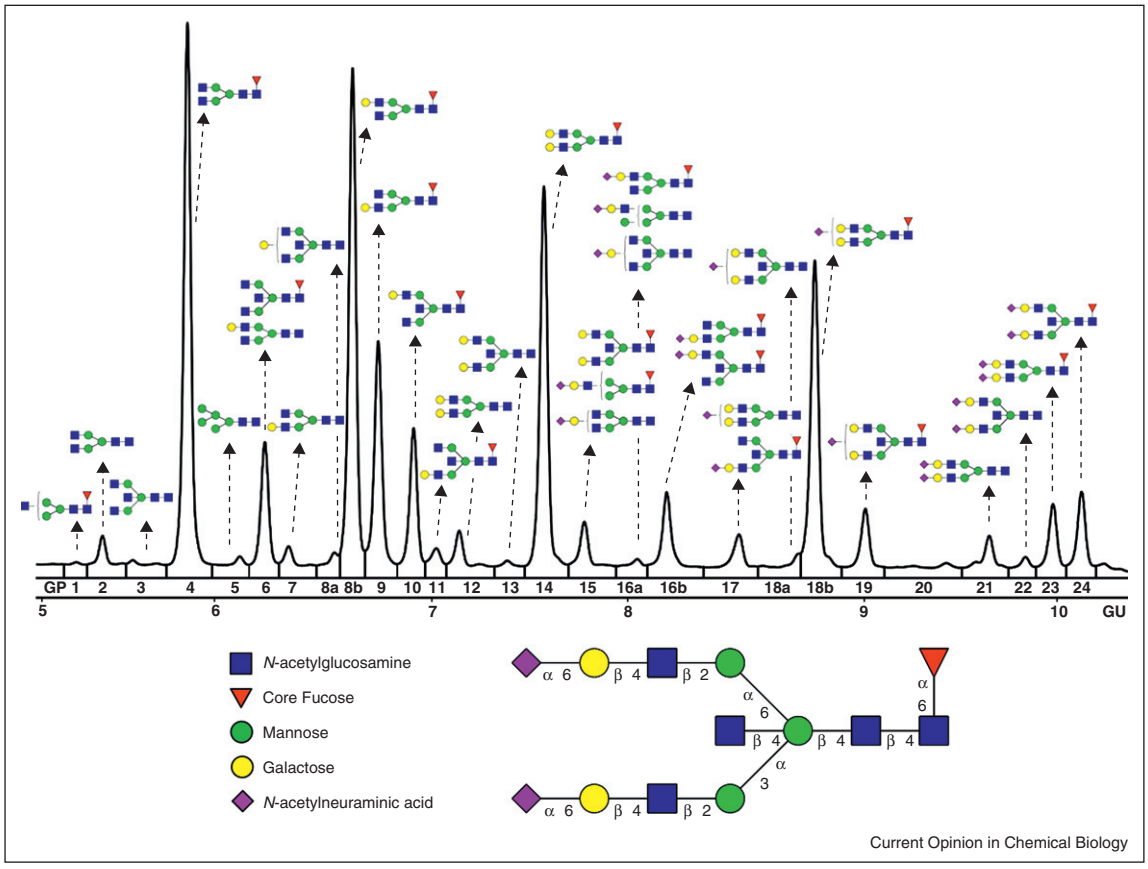
Several large population studies of the plasma glycome have been published until now [13–15]. They have all revealed a high variability in glycome composition between individuals. However, besides age, which significantly affected galactosylation, all other environmental factors individually accounted only for a small fraction of the observed variance, thus the main source of glycome variation between individuals is still not known. Interestingly, when compared to the total plasma glycome, the glycome composition of a single isolated protein (IgG, which is the only glycoprotein that has been analyzed at the population level) varied even more between individuals [10•]. This indicates that varying concentrations of plasma proteins and different structural and functional roles of the same glycans on different plasma proteins actually blur protein-specific regulation of glycosylation of individual proteins.

High-throughput DNA analysis enabled one of the largest bursts of discovery in human history

Until middle of last decade two approaches were used to study the effects of genetic variation on complex human diseases: genetic linkage studies and the candidate-gene approach. Studies of genetic linkage used sets of scarce genetic markers and were consequently successful only in identifying rare genetic mutations with high penetrance and large-scale effects (monogenic, Mendelian disorders). These studies were generally underpowered to detect effects of common genetic variants [16]. On the other hand, in majority of candidate gene studies, which were based on *a priori* hypotheses of biological plausibility, used sample sizes of only up to several hundred cases and controls and this led to results that were generally not replicated in future studies [17,18••].

Please cite this article in press as: Zoldoš V, et al.: Glycomics meets genomics, epigenomics and other high throughput omics for system biology studies, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.12.007

Figure 2



Ultra performance liquid chromatography (UPLC) separation of the IgG glycome. Different chromatographic techniques can be used to separate glycome into individual glycans or groups of similar glycans. UPLC separation is frequently used to separate glycans attached to human IgG in a way that enables reliable quantification of nearly all individual components of the IgG glycome. However, UPLC analysis does not reveal structural details, thus individual chromatographic peaks had to be collected and subjected to structural analysis by MS as described in Pucic *et al.* [10•].

The development of high-throughput genotyping technologies enabled radical changes in experimental design [19]. Several hundred thousands of particularly useful single nucleotide polymorphism (SNP) markers spanning across the entire genome (haplotype tagging) were chosen to develop high density SNP arrays, which made hypothesis-free genome-wide association studies (GWAS) possible [20,21•]. Over 500,000 individuals were genotyped in this way enabling identification of loci associated with complex diseases such as breast and colorectal cancers [22,23], inflammatory bowel disease [24,25], type-2 diabetes [26], and a variety of other diseases and traits [27]. However, individual effects of the identified loci are generally small, explaining only a small fraction of the trait or disease variation [21•]. As such, they do not substantially improve predictions over those based on known factors such as family history [28–30]. The discrepancy between population genetic models and family-based estimates of heritability can be overcome by inclusion of additional measurements such as epigenetic measurement which could substantially contribute to understanding complex diseases if combined with genetic associations and phenotype information [31,32,33•].

Epigenetic mechanisms account both for cellular state within an individual (cellular memory) and for non-DNA sequence-based trans-generational inheritance, which can account for discrepancy between genetic (population) and familial inheritance. First genome-wide chromatin assays have already enabled quantification of interindividual differences in chromatin state, which provided an initial insight into epigenetic contribution to human variation of complex phenotypes [34,35]. The new field of epigenetic epidemiology, which aims at examining time-dependence, heritability, and environmental relationship of epigenetic marks and at integrating novel genome-wide methylation scans with GWAS, is emerging as an interface between genetics, epigenetics and human disease variation [36,37]. Although still very challenging due to the nature of the epigenetic information, epigenome-wide association studies (EWAS) are already

Please cite this article in press as: Zoldoš V, et al.: Glycomics meets genomics, epigenomics and other high throughput omics for system biology studies, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.12.007

Table 1

Comparison of different technologies for glycan analysis

	UPLC	CE-LIF	MS
Acceptance/usage	Very widely used	Not used much, but gaining importance	Widely used
Throughput	High, approximately 50 samples per UPLC per day without multiplexing	Very high, multiplexing with up to 96 capillaries allows the analysis of thousands of samples per day.	Very high, as single mass spectra can be obtained at a subminute time scale. Analysis of thousands of samples per day is possible
Expertise required for instrument operation	Medium	Medium	High: both MALDI-MS and ESI-MS require quite some experience
Resolution	High	High	Very high
Capacity for isomer separation	Very good	Very good	None, but distinguishing of isomers may be achieved by tandem MS
Capacity for reliable quantification	Very good	Medium	Medium/low
Costs of equipment	High	Medium (due to cheap multiplexing)	Very high
Costs per sample in high throughput mode	Rather high costs, mainly due to low throughput (multiple UPLC systems required) and costs of solvents	Low costs per sample	Rather high costs, mainly due to expensive equipment

Three main technologies are today being used to analyze glycome in a high-throughput manner. Some main features of ultra-performance liquid chromatography (UPLC), capillary electrophoresis with laser induced fluorescence (CE-LIF) and mass spectrometry (MS) are presented.

underway and, integrated with GWAS, are expected to reveal epigenetic factors which contribute to disease phenotype [38].

Genome wide association studies of the human glycome

Only three GWA studies of glycosylation-related traits have been published so far [39,40•,41]. The first one determined composition of desialylated total plasma *N*-glycome by HPLC analysis in 2705 individuals from the islands of Vis and Korčula in Croatia and Orkney Islands in Scotland. The percentage of glycan structures containing core or antennary fucose was calculated as well. Out of 13 identified HPLC peaks, significant associations with particular SNPs were found for five of them, as well as for antennary fucose [40•]. The identified SNPs were located within known glycosyltransferase genes, *FUT6* and *FUT8*, and were associated with glycan structures along biosynthetic pathways of these two glycosyltransferases; thus, the molecular mechanisms behind these associations were clear. However, this approach led to a discovery of yet another gene, previously unrelated to protein glycosylation, coding for the transcription factor HNF1A. Subsequent functional studies revealed the central role of HNF1A in antennary fucosylation of plasma proteins exercised by promoting both *de novo* and salvage pathways of GDP-fucose synthesis, expression of *FUT3*, *FUT5* and *FUT6*, and the suppression of expression of *FUT8* [40•].

As an extension to the first GWA study, the second one included more individuals (3533) and a more detailed plasma glycome analysis [39], confirming all of the previous observations and revealing a new association of *HNF1A* with glycan branching. In addition, three new genes which associate with plasma glycome composition have been identified: *MGAT5*, *B3GAT1* and *SLC9A9*. In accordance with its biological function, *MGAT5* has been associated with highly branched glycans. *B3GAT1* is a member of glucuronyltransferase gene family. Glucuronic acid is known to exist on a subset of human lymphocytes, but was never reported on plasma proteins. A detailed structural analysis confirmed the existence of glucuronic acid on a subset of N-glycans released from human plasma glycoproteins [10•], indicating that results from GWAS can lead structural glycomic studies. The third identified gene, *SLC9A9*, was also previously not related to glycosylation. It codes for a proton pump which regulates pH in endosomes [42]. Since changes in Golgi pH can impair protein sialylation [43], the association between *SLC9A9* and tetrasialylated glycans makes biological sense.

Epigenetics of protein glycosylation

Longitudinal studies revealed high stability of an individual plasma N-glycome over a period of weeks and even after one year [44]. Heritability of individual glycan levels was generally below 50% [15], indicating presence of nonheritable mechanisms that provide temporal stability in the control of hundreds of different enzymes participating in glycan biosynthesis. Regulatory elements, such as promoters, enhancers and insulators, are major targets for various epigenetic modifications [45], which modulate function of these elements in the regulation of gene

Please cite this article in press as: Zoldoš V, et al.: Glycomics meets genomics, epigenomics and other high throughput omics for system biology studies, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.12.007

expression. Recent development of new-generation sequencing technologies enabled whole-genome expression and epigenetic profiling analyses, which provided insight into associations between chromatin signatures and gene expression [46].

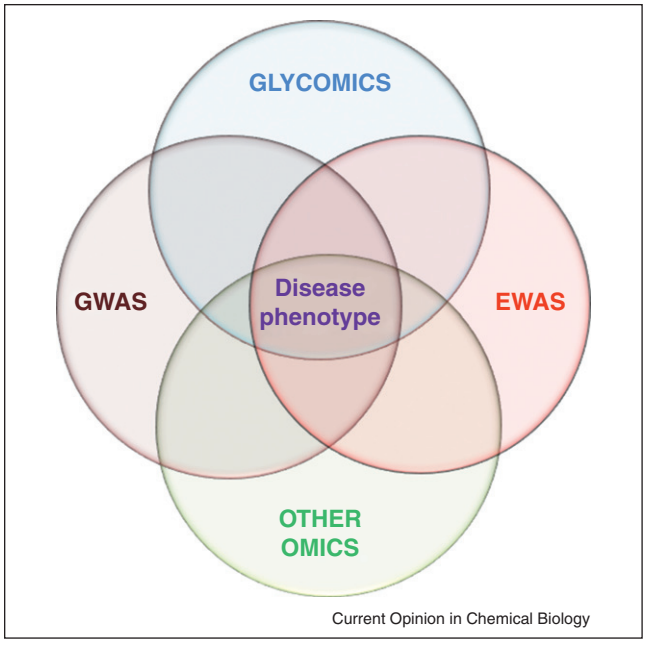
Epigenetic deregulation of many individual glyco-genes was reported so far only in cancer and majority of studies analyzed only methylation status of the gene promoter regions [47–49] or rarely combined analysis of DNA methylation with the analysis of histone modifications [50]. However, comprehensive analyses linking epigenetic modifications with the expression of glyco-genes are still lacking. So far, particular glyco-genes were mostly included only as a part of studies that aimed at elucidating global epigenetic changes during tumor development, such as in the case of DNA methylation profiling of chronic lymphocyte leukemia or breast cancer [51,52]. Therefore, future efforts need to be directed towards mining and analyzing publicly available data in order to generate novel findings for the subset of glyco-genes. An alternative approach is the use of epigenetic inhibitors, which modulate gene expression levels, followed by identification of genes, which affect composition of the N-glycome [53,54]. Targeted genome editing followed by functional studies will open many new opportunities in the future, but this technology is still not fully adapted to epigenetic studies [55,56•].

The importance of epigenetic regulation of glyco-genes is still underestimated and represents an intriguing challenge for further studies. Recently, we started to unravel potential mechanisms by which epigenetic deregulation of the glycome may contribute to disease development. By applying GWAS to human plasma N-glycome, we identified the transcription factor HNF1A as the master regulator of plasma protein fucosylation [40•]. After analyzing methylation levels on four CpG sites, which affect HNF1A expression in over 800 individuals, we were able to show statistically significant associations between epigenetic status of the promoter of this gene and highly branched glycan structures in the plasma glycome [57]. We obtained further support for this association by showing the same type of glycosylation changes in plasma of patients with Maturity Onset Diabetes of the Young (HNF1A-MODY), a subset of diabetes caused by inactivating mutation in HNF1A.

Conclusions

The complex interplay of multiple molecular systems is an inherent property of all biological processes. However, the traditional reductive approach is unable to reveal the complexity or determine the true role and importance of different individual components. This is clearly evident from the poor performance of genetic variants compared to the traditional epidemiological risk factors in predicting the development of even the most heritable illnesses [28].

Figure 3



Multi-dimensional omics. The phenotype of complex diseases can be understood only by considering new research horizons that integrate genomics, epigenomics, glycomics and other omics, which all contribute to the development and course of disease.

The development of a complex human disease is a combination of genetic background and numerous environmental factors that act during lifetime. While the link between genes and protein function is relatively straightforward, this is not the case for glycans or lipids, which are produced in complex biological pathways affected by both genes and the environment. Some of the environmental factors are perpetuated through epigenetic mechanisms, which modulate gene expression and help an individual to adapt to changing environment (Figure 1). This results in a very complex phenotype, which is hard to predict from individually measured genetic or environmental components, and a new research horizon that integrates genomics, epigenomics and high-dimensional phenotype omics is emerging [31,58,59]. Glycomics is here particularly interesting because it already integrates genetic background with environmental factors acting during lifetime of an organism [60]. However, data integrating glycomics with other omics are very limited. One recent study integrating glycomics and lipidomics resulted in identification of a very interesting association between glycan branching and polyunsaturated lipids, which might be part of the mechanism involved in the regulation of membrane

Please cite this article in press as: Zoldoš V, et al.: Glycomics meets genomics, epigenomics and other high throughput omics for system biology studies, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.12.007

dynamics [61]. Multilevel omics datasets offer an opportunity to study how individual parts of a biological system work together in transition from health to disease (Figure 3), but further progress is needed to develop both analytical and statistical methods which would enable the analysis of different types of omics and their interactions for high-dimensional predictive modeling.

Acknowledgements

Work in authors' laboratories is supported by the Croatian Ministry of Science, Education and Sport grants #006-0061194-2023 (to GL), #119-1191196-1224 (to VZ); by the European Commission GlycoBioM (contract #259869), HighGlycan (contract #278535), IBD-BIOM (contract #305479) and MIMOmics (contract #305280) grants and by AUF PSCI grant.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. •• Walt D, Aoki-Kinoshita KF, Bendiak B, Bertozzi CR, Boons GJ, Darvill A, Hart G, Kiessling LL, Lowe J, Moon R *et al.*: *Transforming Glycoscience: A Roadmap for the Future*. Washington: National Academies Press; 2012.
Very important policy document endorsed by the US National Academy of Sciences; promotes investment into glycoscience as a priority for the next decade.

2. • Moremen KW, Tiemeyer M, Nairn AV: **Vertebrate protein glycosylation: diversity, synthesis and function**. *Nat Rev Mol Cell Biol* 2012, **13**:448-462.
Comprehensive review about protein glycosylation.

3. • Gornik O, Pavic T, Lauc G: **Alternative glycosylation modulates function of IgG and other proteins — implications on evolution and disease**. *Biochim Biophys Acta* 2012, **1820**:1318-1326.
Overview of functional significance of alternative glycosylation of IgG and other proteins.

4. National Research Council: *Transforming Glycoscience: A Roadmap for the Future*. National Academies Press; 2012.

5. Royle L, Campbell MP, Radcliffe CM, White DM, Harvey DJ, Abrahams JL, Kim YG, Henry GW, Shadick NA, Weinblatt ME *et al.*: **HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software**. *Anal Biochem* 2008, **376**:1-12.

6. Schwarzer J, Rapp E, Reichl U: **N-glycan analysis by CGE-LIF: profiling influenza A virus hemagglutinin N-glycosylation during vaccine production**. *Electrophoresis* 2008, **29**:4203-4214.

7. Ruhaak LR, Hennig R, Huhn C, Borowiak M, Dolhain RJ, Deelder AM, Rapp E, Wuhrer M: **Optimized workflow for preparation of APTS-labeled N-glycans allowing high-throughput analysis of human plasma glycomes using 48-channel multiplexed CGE-LIF**. *J Proteome Res* 2010, **9**:6655-6664.

8. Selman MH, McDonnell LA, Palmblad M, Ruhaak LR, Deelder AM, Wuhrer M: **Immunoglobulin G glycopeptide profiling by matrix-assisted laser desorption ionization Fourier transform ion cyclotron resonance mass spectrometry**. *Anal Chem* 2010, **82**:1073-1081.

9. Knezevic A, Bones J, Kracun SK, Gornik O, Rudd PM, Lauc G: **High throughput plasma N-glycome profiling using multiplexed labelling and UPLC with fluorescence detection**. *Analyst* 2011, **136**:4670-4673.

10. • Pucic M, Knezevic A, Vidic J, Adamczyk B, Novokmet M, Polasek O, Gornik O, Supraha-Goreta S, Wormald MR, Redzic I *et al.*: **High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations**. *Mol Cell Proteomics* 2011, **10**:M111010090.
First large population study of the human IgG glycome. Presents glycome variation in nearly 2300 people from three isolated populations.

11. Rakus JF, Mahal LK: **New technologies for glycomic analysis: toward a systematic understanding of the glycome**. *Annu Rev Anal Chem (Palo Alto Calif)* 2011, **4**:367-392.

12. • Cummings RD: **The repertoire of glycan determinants in the human glycome**. *Mol Biosyst* 2009, **5**:1087-1104.
An interesting article that presents variety of glycan epitopes present in the human glycome.

13. Ruhaak LR, Uh HW, Beekman M, Koeleman CA, Hokke CH, Westendorp RG, Wuhrer M, Houwing-Duistermaat JJ, Slagboom PE, Deelder AM: **Decreased levels of bisecting GlcNAc glycoforms of IgG are associated with human longevity**. *PLoS ONE* 2010, **5**:e12566.

14. Knežević A, Gornik O, Polašek O, Pučić M, Novokmet M, Redžić I, Rudd PM, Wright AF, Campbell H, Rudan I *et al.*: **Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans**. *Glycobiology* 2010, **20**:959-969.

15. Knežević A, Polašek O, Gornik O, Rudan I, Campbell H, Hayward C, Wright A, Kolčić I, O'Donoghue N, Bones J *et al.*: **Variability, heritability and environmental determinants of human plasma N-glycome**. *J Proteome Res* 2009, **8**:694-701.

16. Wright A, Charlesworth B, Rudan I, Carothers A, Campbell H: **A polygenic basis for late-onset disease**. *Trends Genet* 2003, **19**:97-106.

17. Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE *et al.*: **Replicating genotype–phenotype associations**. *Nature* 2007, **447**:655-660.

18. •• Ioannidis JP: **Why most published research findings are false**. *PLoS Med* 2005, **2**:e124.
Very important paper discussing why the majority of published candidate-gene studies reported false results.

19. Couzin J: **Genomics. The HapMap gold rush: researchers mine a rich deposit**. *Science* 2006, **312**:1131.

20. Cardon LR, Bell JI: **Association study designs for complex diseases**. *Nat Rev Genet* 2001, **2**:91-99.

21. • Visscher PM, Brown MA, McCarthy MI, Yang J: **Five years of GWAS discovery**. *Am J Hum Genet* 2012, **90**:7-24.
A short overview of acomplishments in the GWAS field in the past five years.

22. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R *et al.*: **Genome-wide association study identifies novel breast cancer susceptibility loci**. *Nature* 2007, **447**:1087-1093.

23. Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, Spain S, Penegar S, Chandler I, Gorman M, Wood W *et al.*: **A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21**. *Nat Genet* 2007, **39**:984-988.

24. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R *et al.*: **Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci**. *Nat Genet* 2010, **42**:1118-1125.

25. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, Lee JC, Goyette P, Imielinski M, Latiano A *et al.*: **Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47**. *Nat Genet* 2011, **43**:246-252.

26. McCarthy MI, Zeggini E: **Genome-wide association studies in type 2 diabetes**. *Curr Diab Rep* 2009, **9**:164-171.

27. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA: **Potential etiologic and functional implications of genome-wide association loci for human diseases and traits**. *Proc Natl Acad Sci U S A* 2009, **106**:9362-9367.

28. Clayton DG: **Prediction and interaction in complex disease genetics: experience in type 1 diabetes**. *PLoS Genet* 2009, **5**:e1000540.

Please cite this article in press as: Zoldoš V, et al.: Glycomics meets genomics, epigenomics and other high throughput omics for system biology studies, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.12.007

29. Aulchenko YS, Struchalin MV, Belonogova NM, Axenovich TI, Weedon MN, Hofman A, Uitterlinden AG, Kayser M, Oostra BA, van Duijn CM *et al.*: **Predicting human height by Victorian and genomic methods**. *Eur J Hum Genet* 2009, **17**:1070-1075.

30. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH: **Missing heritability and strategies for finding the underlying causes of complex disease**. *Nat Rev Genet* 2010, **11**:446-450.

31. Birney E: **Chromatin and heritability: how epigenetic studies can complement genetic approaches**. *Trends Genet* 2011, **27**:172-176.

32. Feinberg AP: **Epigenetics at the epicenter of modern medicine**. *JAMA* 2008, **299**:1345-1350.

33. • Suhre K, Shin SY, Petersen AK, Mohney RP, Meredith D, Wagele B, Altmaier E, Deloukas P, Erdmann J, Grundberg E *et al.*: **Human metabolic individuality in biomedical and pharmaceutical research**. *Nature* 2011, **477**:54-60.
An important paper which stresses the importance of human diversty and the need for individual approach.

34. Birney E, Lieb JD, Furey TS, Crawford GE, Iyer VR: **Allele-specific and heritable chromatin signatures in humans**. *Hum Mol Genet* 2010, **19**:R204-R209.

35. McDaniell R, Lee BK, Song L, Liu Z, Boyle AP, Erdos MR, Scott LJ, Morken MA, Kucera KS, Battenhouse A *et al.*: **Heritable individual-specific and allele-specific chromatin signatures in humans**. *Science* 2010, **328**:235-239.

36. Jaffe AE, Murakami P, Lee H, Leek JT, Fallin MD, Feinberg AP, Irizarry RA: **Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies**. *Int J Epidemiol* 2012, **41**:200-209.

37. Feinberg AP: **Genome-scale approaches to the epigenetics of common human disease**. *Virchows Arch* 2010, **456**:13-21.

38. Rakyan VK, Down TA, Balding DJ, Beck S: **Epigenome-wide association studies for common human diseases**. *Nat Rev Genet* 2011, **12**:529-541.

39. Huffman JE, Knezevic A, Vitart V, Kattla J, Adamczyk B, Novokmet M, Igl W, Pucic M, Zgaga L, Johannson A *et al.*: **Polymorphisms in B3GAT1, SLC9A9 and MGAT5 are associated with variation within the human plasma N-glycome of 3533 European adults**. *Hum Mol Genet* 2011, **20**:5000-5011.

40. • Lauc G, Essafi A, Huffman JE, Hayward C, Knežević A, Kattla JJ, Polašek O, Gornik O, Vitart V, Abrahams JL *et al.*: **Genomics meets glycomics — the first GWAS study of human N-glycome identifies HNF1alpha as a master regulator of plasma protein fucosylation**. *PLoS Genet* 2010, **6**:e1001256.
First genome wide study of the human glycome. Important both as a proof of principle and because it identified HNF1A as the master regulator of plasma protein fucosylation.

41. Kutalik Z, Benyamin B, Bergmann S, Mooser V, Waeber G, Montgomery GW, Martin NG, Madden PA, Heath AC, Beckmann JS *et al.*: **Genome-wide association study identifies two loci strongly affecting transferrin glycosylation**. *Hum Mol Genet* 2011, **20**:3710-3717.

42. Roxrud I, Raiborg C, Gilfillan GD, Stromme P, Stenmark H: **Dual degradation mechanisms ensure disposal of NHE6 mutant protein associated with neurological disease**. *Exp Cell Res* 2009, **315**:3014-3027.

43. Rivinoja A, Hassinen A, Kokkonen N, Kauppila A, Kellokumpu S: **Elevated Golgi pH impairs terminal N-glycosylation by inducing mislocalization of Golgi glycosyltransferases**. *J Cell Physiol* 2009, **220**:144-154.

44. Gornik O, Wagner J, Pučić M, Knežević A, Redžić I, Lauc G: **Stability of N-glycan profiles in human plasma**. *Glycobiology* 2009, **19**:1547-1553.

45. Mito Y, Henikoff JG, Henikoff S: **Histone replacement marks the boundaries of cis-regulatory domains**. *Science* 2007, **315**:1408-1411.

46. Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG *et al.*: **Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers**. *Cell* 2010, **142**:967-980.

47. Kannagi R, Sakuma K, Miyazaki K, Lim KT, Yusa A, Yin J, Izawa M: **Altered expression of glycan genes in cancers induced by epigenetic silencing and tumor hypoxia: clues in the ongoing search for new tumor markers**. *Cancer Sci* 2010, **101**:586-593.

48. Yusa A, Miyazaki K, Kimura N, Izawa M, Kannagi R: **Epigenetic silencing of the sulfate transporter gene DTDST induces sialyl Lewisx expression and accelerates proliferation of colon cancer cells**. *Cancer Res* 2010, **70**:4064-4073.

49. Kim YS, Deng G: **Aberrant expression of carbohydrate antigens in cancer: the role of genetic and epigenetic regulation**. *Gastroenterology* 2008, **135**:305-309.

50. Caretti A, Sirchia SM, Tabano S, Zulueta A, Dall'Olio F, Trinchera M: **DNA methylation and histone modifications modulate the beta1,3 galactosyltransferase beta3Gal-T5 native promoter in cancer cells**. *Int J Biochem Cell Biol* 2012, **44**:84-90.

51. Tong WG, Wierda WG, Lin E, Kuang SQ, Bekele BN, Estrov Z, Wei Y, Yang H, Keating MJ, Garcia-Manero G: **Genome-wide DNA methylation profiling of chronic lymphocytic leukemia allows identification of epigenetically repressed molecular pathways with clinical impact**. *Epigenetics* 2010, **5**:499-508.

52. Hill VK, Ricketts C, Bieche I, Vacher S, Gentle D, Lewis C, Maher ER, Latif F.: **Genome-wide DNA methylation profiling of CpG islands in breast cancer identifies novel genes associated with tumorigenicity**. *Cancer Res* 2011, **71**:2988-2999.

53. Saldova R, Dempsey E, Perez-Garay M, Marino K, Watson JA, Blanco-Fernandez A, Struwe WB, Harvey DJ, Madden SF, Peracaula R *et al.*: **5-AZA-2′-deoxycytidine induced demethylation influences N-glycosylation of secreted glycoproteins in ovarian cancer**. *Epigenetics* 2011, **6**:1362-1372.

54. Horvat T, Mužinić A, Barišić D, Bosnar MH, Zoldoš V: **Epigenetic modulation of the HeLa cell membrane N-glycome**. *Biochim Biophys Acta* 2012, **1820**:1412-1419.

55. Perez-Pinera P, Ousterout DG, Gersbach CA: **Advances in targeted genome editing**. *Curr Opin Chem Biol* 2012, **16**:268-277.

56. • Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, Amora R, Miller JC, Leung E, Meng X *et al.*: **Targeted genome editing across species using ZFNs and TALENs**. *Science* 2011, **333**:307.
An interesting article that presents potential of new technologies for genome editing.

57. Zoldoš V, Horvat T, Novokmet M, Cuenin C, Mužinić A, Pučić M, Huffman JE, Gornik O, Polašek O, Campbell H *et al.*: **Epigenetic silencing of HNF1A associates with changes in the composition of the human plasma N-glycome**. *Epigenetics* 2012, **7**:164-172.

58. Houle D, Govindaraju DR, Omholt S: **Phenomics: the next challenge**. *Nat Rev Genet* 2010, **11**:855-866.

59. Heard E, Tishkoff S, Todd JA, Vidal M, Wagner GP, Wang J, Weigel D, Young R: **Ten years of genetics and genomics: what have we achieved and where are we heading?** *Nat Rev Genet* 2010, **11**:723-733.

60. Lauc G, Zoldoš V: **Protein glycosylation — an evolutionary crossroad between genes and environment**. *Mol Biosyst* 2010, **6**:2373-2379.

61. Igl W, Polasek O, Gornik O, Knezevic A, Pucic M, Novokmet M, Huffman J, Gnewuch C, Liebisch G, Rudd PM *et al.*: **Glycomics meets lipidomics-associations of N-glycans with classical lipids, glycerophospholipids, and sphingolipids in three European populations**. *Mol Biosyst* 2011, **7**:1852-1862.

Please cite this article in press as: Zoldoš V, et al.: Glycomics meets genomics, epigenomics and other high throughput omics for system biology studies, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.12.007