Glycomics by High-throughput Integrated Technologies

FINAL REPORT

Grant Agreement number: *260600*

Project acronym: GlycoHIT

Project title: *Glycomics by High-throughput Integrated Technologies*

Funding Scheme: *FP7-CP-FP-SICA*

Project Start Date: *1/1/2011*

Project End Date: *30/6/2014*

Name, title and organisation of the scientific representative of the project's coordinator*: Prof. Lokesh Joshi NATIONAL UNIVERSITY OF IRELAND, GALWAY*

Tel: *+353 91 495768*

Fax: *+353 91 495768*

E-mail: *lokesh.joshi@nuigalway.ie*

Project website address: *http://www.glycohit.eu/*

Table of Contents

1. FINAL PUBLISHABLE SUMMARY REPORT 3

1.1 EXECUTIVE SUMMARY 3
1.2 A SUMMARY DESCRIPTION OF PROJECT CONTEXT AND OBJECTIVES 4
1.3 A DESCRIPTION OF THE MAIN S&T RESULTS/FOREGROUNDS 7
1.3.1 WP2 ADVANCED LC-MS GLYCO-ANALYTICAL METHODOLOGIES 7
1.3.2 WP3 MAMMALIAN LECTIN AND LECTIN MIMIC ARRAYS 11
1.3.3 WP4 OLIGONUCLEOTIDE ARRAYS AND GLYCOSIGNATURE GENERATION 13
1.3.4 WP5 LECTIN-GLYCAN INTERACTION ANALYSIS 15
1.3.5 WP6 GLYCOANALYSIS BIOINFORMATICS 16
1.3.6 WP7 LABEL-FREE SENSOR DEVELOPMENT 17
1.3.7 WP8 GLYCOMICS IN CANCER, A PILOT STUDY 22
1.4 IMPACT, DISSEMINATION AND EXPLOITATION. 28
1.4.1 OVERVIEW 28
1.4.2 WP2 ADVANCED LC-MS GLYCO-ANALYTICAL METHODOLOGIES 28
1.4.3 WP3 MAMMALIAN LECTIN AND LECTIN MIMIC ARRAYS 29
1.4.4 WP4 OLIGONUCLEOTIDE ARRAYS AND GLYCOSIGNATURE GENERATION 29
1.4.5 WP5 LECTIN-GLYCAN INTERACTION ANALYSIS 30
1.4.6 WP6 GLYCOANALYSIS BIOINFORMATICS 30
1.4.7 WP7 LABEL-FREE SENSOR DEVELOPMENT 30
1.4.8 WP8 GLYCOMICS IN CANCER – A PILOT STUDY 31
1.4.9 PUBLICATIONS AND DISSEMINATION ACTIVITES 32
1.5 WEBSITE AND CONTACT DETAILS 39

1. Final publishable summary report

1.1 *Executive summary*

GlycoHIT's aim has been to develop high throughput (HTP) technologies for analyzing protein glycosylation in clinical samples, specifically for the **rapid, accurate, early diagnosis of cancer**. Glycosylation is a type of protein post-translational modification (PTM) where cell surfaces and proteins are modified with sugar molecules. Glycosylation is critical for the physiological state of cells, systems and organs and for relating genomic and proteomic measurements to their biological and functional contexts. Understanding glycosylation is essential for deciphering normal and diseased physiological processes, diagnostic markers and for developing effective treatments for a range of diseases.

Despite extensive research on cancer over the past decades, cancer glycobiology has received comparatively little attention primarily due to the lack of robust high-throughput (HTP) tools for glycoanalysis. The few cancer biomarker assays that are currently used or are undergoing validation involve differentially glycosylated glycoconjugates on the tumour cell surface or in the serum. Thus, it is highly likely that a considerable number of glycobiomarkers remain to be discovered and validated for various cancers. Furthermore, glycosylated membrane receptors are a very important class of potential cancer drug targets. The development of alternative HTP approaches for analysis of serum glycoconjugates in health and disease is imperative if the goals of early detection and improved therapies for cancer are to be realised.

Existing technologies for analyzing protein glycosylation, such as high performance liquid chromatography (HPLC) and mass spectrometry (MS), are suitable for discovery research purposes. However, further advances in existing technologies and novel technical platforms are needed that can perform HTP analysis of protein glycosylation. The GlycoHIT consortium is advancing the state of the art to **develop future technologies that enable faster and multiplexed analysis of glycosylation** on proteins from clinical and biopharmaceutical samples.

The overall focus of this project has been to provide an array of complementary and innovative HTP glycoanalytical technologies, for the analysis of cancer-associated glycobiomarkers and glycoprotein therapeutics. During the project, the following tools and technologies were developed:

- Improved **sample preparation procedures, pre-analysis separation steps, chromatographic separation, and peak identification algorithms employed in HPLC/LC-MS methods**, facilitating higher throughput of samples during discovery experiments, more reliable identification of potential glycobiomarkers, and generation of validation data. Together, these approaches have resulted in more focused glycoanalysis, improved specificity of structural determination and faster throughput.
- **A new generation of lectin array platforms** were developed for rapid, robust and reliable detection of glycobiomarkers and analysis of glyco-profiles in large sample sets. The main method was by exploring novel sources of specific glycan recognition molecules, including recombinant mammalian lectins, combinatorial libraries of antibody fragments/peptides and random oligonucleotide libraries. Selection of optimal lectin mimics was also done employing a **rational design approach using lectin-protein modeling and experimental data from lectin-protein analysis**.
- **Label-free sensor platforms for glycan-lectin-based assays** were developed to facilitate the analysis of cancer-associated glycobiomarkers in a rapid, specific and sensitive manner, opening the door to routine use in clinical practice.

During the project, these new technologies have been tested for use in routine clinical diagnostics and therapeutic monitoring using known glycobiomarkers. During validation, novel cancer glycobiomarkers candidates were discovered. GlycoHIT has developed and tested innovative HTP glyco-analytical platforms, based on the novel application of specific glycan binders, and has significantly advanced the state of the art. Combination of the GlycoHIT technologies with other *omic technologies will open new doors in the life sciences and the impact in understanding disease processes is expected to be substantial.

1.2 *A summary description of project context and objectives*

Glycosylation is a type of protein post-translational modification (PTM) where cell surfaces and proteins are modified with sugar molecules. Glycosylation is critical for the physiological state of cells, systems and organs and for relating genomic and proteomic measurements to their biological and functional contexts. Understanding glycosylation is essential for deciphering normal and diseased physiological processes, diagnostic markers and for developing effective treatments for a range of diseases.

Despite extensive research on cancer over the past decades, cancer glycobiology has received comparatively little attention primarily due to the lack of robust high-throughput (HTP) tools for glycoanalysis. The few cancer biomarker assays that are currently used or are undergoing validation involve differentially glycosylated glycoconjugates on the tumour cell surface or in the serum. Thus, it is highly likely that a considerable number of glycobiomarkers remain to be discovered and validated for various cancers. Furthermore, glycosylated membrane receptors are a very important class of potential cancer drug targets. The development of alternative HTP approaches for analysis of serum glycoconjugates in health and disease is imperative if the goals of early detection and improved therapies for cancer are to be realised. GlycoHIT's aim has been to develop high throughput (HTP) technologies for analyzing protein glycosylation in clinical samples, specifically for the **rapid, accurate, early diagnosis of cancer**.

In excess of 10 million new cases of cancer are reported globally every year, making it one of the most devastating diseases worldwide. Early detection of cancer, accompanied by prompt initiation of treatment, can result in increased survival rates, higher rates of remission/lower rates of relapse, reductions in invasive testing and improved quality of life for the patient. For these reasons, sensitive and accurate diagnostic tests for detecting specific cancers are required. A **considerable knowledge gap remains in the field of tumour detection, diagnosis and prognosis, and validation of therapeutic and disease management** due to a lack of HTP technologies.

Existing technologies for analyzing protein glycosylation, such as high performance liquid chromatography (HPLC) and mass spectrometry (MS), are suitable for discovery research purposes. However, further advances in existing technologies and novel technical platforms are needed that can perform HTP analysis of protein glycosylation. The GlycoHIT consortium is advancing the state of the art to **develop future technologies that enable faster and multiplexed analysis of glycosylation** on proteins from clinical and biopharmaceutical samples.

The overall focus of this project has been to provide an array of complementary and innovative HTP glycoanalytical technologies, for the analysis of cancer-associated glycobiomarkers and glycoprotein therapeutics. During the project, these new technologies have been tested for use in routine clinical diagnostics and therapeutic monitoring using known glycobiomarkers. During validation, novel cancer glycobiomarkers candidates have also been discovered. GlycoHIT has developed and tested innovative HTP glyco-analytical platforms, based on the novel application of specific glycan binders, and has significantly advanced the state of the art.

The main objectives of GlycoHIT have been as follows:

- To improve the **sample preparation procedures, pre-analysis separation steps, chromatographic separation, and peak identification algorithms employed in HPLC/LC-MS methods,** facilitating higher throughput of samples during discovery experiments, more reliable identification of potential glycobiomarkers, and generation of validation data. Most cancer biomarkers to date have been identified using HPLC and MS and there is no obvious alternative to these methods for structural characterisation of glycans. A sensitive, robust ultra performance liquid chromatography (UPLC)-hydrophilic interaction chromatography (HILIC) method for *N*-linked glycan analysis and automated methodology for analysis of LC-MS data from glycosylated proteins were developed. Both manual and robotic platforms linked to UPLC for precise and HTP *N*-linked glycan screening of clinical samples were developed. Using these platforms, all peaks for *N*-linked glycans in human serum were assigned. Pre-fractionation of the serum by isoelectric focusing (IEF) was combined with LC-MS and

optmised informatic approaches. A HTP method that identifies *N*-linked glycosites on proteins while simultaneously identifying the precise glycan structure attached to each site was also developed and utilised in these analyses. Together, these approaches have resulted in more focused glycoanalysis, improved specificity of structural determination and faster throughput.

- To contribute to the **development of a new generation of lectin array platforms** for rapid, robust and reliable detection of glycobiomarkers and analysis of glyco-profiles in large sample sets. The main method was by exploring novel sources of specific glycan recognition molecules, including recombinant mammalian lectins, combinatorial libraries of antibody fragments/peptides and random oligonucleotide libraries. A mammalian lectin microarray was successfully established and lectins were shown to be functional and carbohydrate-specific in this format. Phage-display-derived single chain variable fragments (scFvs) were generated against specific glycan targets and were shown to be exquisitely specific. These scFvs were incorporated into the microarray platform and shown to be functional. They were used to determine the substitution value for a natural glycoprotein in this platform, which was comparable to that obtained using the traditional method of enzymatic digestion and HPLC analysis. Using SELEX technology, aptamers were generated with relative specificity and high binding affinities for glycan targets. Selection of optimal lectin mimics was also done employing a **rational design approach using lectin-protein modeling and experimental data from lectin-protein analysis**.`
- To **develop label-free sensor platforms for glycan-lectin-based assays** to facilitate the analysis of cancer-associated glycobiomarkers in a rapid, specific and sensitive manner, opening the door to routine use in clinical practice. A range of potential carbohydrate biosensor platforms were established and evaluated, with detection based on EIS, ECL with and without nanoparticle-enhancement, plasmonic coupled SERS, and a flow cytometry-based suspension array assay. The use of some of these platforms for the *in situ* and dynamic evaluation of cell surface carbohydrate expression have been demonstrated.
- To **validate the assays developed during the project using clinically relevant sample sets** from fully characterised patient (breast, colon, and head and neck cancer) and control groups and to validate the application of these analytical technologies to glyco-profiling of recombinant glycoprotein drugs. Four HTP techniques to analyse glycan-related features have been applied to profile clinical materials for validation. The identified profiles include several traits not previously identified. Preliminary analyses indicate that lectin microarray profiling of serum glycans and HTP proteomics analyses of tumour proteins may be used to identify subtype-specific alterations in breast cancer patients. Treatment-specific alterations in the serum *N*-glycome has been identified for neoadjuvantly treated patients. Thus, technologies developed in this consortium have shown sensitivity and specificity sufficient to reveal clinically significant biological alterations in the proteome and *N*-glycome. Immunostaining technique used to analyse diverse clinical data (naso-sinusal diseases, thyroid lesion and head and neck cancers) allowed the identification of several interesting phenomena. Gal-9 is over-expressed in naso-sinusal squamous cells carcinomas, biotinylated galectin (-1, -3, -7 and -8) expression allows a differential diagnostic in thyroid lesions between follicular adenoma and papillary carcinomas and in head and neck cancers (oral cavity, pharynx and larynx) galectin-2 expression increases with tumour progression. In addition, CHO cell glycosylation was analysed by UPLC and CHO cell supernatant from three processes were analysed by a combination of UPLC and lectin microarray technologies. Purified biopharmaceutical samples from various stages of the production were analysed using the novel carbohydrate recognition molecules developed during this project.

The combined approaches undertaken in GlycoHIT have been applied to a specific area in understanding the glycoprofiles of clinical material from patients with cancer. This has helped to generate data that will contribute to the rapid and accurate diagnosis of diverse clinical conditions associated with various cancers that had previously not been possible. This will have a direct consequence in being able to guide

the utilisation of specific therapies for the relevant clinical stage of disease. GlycoHIT has contributed to an increase in the basic and fundamental knowledge of cell surface cancer-associated biomarkers which may have implications in the general understanding of cancer-associated disease that will have a significant impact in the development of specific intervention therapies.

The platforms and tools developed as part of this project will also have direct relevance to the research and development activities in the diagnostics and biopharmaceutical sectors. Identification and detection of disease-associated glycobiomarkers and the monitoring of drug reactions will be greatly facilitated once the information associated with the glycosylation of glycoproteins and glycolipids can be rapidly extracted from samples.

The technologies developed by GlycoHIT will have a major impact on life sciences research. Whereas HTP proteomic and genomic analysis have provided new tools to researchers, similar tools have not previously been available for glycans due to their inherently complex nature. GlycoHIT has now addressed this technology gap and has added an important group of tools to the life sciences research toolkit. This HTP approach to glycoanalysis will now enable researchers to monitor the presence and evolution of biomarkers. Combination of the GlycoHIT technologies with other *omic technologies will open new doors in the life sciences and the impact in understanding disease processes is expected to be substantial.

1.3 *A description of the main S&T results/foregrounds*

1.3.1 WP2 Advanced LC-MS glyco-analytical methodologies

Highly reproducible high-throughput manual and robotic glycan analysis using HILIC-UPLC

We have developed both a manual and highly reproducible robotic methods utilizing UPLC which were used for analysis of samples from this WP2 and as well as in WP8. Figure 1 shows the reproducibility of replicate serum samples run by the manual and robotic method and their comparison.

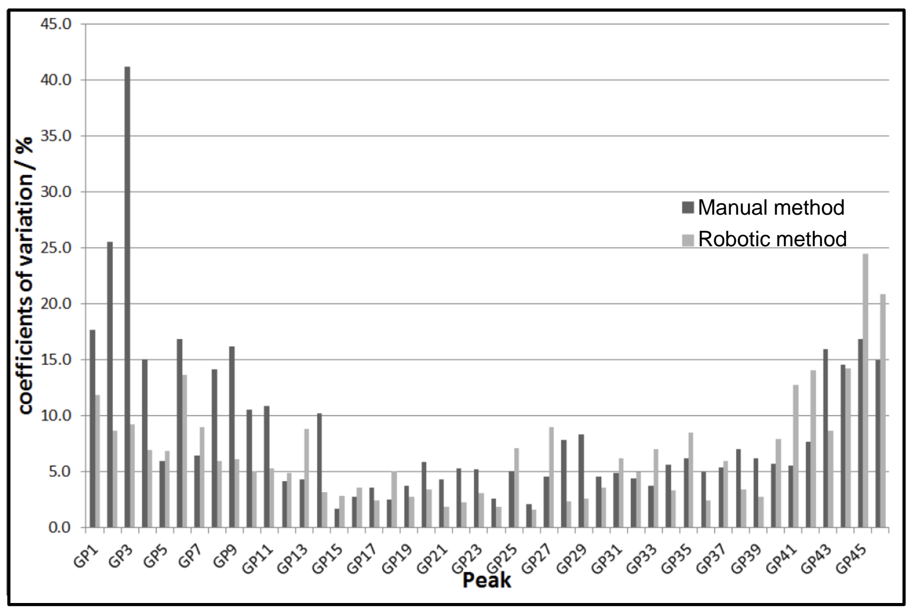


Figure 1 Comparison of the glycan area's coefficients of variation between the automated assay and the manual method. Samples were prepared automatically or using the manual method in quadruplicate on four different days and analysed by UPLC (Stockmann et al, in preparation).

Assignments of serum N-glycome were also finalized (Figure 2) and published in Saldova et al. (*Saldova R, Asadi Shehni A, Hakanssen VD, Hiliard M, Steinfeld I, Yakhini Z, Borresen-Dale AL, Rudd PM. Association of N-Glycosylation with Breast Carcinoma and Systemic Features Using High-Resolution Quantitative UPLC, JPR, 2014, 13: 2314-2327*).

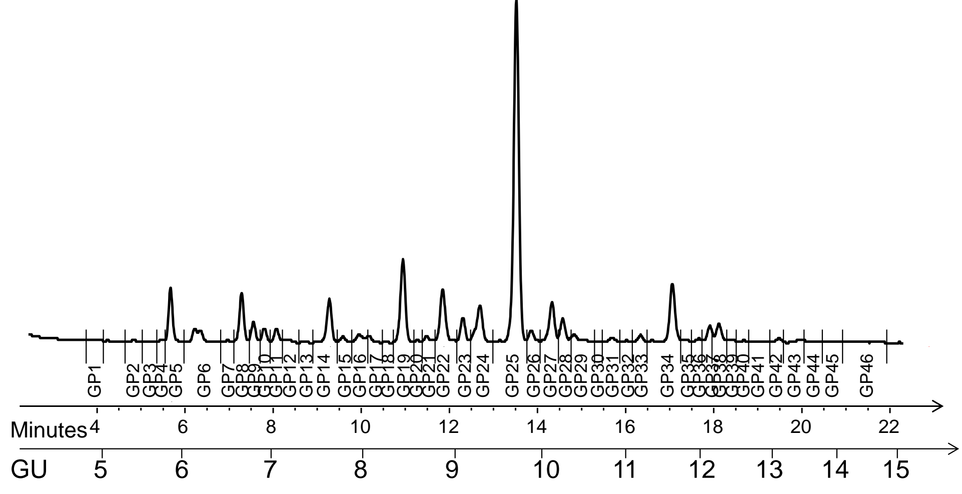


Figure 2 HILIC–UPLC chromatogram from NHS pooled serum sample and its separation into 46 peaks and significant differences between patients and control groups of interest.

A novel method to sign glycan structures to respective glycoproteins and glycosylation sites

Collaboration among Partner 4 (NIBRT), Partner 5 (KI) and Partner 2 (Agilent) we have developed a high throughput method (glycan-UPLC combined with glycopeptide IEF-LC-MS) that identifies N-linked glycosites on proteins while simultaneously identifying the precise glycan structure (or structures if a variety of glycans are present) attached to each site. We have proved this principle on the 14 most abundant proteins in human serum sample where multiple aliquots of blood plasma were passed through an Agilent MARS-Hu14 depletion column, the eluted fraction was collected containing the 14 highly abundant proteins and separated using high resolution IEF into 72 fractions containing glycopeptides, part of which was analysed on HILIC-UPLC for N-glycosylation analysis for each fraction. The N-glycome was separated into 46 peaks (GP1-46) which was the same number of peaks detected in the NHS control serum sample (Figure 2) assigned and published in Saldova et al. (*Saldova R, Asadi Shehni A, Hakanssen VD, Hiliard M, Steinfeld I, Yakhini Z, Borresen-Dale AL, Rudd PM. Association of N-Glycosylation with Breast Carcinoma and Systemic Features Using High-Resolution Quantitative UPLC, JPR, 2014, 13: 2314-2327*). The remaining sample was fractionated by HiRIEF into 72 fractions and then separated into two parts- one digested with PNGase F yielding non-glycosylated image and the second part yielding glycosylated image. The N-glycan analysis of fractions was performed by HILIC-UPLC, the peptides were analysed using HiRIEF-LC-MS/MS analysis and then matched with glycopeptides. Protein- and site-specific glycosylation was then bioinfomatically assessed in HTP algoritm developed by Partner 2 (Agilent). This computational analysis resulted in 147 different identified glycoforms. In total, 29 distinct glycosites and 22 distinct glycosylated proteins were identified, mostly the ones directly targeted by the MARS14 column, as well as a few other low abundance glycoproteins. We illustrate our findings with one example from the protein, Haptoglobin, the glycan-harbouring peptide VVLHPnYSQVDIGLIK, and show the relevant extracted ion chromatograms and spectra (Figure 3). Several novel glycoforms were detected in this fractionation for several glycosites that have been previously published confirming the validity of the method.

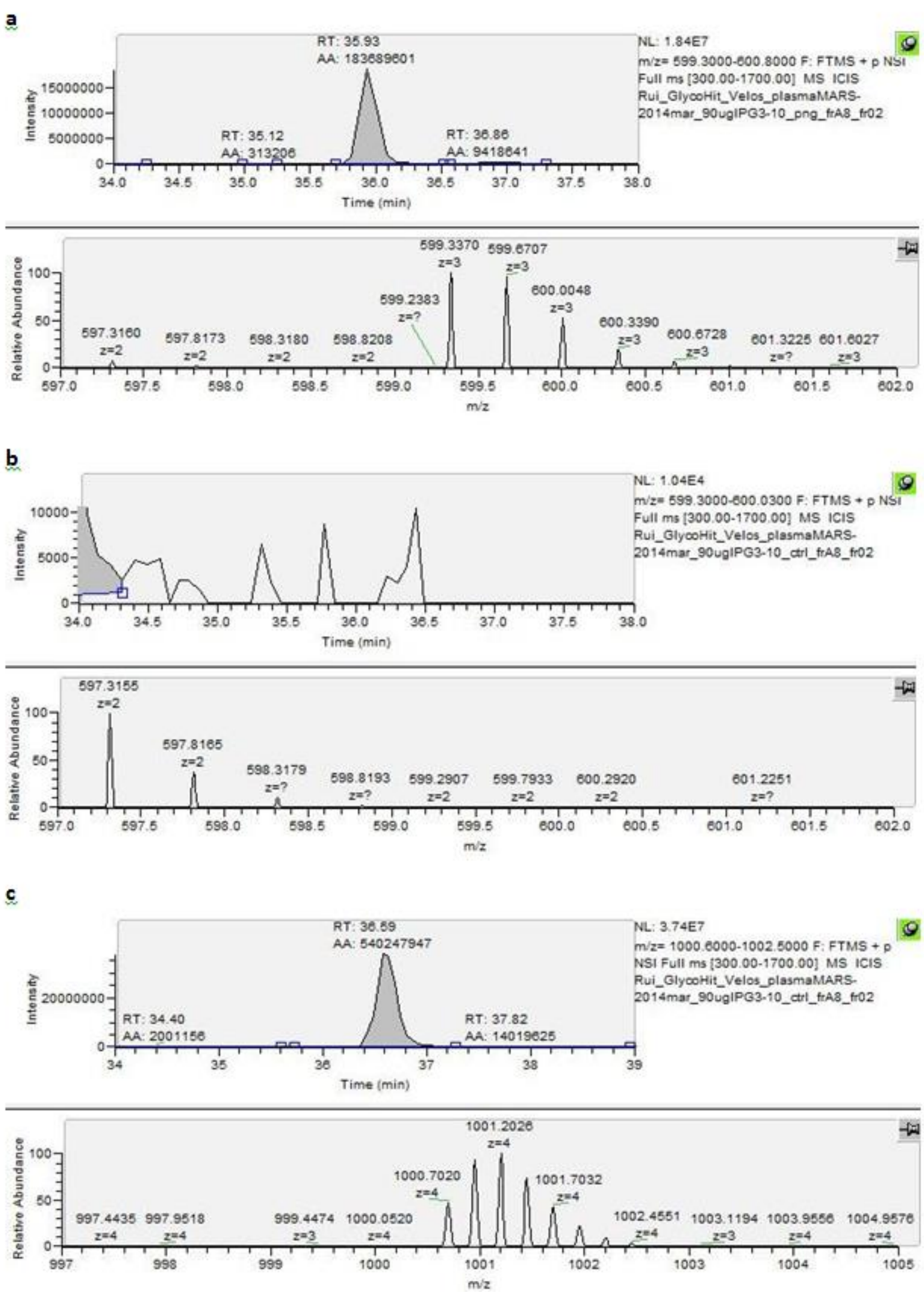


Figure 3 Extracted ion chromatograms (XIC) and isotope cluster spectra for the deamidated and glycosylated forms of peptide VVLHPnYSQVDIGLIK. a) Deamidated form of the peptide in the PNGase F treated sample. b) Such deamidated form is absent from the control sample (only baseline noise is shown). c) Glycoform of the peptide (with glycan A2G2S2) in the control sample

A successful method to increase analytical depth in proteomics (HiRIEF LC-MS/MS) to allow detection of glycosylation enzyme machinery was achieved, both on analysis of cellular proteome (WP2, Branca et al., Nature Methods, 2014) and plasma proteome (WP2, Pernemalm and Lehtiö, JPR 2013). This method was

used to combined analysis of proteome and related changes glycan changes when modulating p16 tumor suppressor expressions in cancer cells (Amano et al. FEBS 2013). This work was carried out with Partners 3 and 16 (LMUM,HUJ).The in-depth proteomics method was also applied to the analysis of breast cancer tumor samples (WP8) together with Partner 7 (Oslo), where this proteomics dataset complements glycomics datasets and on yeast proteome analysis together with Ping Xu, of the Beijing Proteome Institutet (collaboration with Chinese sister project to GlycoHIT). Further, a novel method to achieve complete glycoprotein analysis (matching specific glycan structures with protein and glycosylation site), in parallel up to 10 proteins, was achieved (Manuscript in preparation (Branca et al.). In summary, this work has contributed significantly with novel methods to connect glycan analysis to proteome analysis to achieve multilayer omics analysis and help to connect the researchers cross these fields to solve systems biological questions related to glycoproteomics.

In-depth quantitative proteomics methods using peptide isoelectric focusing

The aim of the method is to perform comprehensive quantitative proteomics to allow informative multilayer omics data analysis including proteome data as demonstrated by a breast cancer dataset as well as A431 and Capan-1 cell line datasets generated. We have developed a method High Resolution Isoelectric focusing (HiRIEF) based prefractionation of complex proteome samples prior LC-MS/MS analysis. Briefly the method is based on the following steps: 1) Generation of experimental fractions containing peptides within a defined isoelectric point (pI)-range though high resolution peptide isoelectric focusing (HiRIEF). 2) Generation of LC-MS/MS data from pI-range defined fractions. 3) Prediction of the pI of all theoretical peptides generated by the 6-ORF translation of the human genome using a peptide pI prediction algorithm developed by us. 4) Generation of pI-range restricted databases from all theoretical peptides in the human genome. 5) Search the LC-MS/MS data from pI-range defined fractions against the corresponding pI-range restricted database. This approach results in a dramatic search space reduction, which for the first time ever enables unbiased matching of experimental peptide spectral data to the entire human genome sequence. The power of this method was proven as we were able to find 99 novel protein coding loci in human cells (Branca et al., Nature Methods). In addition, this experiment provided one of the deepest proteome coverage of annotated genes in human and mouse up to date, further demonstrating the capacity of the HiRIEF-LC-MS/MS workflow. The robustness of the method on unbiased novel protein search was demonstrated by us performing the analysis on mouse discovering several new protein coding genes in mouse as well.

Use of the in-depth method for parallel glycan and proteome analysis to elucidate mechanism of N-glycosylation changes detected by p16 deletion

The work solidifies the physiological impact of a tumor suppressor on modulating glycan expression, (joint research with **Partner 3**, LMUM). The tumor suppressor $p16^{INK4a}$ is known to exert cell-cycle control via cyclin-dependent kinases, N/O-glycosylation and galectin expression in an orchestrated manner. Hereby, $p16^{INK4a}$ restores susceptibility to anoikis induction in human Capan-1 pancreatic carcinoma cells. Using chemoselective N/O-glycan enrichment technology (glycoblotting) and product characterization to define structural changes, we first verified a substantial decrease in sialylation for both classes of oligosaccharides. Tests with genetic (transfection with α2,6-sialyltransferase-specific cDNA) or metabolic (medium supplementation with *N*-acetylmannosamine to track down a bottleneck specific for sialic acid biosynthesis) engineering and also cytofluorometric analysis of lectin binding pointed to a role of limited substrate availability, especially for α2,6-sialylation, which switches off reactivity for anoikis-triggering homodimeric galectin-1. Quantitative mass spectrometry analysis of protein level changes confirmed enhanced galectin-1 presence along with an influence on glycosyltransferases (β1,4-galactosyltransferase-IV, α2,3-sialyltransferase-I) and detected $p16^{INK4a}$-dependent downregulation of two enzymes in the biosynthesis pathway for sialic acid, i.e. the bifunctional UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase (GNE) and N-acetylneuraminic acid 9-phosphate synthase ($P < 0.001$). In contrast, quantitative assessment for presence of nuclear CMP-N-acetylneuraminic acid synthase, responsible for providing the donor for enzymatic sialylation which also acts as feed-back inhibitor of the epimerase activity of GNE, revealed a trend for an increase. GNE transfection restored partially the level of Sialylation and decreased

the extent of anoikis induction in Capan-1 cell with stable expression of p16^{INK4a}. Thus, downregulation of two enzymes in sialic acid biosynthesis pathway acting consecutively but not an effect on gene expression of α2,6-sialyltransferase reduced cell reactivity to the growth effector galectin-1.

In-depth plasma analysis

To increase sensitivity and analytical depth in shotgun proteomics, pre-fractionation of complex samples is often used. We developed a novel pre-fractionation method, Sandwich high resolution isoelectric focusing, which combines both protein and peptide isoelectric focusing and demonstrate its use in plasma analysis (Pernemalm and Lehtiö, JPR, 2013). In the first step, intact proteins are separated based on isoelectric point (pI) using traditional immobilized pH gradient (IPG) strips. Segments in the IPG-strip containing proteins of interest are subsequently cut out and applied to in-strip digestion, without subsequent peptide elution. In the second peptide isoelectric focusing step, the strip segments are used as loading bridges. The peptides are thereby directly applied to the peptide isoelectric focusing, without an intermediate elution step, and separated on narrow range IPG strips to reduce the complexity on the peptide level. In the final step, the peptides are eluted into 96-well plates and analyzed with mass spectrometry. In a proof of principle experiment, using this method to zoom in on pI regions of interest in human plasma, we identify over 800 proteins, with concentrations spanning over six orders of magnitude.

1.3.2 WP3 Mammalian lectin and lectin mimic arrays

Glycan-binding proteins or lectins have been used for over four decades for the characterisation of cell surface glycans and glycoproteins, based on their selectivity for certain glycan structures and linkages. Due to their robust nature, plant lectins are widely used in diverse settings, such as blood group typing, bacterial identification, cell selection and glycoprotein purification. Lectin microarray-based glycan profiling has given promising results with a range of sample types, including glycopeptides, glycoproteins, cell surface glycome, tissue sections and bacteria. In most cases, the target glycans or glycoconjugates are fluorescently labelled and the bound target is detected using laboratory microarray scanners. However, the small number of plant lectins (typically 20 to 60) used in these arrays and the common problem of the wide carbohydrate specificity of plant lectins limit the structural information that they can produce. Most importantly for clinical applications, plant lectins do not share the specificity profiles of human lectins, are not specific enough and thus may not always detect functionally relevant glycosylation alterations.

In this WP, **Partner 1** (NUI Galway) are pioneering the use of recombinant, clinically relevant human lectins in lectin microarrays on the premise that the binding specificities of recombinant human lectins are much more relevant to clinical sample analysis than plant lectins. This WP formed a central part of the GlycoHIT programme. A panel of 16 mammalian lectins involved in adhesion, migration, proliferation inflammation and immune response including recombinant selectins, siglecs, C-type lectins and ficolins, were arrayed and the function of each lectin tested. Conditions of printing, storage and incubation were optimised to generate a functional platform which was inhibitable with appropriate carbohydrates. This platform was used to glycoprofile healthy and breast cancer patient serum samples in a HTP manner and this work is detailed in WP8 below.

In-house production of labeled recombinant galectins and galectin fragments was optimised by **Partner 3** (LMUM) and these tools were used as tumour markers in naso-sinusal disease, benign and malignant thyroid lesions and head and neck cancers by **Partners 8 and 9**, which is further detailed in WP8 below. In addition, the glycan interactions of the recombinant lectins were studied by **Partner 10** (CSIC Madrid) which is detailed in WP5. **Partner 10** also constructed microarray composed exclusively of recombinant galectins and galectin fragments generated by **Partner 3** (LMUM). Inhibition assays with lactose, asialofetuin or a mixture of both were done to assess the specificity of the interaction. The behaviour of different galectins under the assayed conditions was galectin dependent. Some galectins, as for example HG1, completely lost their carbohydrate-binding capability when immobilized. In contrast, other galectins, e.g. HG8 and HG9-N, gave strong binding signals. However, the binding was only partially inhibited, up to different degrees, by lactose/ asialofetuin under the tested conditions.

In addition, **Partner 1** used phage display technology to identify glycan-binding molecules to augment the feature space on their lectin microarray platform and for use in the development of convenient assays for specific glycan motifs. The generation and characterisation of three anti-Gal-a-(1→3)-Gal single chain variable fragments (scFvs) was published (*Cunningham, S., Starr, E., Shaw, I., Glavin, J., Kane, M., & Joshi, L. Development of a convenient competitive ELISA for the detection of the free and protein-bound nonhuman galactosyl-α-(1,3)-galactose epitope based on highly specific chicken single-chain antibody variable-region fragments. Analytical Chemistry, 201385(2), 949–55).* These scFvs was found to be extremely specific upon screening on a neoglycoconjugate (NGC) microarray. The concentrations of the Gal-α-(1→3)-Gal-containing free disaccharide, trisaccharides and tetrasaccharide that caused 50% displacement (i.e. the ED_{50}) were calculated in a competitive ELISA assay (Table 1). This suggested that all scFvs had a higher affinity for the larger oligosaccharides than for the disaccharide.

Table 1 Specificity study of the three competitive ELISAs based on the anti-Gal-α-(1→3)-Gal-BSA scFvs.

		ED_{50} (mM)		
free sugar	M_r	scFv A11	scFv A4	scFv G12
Gal-α-(1,3)-Gal	342.3	3.58	0.91	0.46
Gal-α-(1,3)-Gal-β-(1,4)-Gal	504.4	0.08	0.16	0.21
Gal-α-(1,3)-Gal-β-(1,4)-Glc	504.4	0.07	0.12	0.20
Gal-α-(1,3)-Gal-β-(1,4)-Gal-α-(1,3)-Gal	666.5	0.14	0.26	0.12

No inhibition was observed in the presence of any monosaccharide. ScFv A4 gave the most consistent readings and was used to detect and quantify the Gal-α-(1→3)-Gal motif when bound to proteins in the form of NGCs and on both the glycoproteins tested, Erbitux and laminin (Table 2).

Table 2. Comparison of the molar ratios of Gal-α-(1,3)-Gal to protein in selected glycoproteins.

	Gal-α-(1,3)-Gal:protein molar ratio	
glycoprotein	competitive ELISA	MALDI/other
Gal-α-(1,3)-Gal-BSA	9.4	16 (10–25)[a]
Gal-α-(1,3)-Gal-BSA	19.7	23 (15–31)[a]
laminin	38.9	ND[b]
Erbitux	3.9	1.13[c]

[a]*Mean (range) determined by MALDI analysis.*
[b]*The Gal- α -(1→3)-Gal content of native laminin has not been reported*
[c]*Value determined following detergent denaturation and enzymatic release of oligosaccharides*

The ELISA-estimated values for the NGCs were very close to or within the range of the MALDI results provided by the manufacturer. Both laminin and Erbitux in their native state were too viscous to be analysed in the competitive ELISA. However, heat treatment reduced the viscosity, and assay results indicated that there were 39 mol of the Gal-α-(1→3)-Gal residue/mol of murine laminin and approximately 3.9 mol/mol of Erbitux, in the same range as recently reported. The content of Gal-α-(1→3)-Gal in murine laminin has not previously been reported to the best of our knowledge.

A panel of three clones was generated from a chicken immune scFv library prepared from birds immunized with the non-human sialic acid, *N*-glycolylneuraminic acid (Neu5Gc). Characterisation studies are on-going but the clones appear to specifically binding to Neu5Gc.

The anti-Gal-α-(1→3)-Gal scFvs were also covalently immobilised on a microarray surface and several immobilisation conditions were assessed and optimised. Different conditions for scFv performance were also assessed and optimised and the scFvs were found to maintain their function and specificity in this format. In this platform, scFv G12 performed best, followed by A4 and finally A11. Inhibition studies were performed to generate half maximal inhibitory concentration (IC_{50}) for the inhibition of binding of labelled Gal-α-(1→3)-Gal–BSA NGC to the immobilised scFv using Gal monosaccharide and Gal-α-(1→3)-Gal-containing disaccharide, trisaccharides and tetrasaccharide and no inhibition was observed using the monosaccharide. As quantification of the substitution of the Gal-α-(1,3)-Gal structure on glycoproteins is of great importance in the biopharmaceutical industry, Gal-α-(1,3)-Gal substitution on the natural glycoprotein bovine thyroglobulin (BT) was successfully quantified using the scFv microarray. A manuscript describing this work is in preparation for publication.

Partners 1 and 2 (Agilent) have undertaken the rational re-design of existing viral sialic acid binding proteins to generate modified proteins that differentiate between the most common sialic acid variants, Neu5Ac and Neu5Gc. This work is detailed in WP6.

1.3.3 WP4 Oligonucleotide arrays and glycosignature generation

The overall goal of this WP was to develop an oligonucleotide microarray-based platform for the generation of glycosignatures for the total serum glycoprotein pool, selected sub-fractions or specific serum glycoproteins, to enable discrimination of diseased and control individuals. The first stage in this process was the selection and characterization of specific oligonucleotide sequences that bind selected glycan moieties from random oligonucleotide libraries using Systematic Evolution of Ligands by EXponential enrichment (SELEX) technology. These aptamers were to act as controls during development of the microarray platform and also represented the starting sequences for further expansion of the oligonucleotide aptamer microarray feature space by rational design.

Random DNA and RNA oligonucleotide libraries were screened using standard SELEX procedures against selected glycan motifs and RNA and DNA-based aptamers which recognised H type 2 antigen (H2), the Lewis b determinant and the non-human sialic acid, Neu5Gc were selected during this WP. The former were targets for selection of RNA aptamers and the latter two for DNA aptamers.

The proof-of-concept experiments for off the shelf oligonucleotide microarrays from Agilent Technologies, Inc., were performed using structurally related oligosaccharides labeled with nanoparticles and fluorescently labeled glycoforms of bovine fetuin, a well-characterised model glycoprotein with both *N*- and *O*-linked glycosylation. which enabled the identification of bottlenecks to overcome in order to progress in the use of oligonucleotide microarrays for glycosignaturing. Most importantly, the development of a robust aptamer assay is required as it is difficult to pinpoint optimized steps with no assay or ability to check optimal aptamer performance. In the process of developing this HTP aptamer microarray approach, several additional factors have been identified that also require further optimisation, but are similarly dependent on a successful validation assay; 1) efficient carbohydrate amination, 2) optimisation of incubation conditions and reagents, 3) identification and use of positive controls, 4) enhancement of signal to noise ratio and 5) choice of label and detection method for binding event.

Thus, further assay development using the DNA aptamers against Neu5Gc was undertaken during this period. Several strategies were pursued to develop appropriate aptamer assays including biotinylated aptamers used in ELISA-type assays and surface plasmon resonance (SPR) instruments, and unlabelled aptamers detected with a DNA-binding fluorescent dye in binding and inhibition assays. A most desirable targeted assay format would be an ELISA type assay which would be transferable to several platforms including SPR. The choice of label and detection method for monitoring the binding event is a very important issue for aptamer assay development. The conjugated π-systems in the vast majority of fluorescent labels interfere with DNA structure and thus binding despite fluorescent labels usually allowing the most sensitive detection in assays. Biotin is a frequently used alternative, but a secondary detection step is necessary and only strong binding can be detected using this label. A binding assay based on the selection method was the

final successful assay which made use of the OliGreen ssDNA dye as an ultra sensitive fluorescent nuclei acid stain for quantification. Successful assay development will pave the way for use of the oligonucleotide microarray as a glycosignaturing and screening platform.

The assays demonstrated relative specificity and high binding affinities for Neu5Gc. To our knowledge this is the first report of DNA aptamer binding assay that shows differentiation between Neu5Gc and Neu5Ac. The aptamer R9-Gc can detect free Neu5Gc at a very low level in ng/mL. We have also demonstrated the detection of Neu5Gc in a glycoprotein sample. Thus, the assay provides the option of either direct analysis of the Neu5Gc on a glycoprotein of interest or measurement of the free Neu5Gc in solution.

To progress the overall goals of GlycoHIT, plant lectin microarrays were validated as glycosignaturing platforms and this work was published *Gerlach, J.Q., Kilcoyne, M., Joshi, L. Microarray evaluation of the effects of lectin and glycoprotein orientation and data filtering on glycoform discrimination. Anal. Methods, 2014, 6, 440-449).* Bovine fetuin (FET) was selected as the model glycoprotein and various glycoforms were generated. In addition, to evaluate the possibility of real work interference with lectin binding by the presence of free mono- and oligosaccharides which compete with analytes, five partially-inhibitory conditions were also included to evaluate the impact of blocking select lectin interactions on glycoform discrimination. Stringent filtering was applied to allow selection of a subset of lectins which could be used for glycoform detection even under these interference conditions. The lectin microarray data for the selected lectin subset was compared to the reported specificities of these lectins to evaluate their correlation to presented glycoform structures.

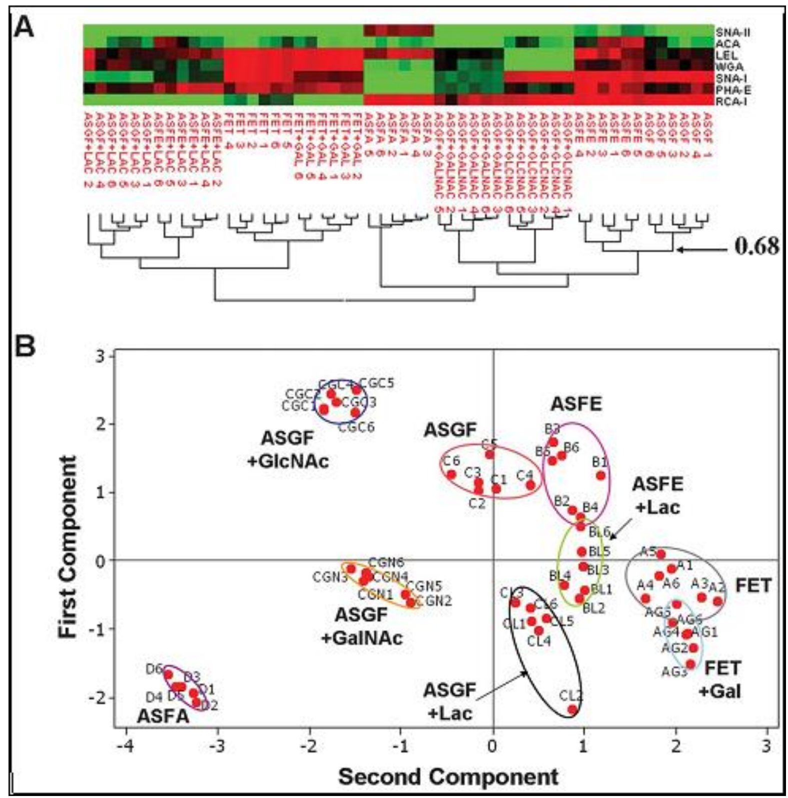


Figure 4 (A) Heat map and unsupervised clustering result for all glycoforms (FET, ASFE, ASGF, and ASFA) and interference conditions (FET + Gal, ASFE + Lac, ASGF + Lac, ASGF + GalNAc, and ASGF + GlcNAc) produced with data from seven lectins. (B) Principal component analysis (PrCA) score plot from microarray data of seven lectins which distinguished the nine test samples.

With the lectin microarray approach, approximately half of the 43 printed lectins demonstrated some level of binding to one or more of the bovine fetuin glycoforms profiled. This work showed that increasing the

stringency of data pre-filtering generally increased direct correlation of lectin microarray binding to known glycan structures from Fet and the various fetuin glycoforms. Furthermore, these findings also demonstrated the ability of the lectin microarray to discriminate between glycoforms under conditions in which a portion of the lectin population was competitively inhibited (Figure 4).

1.3.4 WP5 Lectin-glycan interaction analysis

A deep knowledge of the fundamentals underlying the interaction between glycans and their recognition partners, e.g. lectins, antibodies or enzymes, is essential to rationalize binding and selectivity data at atomic/molecular level and to approach if necessary a structural optimisation of either binding partner. Within WP5, we have aimed to examine the interaction between selected lectin mimics and glycan ligands, using both experimental interaction analysis and *in silico* modelling. The output of these studies may assist in the selection of an optimal panel of both protein and oligonucleotide-based lectin mimics for array platforms. In particular, we have focused on carbohydrate-binding proteins showing specificity towards galactosides, which play a prominent role as tumour-associated antigens. Thus, the study has included different galactose-specific antibodies and lectins exhibiting fine differences in oligosaccharide-binding that could be exploited in the development of sensors for glyco-phenotyping.

The study of the interaction of galactosyl-containing glycans with immunoglobulin fragments has involved the joint efforts of the groups at Madrid and Galway (Partners 10 and 1, respectively). In detail, we have performed a meticulous analysis of the recognition features of the non-human Gal-α-(1→3)-Gal disaccharide epitope (or glycotope), directly responsible for the hyperacute rejection (HAR) of porcine organ xenografts, by specific scFv G12, A4 and A11 antibody fragments (prepared by partner 1). Saturation transfer difference (STD) nuclear magnetic resonance (NMR) spectroscopy methods have been employed to access residue-specific binding information from the point of view of the ligand, while circular dichroism (CD) measurements have been used to monitor conformational and thermal stability changes in the protein due to the presence of the ligands. The STD analysis has been also applied to investigate the interaction with the scFvs of three other related oligosaccharides, i.e., the linear-B-trisaccharide, 3α-4β-D-Galactotriose and 3α-4β-3α-D-Galactotetraose in order to examine the involvement of each residue in the recognition process. In all cases, the presence of the 3-O-α-D-Gal A' moiety is a decisive feature for the binding process to take place. In contrast, the other residues at the reducing end merely display transient interactions with the carbohydrate-binding site.

Regarding lectins, the studies have covered a large panel of targets, including human galectins 1-4 and 7-9, the human macrophage Gal-specific C-type lectin, also a plant toxic lectin (viscumin) and other model mannose-specific lectins (Concanavalin A, pea lectin), all prepared by **Partner 3**.

Diverse NMR methodologies have been employed to analyse the structures (NOESY, TOCSY) and to answer questions concerning potential aggregation (DOSY) and folding (temperature-dependent experiments) states, presence of conformational flexibility or defined rigid structures in solution (relaxation time measurements), localisation of the binding sites (HSQC) and stoichiometry (DOSY, chemical shift titration). From the ligand perspective, the bound conformation and binding epitope have been determined using TR-NOE and STD experiments, respectively. Different glycomimetics have been exploited, from galactose derivatives to others mimicking glucose or mannose, including fluorine-containing molecules. In parallel, information on the structural organization, hydrodynamic properties and thermal stability of the proteins has been obtained using CD, size exclusion chromatography, sedimentation equilibrium and sedimentation velocity experiments, while the thermodynamic parameters of ligand binding and crosslinking activity have been assessed by isothermal titration calorimetry (ITC) and asialofetuin precipitation experiments.

Using a combination of docking and molecular modelling procedures, structural models for different lectin-ligand complexes have also been established. The lectin-ligand interaction involves hydrogen bonds, van der Waals contacts and carbohydrate-aromatic stacking interactions. In particular, from the lectin side, strategically positioned aromatic residues, especially Tyr and Trp, as well as key histidines are involved in the interaction. Additionally, glutamine and arginine residues also play a paramount role in the stability of the

complexes. From the sugar and mimetics perspective, non-polar patches with at least three C–H bonds are involved. Moreover, at least two key positions should be hydroxylated since they are essential to establish hydrogen bonds.

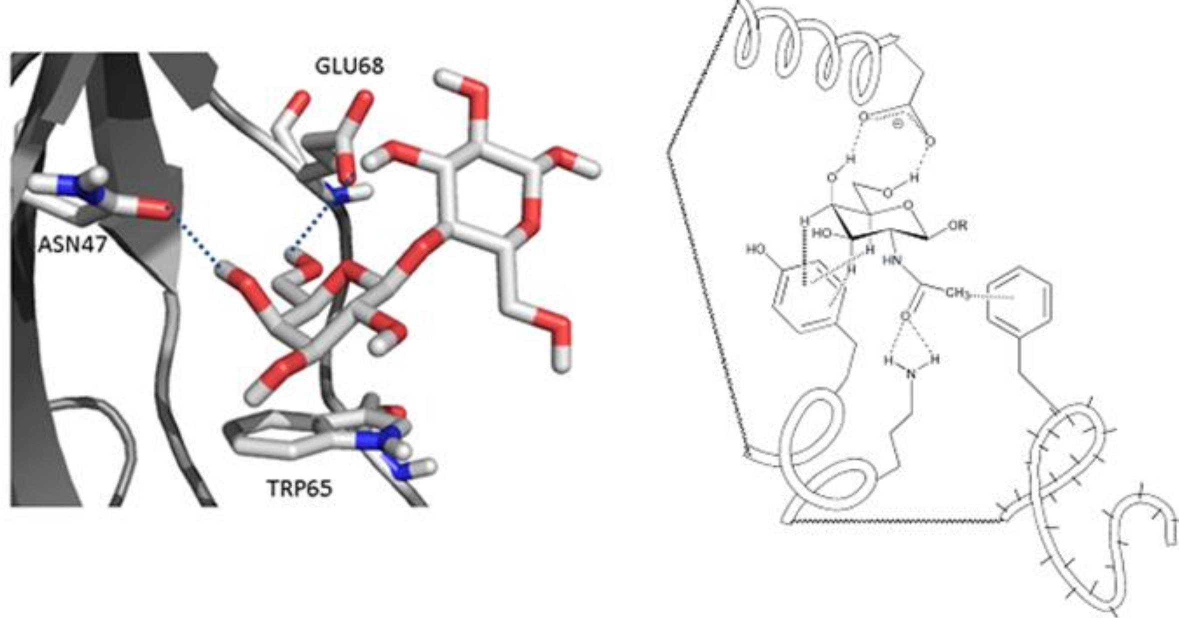


Figure 5 Right panel, schematic view of the most frequent interactions in the recognition of carbohydrates by proteins and by their corresponding mimetics. Left panel, the specific example of the recognition of galactose derivatives and glycomimetics by the galectin family.

1.3.5 WP6 Glycoanalysis bioinformatics

In summary, the work of WP6 has included the following:

- WP6 has focused on supporting the bioinformatics needs of all other WPs. This includes the development of data analysis methods to address the following types of data:
    - Lectin arrays – adapting standard statistical methods to compare binding signals between different samples. The methods are described in NUIG publications.
    - HPLC glycomics - adapting survival analysis and comparative statistical techniques to study the potential signal of serum glycobiomarkers in terms of differentiating tumor types and survival classes.
    - LC/MS glycomics - we developed methods for identifying peaks and enabling a subtractive approach to glycoproteomics, (Agilent in collaboration with KI and NIBRT). The approach is based on comparing pre PNGaseF to post PNGaseF measurements of the same sample, in an automatic manner. IEF fractionation increases the potential coverage and accuracy of the technique. Successful measurement results were obtained for several nid complexity samples, including human MARS14.
    - Study of p16 models, (Agilent jointly with KI and LMUM).
    - Joint analysis of glycomics data with other types of measurements on the same cohort - Agilent developed methods and software to support joint data analysis and encapsulated them, in collaboration with other Agilent groups, into a Cytoscape plug-in, called ENViz, which was released in late 2013. Agilent used their methods to analyse the cohorts provided by OUS and are writing a manuscript to describe the results.

- Agilent also developed algorithmics to optimize fraction selection in IEF. The algorithm was tested on actual measurement data from KI as well as from BPRI (Beijing Protein Research Institute, part of an associated Chinese project). A paper was accepted to BIBM 2014, as a full paper. A component is currently incorporated into SpectrumMill, a commercial proteomics and glyco-proteomics analysis s/w.
- Lectin mimics and optimized lectins - Agilent worked, in collaboration with NUIG, on optimizing naturally occurring mammalian lectins to yield better specificity in terms of differentiating Nu5Ac from Nu5Gc forms of syalic acid. This work required the development of novel applications of libraries of synthetic oligonucleotides, one of Agilent's leading products, and of NGS data analysis techniques, as described in more details in D6.3.
- In WP6, IP and KI have developed 2D and 3D alignment methods to improve the comparison of two or more LC/MS proteomics or glycoproteomics samples.

1.3.6 WP7 Label-free sensor development

WP7 related to the development and validation of label-free sensor platforms for protein-based lectin array assays.

Carbohydrate suspension array

We evaluated several iterations of lectin and glycan suspension arrays. A panel of lectins were then coupled with beads and incubated with FITC-labelled glycoproteins. As only some of these lectins functioned correctly in this format, use of the lectin suspension array was not pursued. Similarly several anti-glycan antibodies, including anti-H2, anti-Le^{b}, anti-H1 and anti-Le^{a} antibodies, were also applied to the suspension array. The coupling efficacy was validated by anti-mouse antibodies and showed good efficiency. However, the antibody suspension array did not show any binding to appropriate carbohydrate-PAA-biotin conjugates or carbohydrate-biotin with streptavidin-FITC. The reasons may include insufficient antibody concentration to saturate the beads or a requirement for antibody presentation at some distance from the bead surface e.g. use of linkers. However, the relative cost was deemed as too great for antibodies as the requirement for antibody concentration was so high and the antibody suspension array was discontinued due to prohibitive cost.

Several different iterations of carbohydrate suspension array were assessed, including immobilisation of streptavidin beads coated with biotinylated carbohydrates, which resulted in low detection intensities. The strategy was altered to increase the distance of the presented carbohydrates from the bead surface and using biotin-PAA-sugar. Anti-glycan antibodies successfully detected specific glycans in this iteration and greater detection sensitivity was afforded compared to biotin-carbohydrate. This was most likely due to the multivalency generated using the BSA-biotin conjugated beads system.

The carbohydrate suspension array was applied to detect the blood type of the 25 serum samples from the MDG cohort transferred from **Partner 7** (OUS) to **Partner 1** (NUI Galway) (as described in WP8 below). The tetrasaccharides of blood type A and B were immobilised on beads and incubated with the serum IgG fractions. IgG binding was detected using goat anti-human secondary antibodies labelled with rhodamine. However, only a few samples had a response with blood type A beads. The high distribution of blood group AB amongst the sample set seems statistically unlikely. The experiment was repeated again with a different secondary antibody labelled with FITC but the results were similar. One possible reason is that the tetrasaccharide of blood type A did not present a good recognition molecule and the trisaccharide may function better in this format. In addition to blood groups A and B, the sialyl Tn and Le^{b} structures were included in the carbohydrate suspension array. The preliminary data also showed some interactions between the cancer antigen and serum antibodies. In summary, the carbohydrate suspension array was successfully established and application to samples will continue beyond the life of the GlycoHIT project.

EIS sensor

A linear relationship between ΔR_{ct} signals and the concentrations of cells was obtained in the range from 1 × 10^4 cell mL^{-1} to 10^7 cell mL^{-1} using K562 cell as a model as shown in Figure 6. The detection limit of K562 cell was 8.1×10^3 cell mL^{-1} (signal-to-noise ratio of 3). The linear relationship can be represented as ΔR_{ct} (Ω)=51.7×Log C_{cell} (cells·mL^{-1}) – 182.1 with a correlation coefficient of R^2 = 0.9900, where ΔR_{ct} is the electron transfer resistance change and c is the K562 cell concentration.

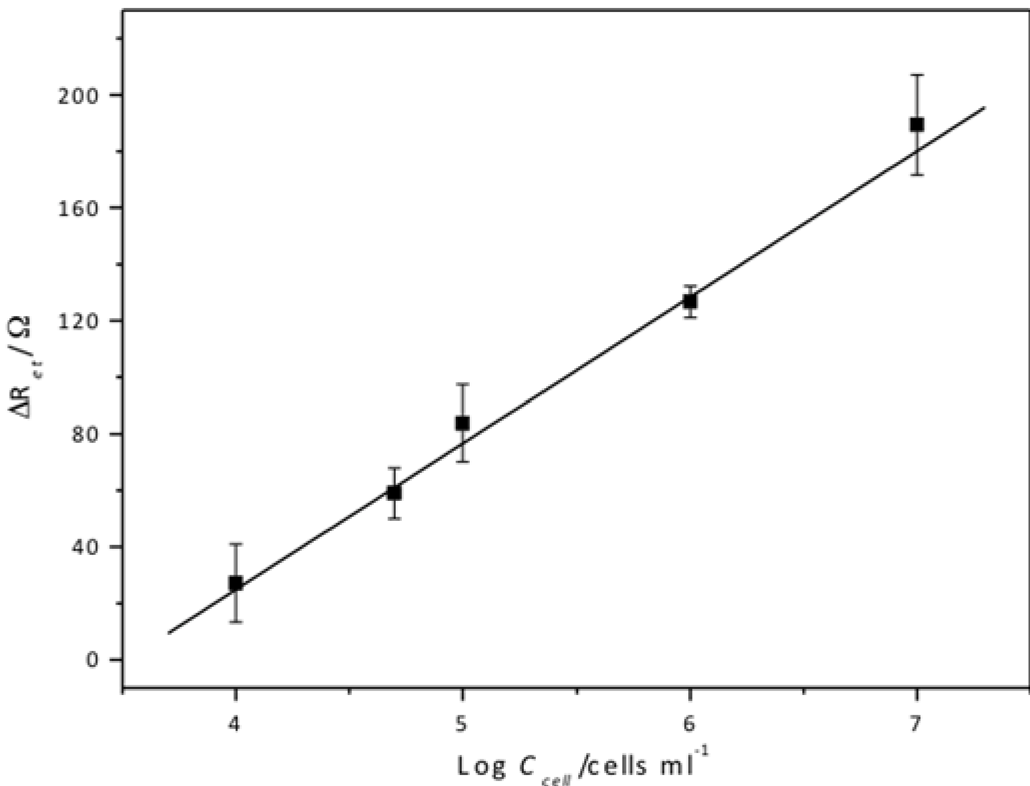


Figure 6 Plots of ΔR_{ct} (Ω) (ΔR_{ct} (Ω) = R_{ct} (GCE\CNT\ConA\Cell) - R_{ct} (GCE\CNT\ConA)) value vs. logarithm value of different K562 cell concentrations. (a) 1.0×10^4, (b) 5.0×10^4, (c) 1.0×10^5, (d) 1.0×10^6 and (e) 1.0×10^7 cells·mL^{-1}. Electrolyte solution: 5 mM K4Fe(CN)$_6$, 5 mM $K_3Fe(CN)_6$ and 0.5 M KCl. Scan rate: 100 mV/s. Frequency range was 0.1~10^5 Hz. Amplitude: 10 mV. The error bars were obtained from three parallel experiments.

ECL competitive-type biosensor

The prepared gold nanoparticles were characterized by transmission electron microscopy (TEM) and UV-vis spectra. Statistical analysis of TEM data revealed that the average diameter of the gold nanoparticles was about 20 nm. The prepared GOx@Au nanoprobes were also characterized by TEM and UV-vis. The probes were about 20 nm from TEM image. From the UV-vis absorption spectrum, the concentration of Au nanoparticles was also estimated to be 1.87 nM. After the conjugation of GOx, the characteristic absorption peak of gold nanoparticles red shift from 520 nm to 525 nm due to the decoration of GOx on Au nanoparticles.

In order to deeply understand the interaction between ConA and mannose, a quantitative analytical technique, based on ECL and EIS techniques, was developed to characterize the affinities between ConA and mannose, mannan, GOx, or GOx-Au NPs. The dissociation constants (K_d) are 0.44 mM, 52.3 nM , 0.30 µM, and 1.64 nM, respectively. Such a carbohydrate biosensor shows excellent sensitivity for the detection of

K562 cells. Figure 7 shows the ECL signal intensity as a function of cell concentration. A detection limit of 18 cells with a linear calibration range from 4.2×10^3 to 3.0×10^6 cells mL-1 was obtained. This strategy presents a promising platform to design biosensors by taking advantage of the interaction between carbohydrates and lectins.

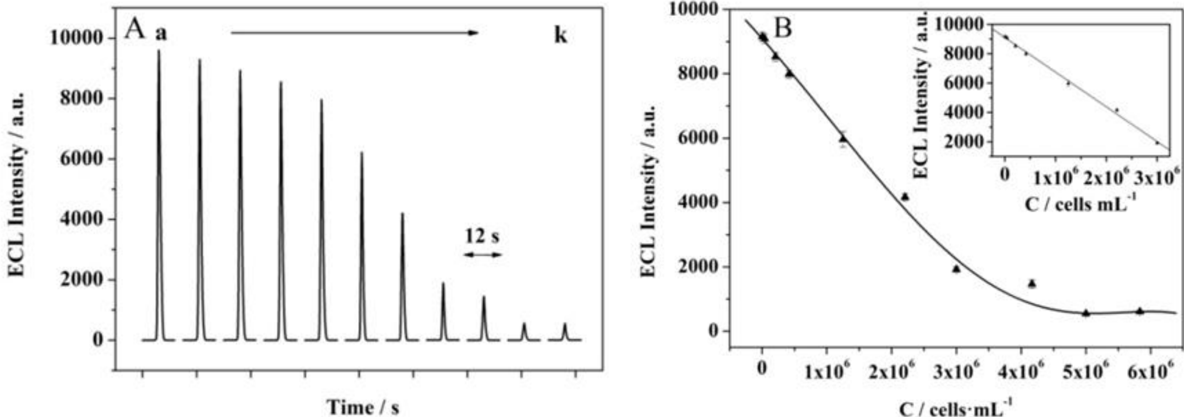


Figure 7 (A) ECL-Time behaviours of GOx@Au/BSA/ConA/GE at different concentrations of K562 cell: (a) 0, (b) 4.2×10^3, (c) 4.2×10^4, (d) 2.1×10^5, (e) 4.2×10^5, (f) 1.3×10^6, (g) 2.2×10^6, (h) 3.0×10^6, (i) 4.2×10^6, (j) 5.0×10^6, (k) 5.8×10^6 cells mL^{-1}. (B) ECL intensity as a function of the concentration of K562 cells. The inset shows the linear relationship between ECL intensity and the K562 cell concentration.

Taking into account that 10 µL of K562 cell suspension was used for incubation, the competitive ECL strategy achieved the limit of detection of only 18 K562 cells. Assuming the cell had the same binding kinetics as the mannose containing glycoproteins, the cell surface mannose quantities could be equal to the active mannose amount entirely replacing the probes on the electrode. In this way, the mannose moieties on K562 cell surface were estimated to be 1.8×10^{10}, which was close to that obtained by electrochemical and enzymatic methods. These results demonstrate that the proposed ECL strategy is suitable for sensitive cell surface carbohydrate expression evaluation. Such a simple and sensitive biosensor is promising in the applications of in situ carbohydrate expression and tumor cell determination.

ECL sandwich-type sensor

The nanoprobes were then used in the ECL sandwich-based assay for determination of the cell surface carbohydrate expression. The Cell/Con A/MWNT/GCE electrode described above was incubated with 10 µL of the ECL nanoprobe ConA@Au-RuSiO_2 NPs solution for 60 min at 37 °C. Finally, the electrode was washed carefully and thoroughly with PBS to remove non-specifically bound nanoprobes to minimize the background response. The resulting electrode was characterized via ECL measurement in a 0.1 M PBS solution (pH 7.4) with 5 mM TPA using Ag/AgCl electrode with saturated KCl solution and platinum wire as the reference electrode and counter electrode, respectively. The ECL measurements were performed from 0.2 to 1.25 V with scan rate of 100 mV s^{-1}. The EIS spectra for the assembly procedure of the electrode and carbohydrate expression were determined. The increased charge transfer resistance indicates that Con A, cell and aptamer are assembled on the electrode due to the electronic inert properties of the biomolecules. It was observed that the intensity of light emission was significantly increased in the presence of cells. Such results indicate that the biosensor can be used for both cell assay and cell carbohydrate expression evaluation.

The surface density of Con A immobilized on the MWNTs modified electrode is an important parameter for capturing cell. As shown in Figure 8, a sharp increase of the ECL signal of the biosensor is observed with the increasing concentration of Con A in the solution while modifying the electrode. Then, it reached a steady value after 2.0 mg mL^{-1}, indicating a tendency of saturated surface density to thoroughly capture cells. In addition, another significant factor for capturing cells on the Con A/MWNT coated electrode is the cell incubation time. With the increasing incubation time with 1.0×10^6 K562cells mL^{-1}, the ECL intensity of the

biosensor increased and then reached a plateau at 80 min (Figure 8). Longer incubation time did not enhance the response further.

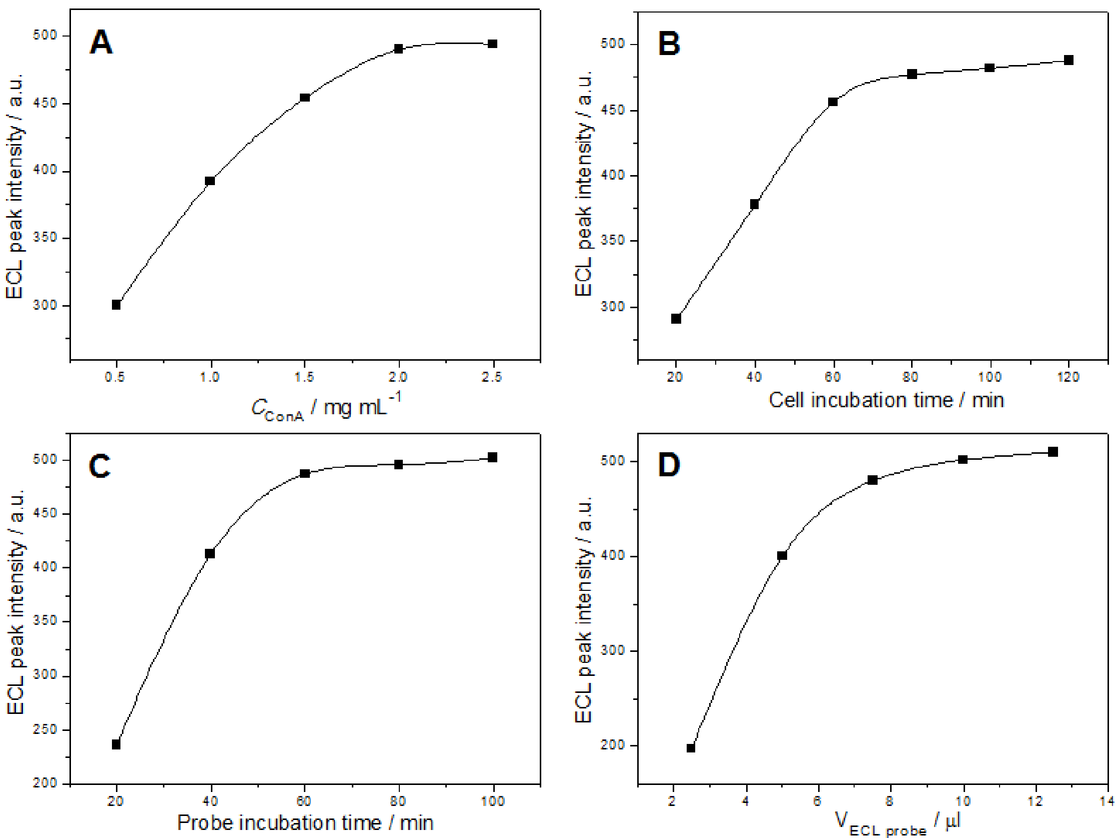


Figure 8 Optimization of (A) Con A concentrations modified on GCE, (B) Cell incubation time, (C) Probe incubation time and (d) Probe volumes of the ECL biosensor. The ECL peak intensity for analysis was obtained at 1.25 V (vs. Ag/AgCl) in 0.1 M PBS (pH 7.4) containing 5 mM TPA. CV Scan rate: 100 mV/s. The PMT voltage: 750 V. The concentration of K562 cell: 1.0×10^{6} cells·mL^{-1}

Aptamer electrogenerated chemiluminescence (ECL) biosensor

The highly sensitive and selective detection of cancer cells from human serum plays an important role in early cancer diagnosis, and thus greatly increases the chance for effective treatment. Under the optimized conditions, we investigated the performance of the developed aptamer sensor in quantitative analysis of cancer cells. Figure 9 presents the ECL signals of the biosensor responding to CCRF-CEM cells. Upon increasing the concentration of CCRF-CEM cell, the ECL intensity decreased due to the ECL inhibition of ALP hydrolysate phenol. The ECL intensity changes between the ECL signal without cells and with cells at different concentrations was recorded. The detection limit is estimated to be 38 cells mL^{-1} at 3σ. The ECL biosensor exhibits higher sensitivity than some earlier reported methods. It is reasonable that the ECL cytosensor is based on the covalent attachment of aptamer to the rGO-DEN. The large number of aptamers mediated by PAMAM affords a multivalent binding effect that significantly enhances the cancer cell capture efficiency.

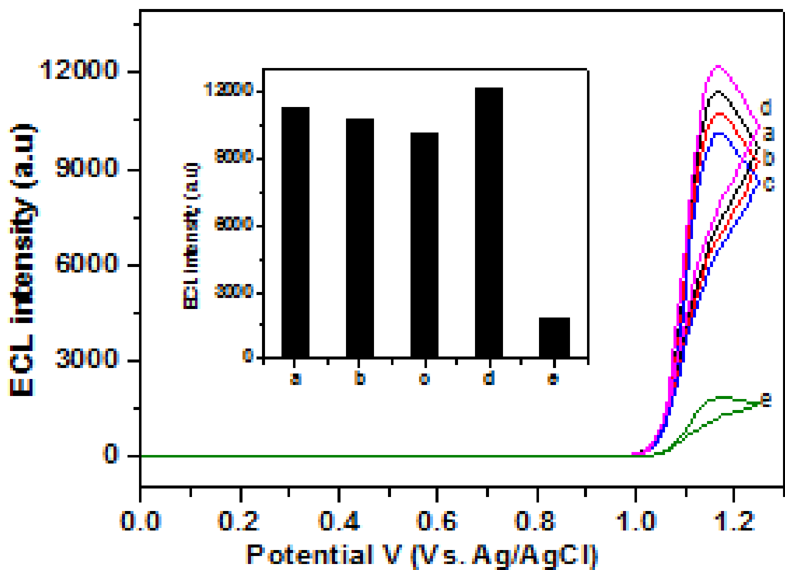
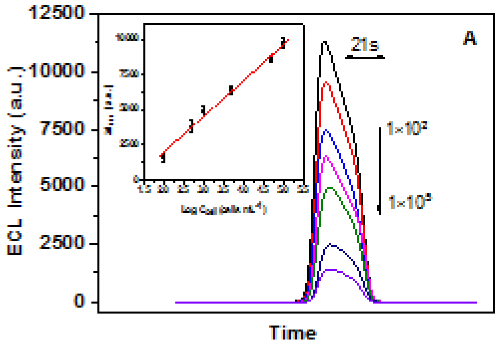


Figure 9 ECL intensity-potential curves of aptamer/rGO-DEN/GCE (a), cell/aptamer/rGO-DEN/GCE (b), ALP-Con A-Au NPs/aptamer/rGO-DEN/GCE (c), ALP-Con A-Au NPs/cell/rGO-DEN/GCE (d), and ALP-Con A-Au NPs/CCRF-CEM cell/aptamer/rGO-DEN/GCE (e) in 50 mM Tris/HCl (pH 9.0) containing 0.25 mM TPA, 75 uM $[Ru(bpy)_3]^{2+}$ and 4mM PPNa. Scan rate, 100 mV/s. The PMT voltage is 600 V. The concentration of CCRF-CEM cells: 1.0×10^5 cells mL^{-1}. Inset: Values of the peak intensities of the corresponding ECL curves at 1.25V.



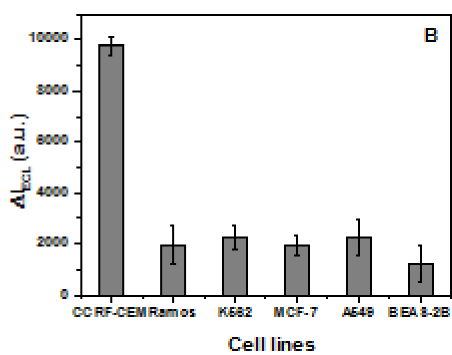


Figure 10 (A) ECL intensity-time curves of the aptasensor after incubation with CCRF-CEM cell at different concentrations (from 0, 1.0×10^2, 5.0×10^2, 1.0×10^3, 5.0×10^3, 5.0×10^4, to 1.0×10^5 cells mL^{-1} respectively) in 50 mM Tris/HCl (pH 9.0) containing 0.25 mM TPA, 75 uM $[Ru(bpy)_3]^{2+}$ and 4 mM PPNa. Inserts are the plots of ECL intensity changes vs logarithm of CCRF-CEM cell concentration. Error bars represent standard deviations from three independent experiments. (B) The ECL changes of the aptasensor treated with various cancer cell lines. All measurements were performed with as cell concentration of 1.0×10^5 cells mL^{-1}. Scan rate was 100 mV/s. The PMT voltage was 600 V.

Electrochemical aptamer biosensor

We demonstrate a multivalent recognition and highly selective aptamer signal amplification strategy for electrochemical cytosensing and dynamic cell surface N-glycan expression evaluation by the combination of Con A, a mannose binding protein as model, conjugated poly(amidoamine) (PAMAM) dendrimer on chemically reduced graphene oxide (rGO-DEN) interface, and aptamer and horseradish peroxidase (HRP) modified gold nanoparticles (HRP-aptamer-AuNPs) as nanoprobes. In this strategy, the rGO-DEN cannot only

enhance the electron transfer ability but also provide a multivalent recognition interface for the conjugation of Con A that avoids the weak carbohydrate–protein interaction and dramatically improves the cell capture efficiency and the sensitivity of the biosensor for cell surface glycan. The high affinity aptamer and HRP modified gold nanoparticles provide an ultrasensitive electrochemical probe with excellent specificity. As proof-of-concept, the detection of CCRF-CEM cell (human Acute Lymphoblastic Leukemia) and its surface N-glycan was developed. It has demonstrated that the as designed biosensor can be used for highly sensitive and selective cell detection and dynamic evaluation of cell surface N-glycan expression. The detection limit of as low as 10 cell mL^{-1} was obtained with excellent selectivity. Moreover, this strategy was also successfully applied for the N-glycan expression inhibitor screening. These results imply that this biosensor has potential for clinical diagnostic and drug screening applications, and endows a feasibility tool for an insight into the N-glycan function in biological processes and related diseases.

1.3.7 WP8 Glycomics in cancer, a pilot study

This work package aimed to validate HTP technologies in a clinically relevant setting and was composed of a number of different work-streams.

In WP8, clinical material from our breast cancer cohorts was analysed by UPLC, plant and mammalian lectin microarrays and HiRIEF LC-MS/MS. The technologies developed in the consortium have shown sensitivity and specificity sufficient to reveal clinically significant biological alterations in glycans and glycan-related structures.

We analysed breast cancer serum samples from Partner 7 (OUS) and CHO cell and secreted samples from Partner 14 (BMS). The cell glycans contained mostly high mannosylated glycans with some complex glycans also present; sialic acid was mostly alpha-(2,3)-linked, galactose mostly beta-(1,4)-linked and they contained core fucose. There was no significant difference among the three processes in cell glycans. Comparing different time-points, a significant decrease of high mannosylation, sialylation and branching with time was observed. The secreted glycans were mostly complex with sialic acid mostly alpha-(2,3)- linked, galactose mostly beta-(1,4)-linked, with core fucose mostly present. There was no significant difference among the three processes in supernatants. Comparing different time-points in secreted glycans, significant changes were observed indicating a decrease in sialylation and branching with time.

Within WP8, using prior MS and MS/MS analysis, we profiled clinical breast samples and successfully achieved a high number of protein identifications and quantifications (cohort of 8 luminal A and 8 basal like breast cancers from Partner 7 (OUS). By coupling isotopic labeling (iTRAQ-Isobaric tags for relative and absolute quantitation) to this peptide fractionation enabled quantitation of proteins in breast cancer tissue. A key finding is the acquisition of information rich data on protein level that allows evaluation of the oncogenic pathway status and to relate the data to mRNA and glycan profiling. The analysed quantitative ration distribution between individual tumour samples and even number of protein identified on the two analysed sets (set A and Set B), indicated robust clinical sample selection, preparation and technically successful experiment. This proteome analysis set is so far the most comprehensive analysis of breast cancer proteome and using the data we can cluster the samples based on protein profile.

Further, we used the method to analyse endocrine therapy resistant breast cancer cell line model and clinical samples. We discovered the retinoic acid receptor alpha (RARA) as a potential marker for Tamoxifen resistance in breast cancer (Johansson, *et al.*, Nature Commun., 2013), and this finding was validated in three independent small cohorts using three techniques (ELISA, IHC and WB). We also demonstrated that another estrogen receptor targeting drug, Fulvestrant, is effective for Tamoxifen resistant cells with high RARA expression. The findings could be used to select patients for Fulvestrant treatment. This new proteomics method has opened up an interesting window for detection of novel protein coding regions that will be used in the coming projects.

Studies were carried out by Partner 9 (UMONS) in WP8 related to the detection of galectins and galectin-reactive glycoproteins (the putative biomarker) by immunohistochemistry (IHC) in head and neck paraffin embedded tumour sections. To validate the IHC technique with galectin antibodies as HTP technologies for

clinical sample analysis, we established different series of naso-sinusal diseases, benign and malignant thyroid lesions and oral cavity, pharyngeal and laryngeal cancer, in collaboration with different hospitals. Semi-quantitative analysis or computer-assisted microscopy were conducted on the immunostaining to establish the MOD (Mean Optical Density), LI (Labelling Index) and QS (Quick Score). Data was analysed with the non-parametric Kruskall-Wallis test (more than two groups). In case of significant results, post-hoc tests (Dunn procedure) were used to compare pairs of groups (to avoid multiple comparison effects).

In addition to the above, purified biopharmaceutical samples from various stages of the production were analysed using the novel carbohydrate recognition molecules developed in WP3 and revealed the presence of Gal-α-(1→3)-Gal and the absence of Neu5Gc on all samples.

Breast cancer patients compared to controls

From our analysis of the serum *N*-glycome we have identified significant differences in breast cancer patients compared with controls. We confirmed previous breast cancer findings, such as increases in sialylation, branching and outer arm-fucosylation. We also found decreases in high mannosylated and biantennary core-fucosylated glycans in breast cancer patients compared to controls (Figure 11). We found bisected biantennary non-fucosylated glycans were decreased in patients with progesterone receptor positive tumours and core-fucosylated biantennary bisected mono-galactosylated glycans were decreased in patients with tumour TP53 mutation (Figure 11).

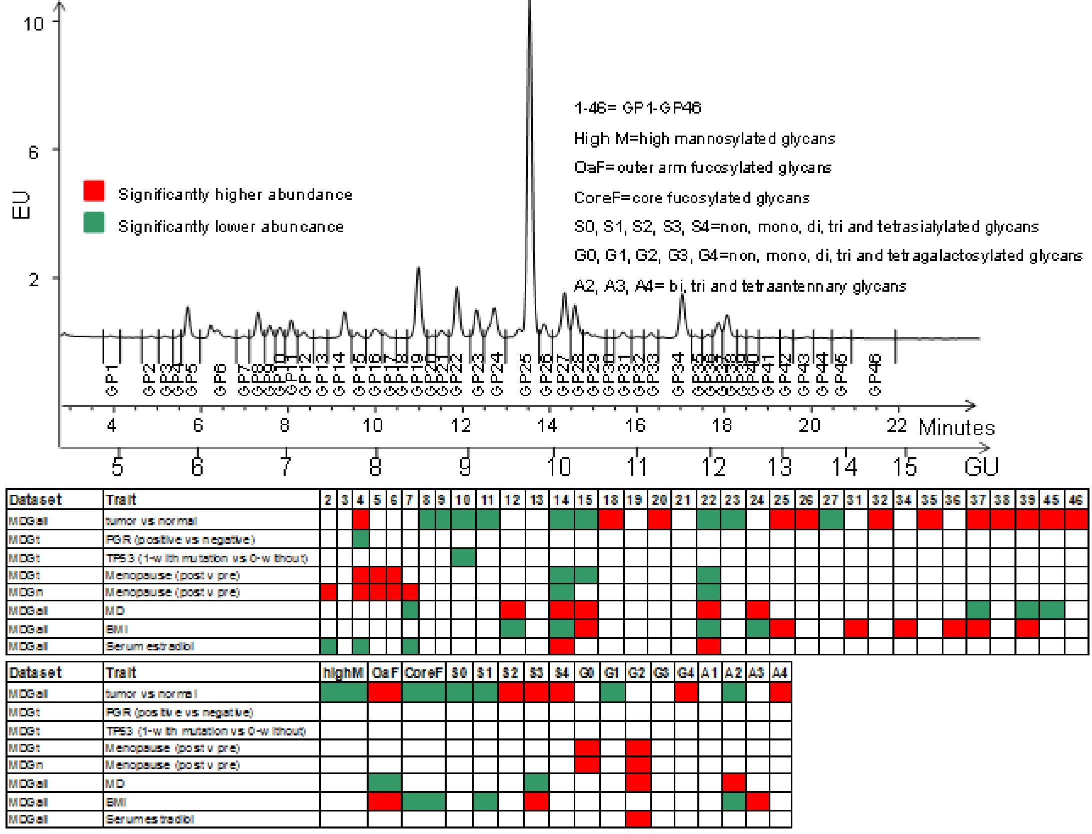


Figure 11 Example of breast cancer sample HILIC-UPLC chromatogram and separation into 46 peaks and significant differences between patients and control groups of interest.

Serum N-glycans related to systemic features in healthy women

Serum N-glycome alterations related to mammographic density, body mass index (BMI) and serum estradiol were identified (Figure 11), of which several had not previously been identified. These are all risk factors for breast cancer.

Identification of alterations specific to breast cancer subtypes

In our published study (Saldova, *et al*., JPR, 2014), we found that systemic features, such as the presence of a carcinoma, serum estradiol level (with related features such as mammographic density) and body mass index were associated with greater alterations in the serum *N*-glycome than features specific to the tumour. Serum estradiol was associated with an increase in di-galactosylated glycans, and higher mammographic density was associated with an increase in biantennary d-igalactosylated glycans and a decrease in tri-sialylated and outer-arm fucosylated glycans (Figure 11). Higher BMI was associated with increase in sialylation, branching, and outer-arm fucosylation and with decrease in core-fucosylation (Figure 11). Nonetheless, we did identify associations between tumour characteristics and serum N-glycans in both the MDG-cohort (Figure 11) and the MicMa-cohort (Figure 12).

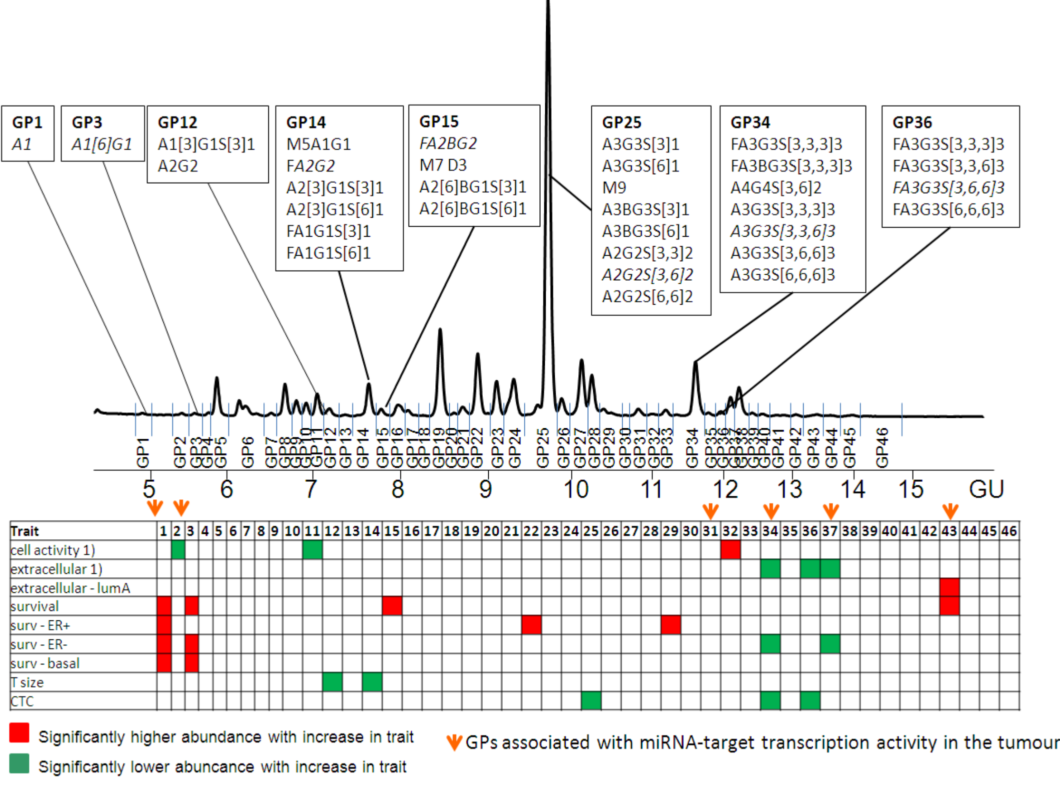


Figure 12 Overview of glycan peaks significantly associated with mRNA-expression, miRNA-targets (orange arrow) and clinical parameters such as survival, tumour size (T size), estrogen receptor status (ER) and circulating tumour cells (CTC) in the MicMa-cohort.

Breast cancer subtype-specific alterations were also identified by high-throughput protein analysis of the tumour proteome and a total of 144 proteins differentially expressed between the good prognosis luminal A subtype and the poor prognosis basal-like subtype (Figure 13).

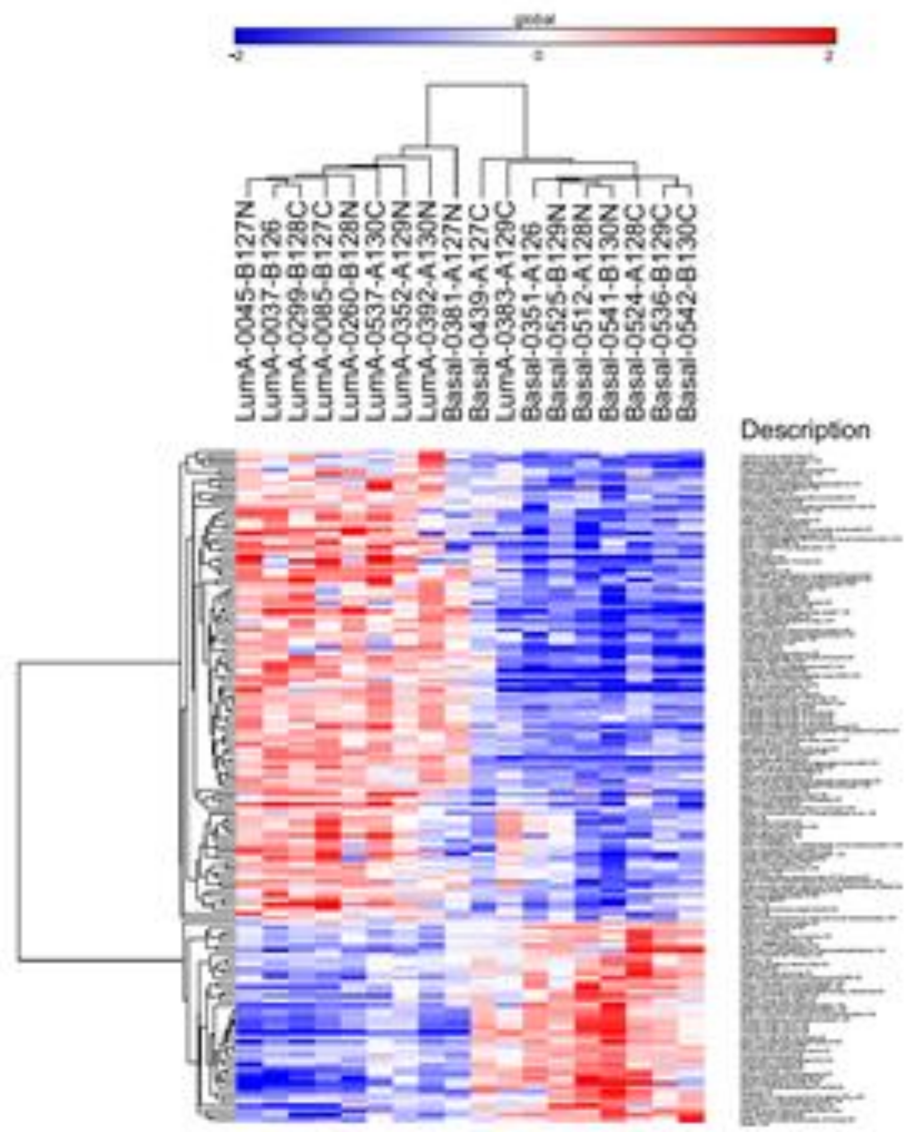


Figure 13 Clustering of the patient samples based on the 144 proteins significantly altered between the luminal A and basal-like breast cancers.

Serum N-glycans associated with tumour gene expression

Using UPLC-analysis of the serum N-glycome of the MicMa-cohort we have demonstrated how integrating gene expression from breast carcinomas may be used to identify serum glycans related to breast carcinogenesis and functional processes in the tumour (Figure 14). This approach may improve the search for biologically relevant serum markers of malignant disease. The strongest associations between tumour gene expression and serum glycan structures were seen for tri-sialylated tri-galactosylated triantennary glycans with or without fucose (GP34, GP36 and GP37) which are correlated with lower adhesion. Reduced adhesion may facilitate invasion and migration of the cancer cells and we do see an insignificant correlation between higher serum levels of GP34 and GP37 and poor prognosis in patients with ER negative tumours (p=0.03 and 0.02 respectively).

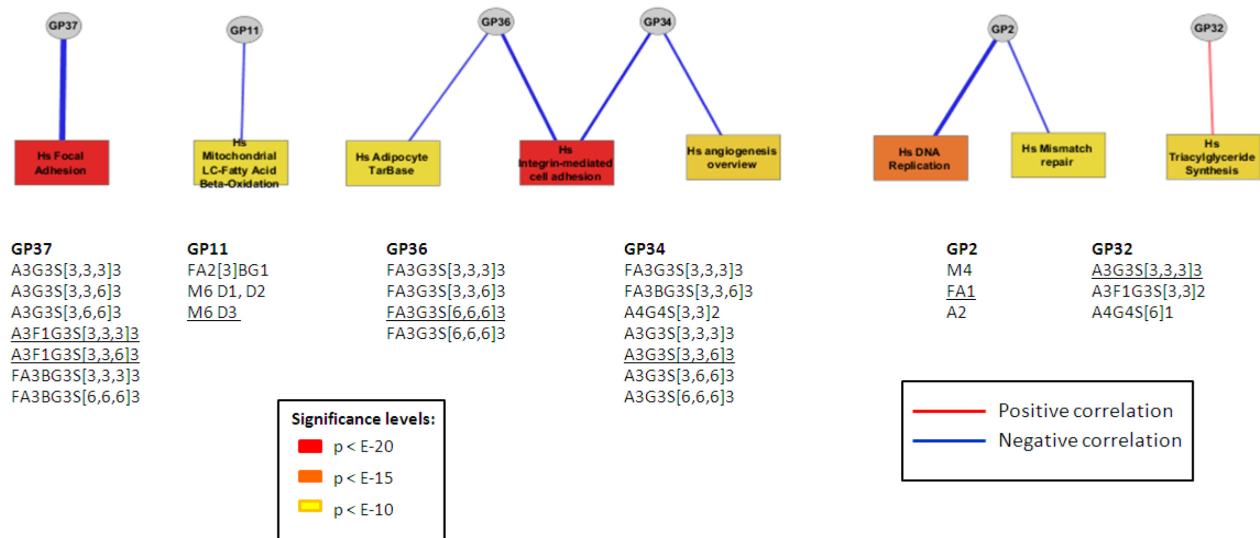


Figure 14 GP-pathways enrichment network. The network is composed of glycan peaks (GP) indicated by oval nodes and wikiPathways indicated by rectangle nodes. An edge in the network represents a significant association between the GP and the connected pathway. Red edge indicates a significant correlation between the GP and the pathway genes, and blue edges indicates a significant anti-correlation between the GP and the pathway genes, FDR<10%. The width of an edge is correlated to the –log(enrichment p-value) and each pathway node is coloured according to the sum of –log(enrichment p-value) of edges connected to it. Glycan structures contributing to each peak is listed with the structure contributing most to the peak in underscore.

Detection of galectins and galectin-reactive glycoproteins by IHC in Head and Neck tumours.

Partner 9 carried out three studies:

(i) Galectin fingerprinting in naso-sinusal diseases (Duray A., et al., 2014)

Galectins immunohistochemical expression profiles in inverted papilloma (IP): In IPs, nucleocytoplasmic immunostaining of Gal-1, -3 and -7 was detected in 100% of epithelial cells. Slightly lower percentages of 96 and 93% were detected for Gal-4 and both Gal-8 and -9, respectively. Galectins were present in both the nuclei and cytoplasm. Only Gal-9 was invariably present in epithelial cells. For epithelial cells, a low to moderate signal was detected for Gal-1, -4 and -8, and a moderate to intense signal was detected for Gal-3, -7 and -9. An analysis of the quantitative data showed an increased percentage of positivity for the following Gals in the epithelium of IPs: Gal-3 compared to allergic nasal polyps (ANPs) and non-eosinophilic chronic rhinosinusitis (NECRS) (post-hoc comparison, P=0.0003 and P=0.002, respectively), Gal-4 compared to allergic and non-allergic nasal polyps (post-hoc comparison, P=0.0004 and P=0.0007, respectively), Gal-7 compared to allergic nasal polyps and eosinophilic chronic rhinosinusitis (ECRS) (post-hoc comparison, P=0.000002 and P=0.0003, respectively), Gal-8 compared to squamous cell carcinoma (post-hoc comparison, P=0.004) and Gal-9 compared to ANPs, ECRS and NECRS (post-hoc comparison, P=0.01, P=0.0006 and P=0.02, respectively). IPs are thus characterized by the overexpression of Gal-3, -4, -7, -8, -9 but not Gal-1.

Galectin immunohistochemical expression profiles in carcinoma***:*** The epithelial cells in naso-sinusal SCCs were positive for all of galectins tested, except Gal-8 (83%). The localization depends on the type of lectin. The nucleocytoplasmic or cytoplasmic expression pattern contrasted with the exclusively cytoplasmic positivity observed for Gal-1 and -9. The signal intensity detected was moderate for Gal-1 and low to moderate for the other Gals. A higher percentage of Gal-9-positive cells was detected in SCCs compared to ANPs, NECRS and ECRS (post-hoc comparisons, P=0.02, P=0.02 and P=0.003, respectively). In contrast, a lower Gal-8 quick score was determined for SCCs compared to IPs (post-hoc comparison, P=0.04). No significant difference was found for the presence of Gal-1, -3, -4 and -7. Thus, the immunohistochemical profile of nasal carcinomas appears to be characterized by a high-level of Gal-9 and a low-level of Gal-8.

(ii) Biotinylated-galectin (-1, -3, -7, -8) expression in benign and malignant thyroid lesions: statistical comparison

Semi-quantitative assessment of acceptor sites expression for galectin-1: The level of acceptor site expression for galectin-1 significantly differed comparing cytoplasmic compartment of adenomas and papillary carcinomas (LI : p= 0.002 ; Mann-Whitney test et MI : p= 0.03 ; Mann-Whitney test). In fact, we revealed the presence of acceptor sites for the galectin-1 in the cytoplasm of follicular cells in papillary carcinomas (LI : median= 3 ; MI : median= 1) whereas cytoplasmic staining was not found in adenomas (LI : median= 0 ; MI : median= 0).

Semi-quantitative assessment of acceptor sites expression for galectin-3: Concerning the results of the glycohistochemical analysis of galectin-3 in benign and malignant thyroid lesions we obtained a significant difference in term of both the percentage of mean immuno-positive cells and the mean staining intensity (LI : p= 0.03 ; Mann-Whitney test et MI : p=0.03 ; Mann-Whitney test). Nucleus of epithelial cells in adenomas expressed acceptor sites for galectin-3 (LI : median= 0,6 ; MI : median= 0,5) in contrast with papillary carcinomas where nuclear staining was not found (LI : median= 0 ; MI : median= 0).

Semi-quantitative assessment of acceptor sites expression for galectin-7: For the acceptor sites expression for galectin-7 in adenomas and papillary carcinomas, a significant difference was detected in the cytoplasmic compartment both in the percentage of mean positive cells and the intensity of cytoplasm staining (LI: p= 0.000012 ; Mann-Whitney test and MI : p= 0.000012 ; Mann-Whitney test). The acceptor sites are present in cytoplasmic compartment of follicular cells in papillary carcinomas (LI: median= 2,5 ; MI : médian = 1). The acceptor sites expression for galectin-7 was negative in adenomas (LI: median= 0 ; MI : median= 0).

Semi-quantitative assessment of acceptor sites expression for galectin-8: The expression of acceptor sites for galectin-8 was also compared between the adenomas and papillary carcinomas. The data revealed a significant difference between the groups, both in the percentage of positive nuclei and the nuclear staining intensity (LI: p= 0.005 ; Mann-Whitney test and MI : p= 0.007 ; Mann-Whitney test). We observed a nuclear staining in epithelial cells of adenomas (LI: median= 1 ; MI : median= 0,9) compared to papillary carcinomas where nuclear immunoreactivity was negative (LI : median= 0,2 ; MI : median = 0,1).

(iii) Galectin-2 expression in head and neck cancers

The LI of galectin-2 expression in peritumoral head and neck epithelium, low and high grade dysplasia and in carcinoma was observed. A significant difference was detected in the percentage of positive cells (LI) between normal epithelium and, high grade dysplasia and carcinoma. Concerning the LI, there is also a significant difference between low grade dysplasia and carcinoma. The figure of the quick score, which takes into account the mean optical density, showed a significant difference between normal epithelium and, high grade dysplasia and carcinoma. And the QS showed a significant difference between high grade dysplasia and, low grade dysplasia and carcinomas.

1.4 *Impact, dissemination and exploitation.*

1.4.1 Overview

Protein glycosylation is a post-translational phenomenon that is involved in most physiological and disease processes including cancer. Most of the known cancer-associated glycobiomarkers were discovered individually using liquid chromatography and mass spectroscopy. The complexity of the glycan modifications of proteins with resulting functional alterations combined with the limitations of the previous techniques employed have hampered the biological understanding of how glycan modifications contribute to carcinogenesis and treatment effect. We have shown that the improvements of the techniques developed in this consortium, can identify biologically relevant and important alterations. This will enable a better characterization of glycans contributing to carcinogenesis and treatment effect. The latter is important in order to identify mechanisms of resistance to treatment. The studies conducted within this consortium have been encouraging, and several new collaborative efforts have been initiated to carry on the work that has been commenced.

1.4.2 WP2 Advanced LC-MS glyco-analytical methodologies

We have generated UPLC technologies with improved HTP resolution which are relatively cheap, quantitative and rapid. This high throughput automated technology enables us to screen hundreds of samples from different collaborative projects both from academia and industry. This UPLC technology, is significantly (6x) faster than previous, it has better resolution and is 10x cheaper than other methods on the market. With this improved speed and crucially improved data quality, we can thus analyse more samples in less time which in turn means we obtain a more robust statistical output.

We have established a HTP glyco-proteomic workflow (collaboration with NIBRT, KI and Agilent). This robust, robotic, user friendly, quantitative glycan analysis platform technology for LC and /or LC/MS is coupled to a dedicated Bioinformatics – Glycobase 3 (recently commercialised by Waters in UNIFI 1.7). This is of significant interest to pharmaceutical companies as well as academic institutions (we have transferred technology through the GlycoHIT programme).

The research carried out in WP2 has provided several new methods to facilitate in-depth proteome analysis and further improved combined proteomics and glycomics analysis. This allows improved analysis of healthy and diseased tissues and other relevant biological material has potential to provide novel combined protein glycan biomarkers. Such a potential biomarker candidate was discovered for endocrine resistant breast, namely retinoic acid receptor (RARA) (WP8). Further, these methods can be used to improve our understanding of disease mechanisms, such as demonstrated by elucidation of p16 related effects on glycosylation and proteome (Amano et al., 2012). The main disease focus in this work has been cancer, where there are virtually no diagnostic markers for plasma based assays. Combined protein and glycan markers has potential to provide more specific and sensitive markers.

A major discovery, using the HiRIEF method developed within WP2, was identification of 99 novel human protein coding genome regions. These novel proteins can be proven to be disease related in a number of diseases or function as biomarkers.

The retinoic acid receptor (RARA) which was identified as a potential marker for endocrine resistant breast cancer (WP8), could also be a potential drug target in the high expressing breast tumors. RARA is a nuclear receptor, a class of proteins that have proven to be very valuable drug targets. Virtually all the receptors for which ligands have been identified are successful therapeutic targets, with either natural or synthetic forms of the ligands having been converted to marketed drug products. Fifteen of these drugs ranked among the top 200 in sales in 2009, with these 15 alone accounted for $27.5 billion sales (Humphreys A. Top 200: World’s best-selling medicines. Med Ad News. Vol. 16, no. 7,

July 2010). Collectively, these agents account for approximately 13% of all drugs approved in the United States and an appreciable portion of global pharmaceutical sales, being second largest family of drug targets.

1.4.3 WP3 Mammalian lectin and lectin mimic arrays

Recombinant mammalian lectins were produced or sourced and a mammalian lectin microarray was successfully generated. This novel microarray was used to glycoprofile serum from healthy and breast cancer patients. Bioinformatic analysis is on-going and a publication on this work is planned.

An optimised procedure was developed for application of phage display technology to the identification of glycan binders. Gal-α-(1→3)-Gal–binding scFvs were identified and characterized, and used for development of a convenient competitive assay for detection of the disaccharide motif free in solution or protein bound. These molecules were patented and the work on scFv development, characterization and the developed ELISA was published. Several other scFvs for glycan motifs were generated and characterization is on-going. An application for a patent on these molecules is in progress.

The scFvs were also arrayed on a microarray platform and the platform was found to be favourable for HTP applications for detection and quantification of defined carbohydrate moieties. It can be used for both clinical and biopharmaceutical samples and is suitable for expansion with additional carbohydrate-specific scFvs as they come on-stream. A manuscript for publication on this work is in preparation.

A panel of next generation lectin mimics were designed, a phage library was built, and a refined set of variants for Neu5Gc was selected by biopanning of the library. Validation experiments are planned and **Partners 1 and 2** are in active pursuit of funding for this work.

1.4.4 WP4 Oligonucleotide arrays and glycosignature generation

Both DNA and RNA SELEX procedures were used to select oligonucleotide which specifically bound three cancer-associated glycan motifs, H type 2 antigen (H2), the Lewis b determinant and the non-human sialic acid, Neu5Gc. Consensus sequences were obtained and specificity and binding affinity of some of these sequences were successfully demonstrated. We are in the process of applying for patent protection for these sequences and then intend to prepare a manuscript for publication.

Proof of concept experiments using an off-the-shelf oligonucleotide microarray highlighted several bottlenecks which must be addressed before these microarrays can be employed for glycoprotein glycosignaturing. One of the main issues was the lack of a robust aptamer assay.

In order to progress the overall goals of GlycoHIT, plant lectin microarrays were validated as glycosignaturing platforms and this work was published (*Gerlach, J.Q., Kilcoyne, M., Joshi, L. Microarray evaluation of the effects of lectin and glycoprotein orientation and data filtering on glycoform discrimination. Anal. Methods, 2014, 6, 440-449*). Accurate correlation to known and determination of unknown glycan structures of glycoproteins from observed binding with carbohydrate-binding molecules is one of the major challenges for affinity-based glycomic methods. With the lectin microarray approach, approximately half of the 43 printed lectins demonstrated some level of binding to one or more of the bovine fetuin glycoforms profiled. This work showed that increasing the stringency of data pre-filtering generally increased direct correlation of lectin microarray binding to known glycan structures from fetuin and the various fetuin glycoforms. Furthermore, these findings also demonstrated the ability of the lectin microarray to discriminate between glycoforms under conditions in which a portion of the lectin population was competitively inhibited. This work showed that accurate interpretation of lectin profiling data will require prior detailed assessment and validation of the performance of individual lectin interactions and the systems in which they will be used.

A robust aptamer assay was also developed to allow the future pursuit of the use of oligonucleotide arrays as glycosignaturing platforms. Various assay formats were developed for generated aptamers and these were used to evaluate the consensus sequences identified in the SELEX procedure. The assays demonstrated relative specificity and high binding affinities for Neu5Gc. To our knowledge this is the first report of DNA aptamer binding assay that shows differentiation between Neu5Gc and Neu5Ac. We have also demonstrated the detection of Neu5Gc in a glycoprotein sample. Thus, the assay provides the option of either direct analysis of the Neu5Gc on a glycoprotein of interest or measurement of the free Neu5Gc in solution.

1.4.5 WP5 Lectin-glycan interaction analysis

Studies on antibody fragments and lectins in WP5 have provided new insights into protein-glycan binding events and structure-function relationships. This aspect is of paramount importance for designing new lectin-mimics. Coordination among partners has been crucial, allowing the exploitation of complementary expertises and synergies. The results reported here provide a valuable background for a continued collaboration with several partners.

1.4.6 WP6 Glycoanalysis bioinformatics

In terms of impact, dissemination and exploitation, highlights from WP6 include:

- Software and data analysis methods of utility to the glycomics, proteomics and glyco-proteomics community:
    - ENViz – a Cytoscape plug-in that can be used to jointly analyse several data types. Described in Cytoscape meetings. Paper under review in Bioinformatics. Methods described in Steinfeld et al, NAR 2013.
    - Peak-calling s/w for .d files, to be used in the subtractive glycoproteomics approach developed by Agilent with KI
    - IEF fraction selection optimization. A module that performs the approach developed in GlycoHIT is incorporated into the next realease of SpectrumMill.
- Lectins with high syalic acid specificity. These reagents are of potential use and demand in the biotherapeutic industry. When the optimization is complete we will protect the leading variants.
- Subtractive IEF-LC/MS for glycoproteomics. An automatic approach to analysing data from pre and post PNGaseF for the same IEF fractions, enabling glycoproteomics of relatively complex mixtures. Internal IP processes at Agilent and at KI will address protection of IP derived from this approach.
- Serum glycobiomarkers for breast cancer. While Saldova et al points to the potential of some glycan groups as serum biomarkers in breast cancer much more work is necessary to turn these preliminary findings into commercializable markers.

1.4.7 WP7 Label-free sensor development

We have established and evaluated a range of potential carbohydrate biosensor platforms with detection based on EIS, ECL with and without nanoparticle-enhancement, plasmonic coupled SERS, and a flow cytometry-based suspension array assay. The use of some of the platforms for the *in situ* and dynamic evaluation of cell surface carbohydrate expression has been demonstrated. All biosensor platforms are dependent on the affinity of lectin or antibody for their target carbohydrates. These biosensors integrate the excellent biocompatibility and multivalent recognition of glycoproteins by lectins or antibodies with significant enzymatic catalysis and nanoparticle signal amplification, and possess high sensitivity, perfect stability and reproducibility for cell detection and cell surface carbohydrate expression evaluation. Such carbohydrate biosensors also provide a useful

tool for antibacterial drug screening and interaction mechanism study. This glycoprotein recognition and biocatalysis biosensing strategy would contribute to the understanding of complex native glycan-related biological processes. This strategy also provides a high sensitive method for dynamically analyzing changes of cell surface glycan in response to inhibitors, medicines and environmental or other stimulations. It can be expanded to provide even more impetus for elucidating the complicated mechanisms underlying glycan related biological and physiological processes.

1.4.8 WP8 Glycomics in cancer – a pilot study

Four HTP techniques to analyse glycan-related features were applied to clinical materials for validation. We have generated potential biomarkers for breast, colorectal and pancreatic cancer (in collaboration with Oslo University Hospital- Drs. Vilde Haakensen and Elin Kure) and examined effect of bioprocessing on glycosylation (in collaboration with BMS- Dr. Kirk Leister). One paper is already published, identifying serum *N*-linked glycan profiles based on UPLC analysis specific for breast cancer patients compared with healthy controls. This profile includes several traits not previously identified. Preliminary analyses indicate that lectin microarrays of serum glycans and HTP proteomics analyses of tumour proteins may be used to identify subtype-specific alterations in breast cancer patients. Treatment-specific alterations in the serum *N*-glycome has been identified for neoadjuvantly treated patients and the serum *N*-glycome of IgG is being explored by both lectin microarrays and UPLC to identify tumour-related alterations. In summary, technologies developed in this consortium have shown sensitivity and specificity sufficient to reveal clinically significant biological alterations in the proteome and *N*-glycome.

CHO cell glycosylation was analysed by UPLC and CHO cell supernatant from three processes were analysed by a combination of UPLC and lectin microarray technologies. The UPLC and lectin microarray results largely correlated for qualitative analyses, with the UPLC method additionally providing detailed quantitative data. It is possible that the combination of these two techniques could provide a powerful and rapid HTP profiling tool capable of yielding qualitative and quantitative data for a defined biopharmaceutical process, which would allow valuable near 'real-time' monitoring of the biopharmaceutical product. A manuscript on this work is currently being jointly compiled for publication between **Partners 1, 4 and 13**.

As a consequence of our collaboration and results, our new research activities as a direct result of GlycoHIT include the profiling of TIF/NIF/FIF (breast interstitial fluid) in breast cancer as a potential biomarker, which is being is carried out in collaboration with Dr. Vilde Haakensen, Oslo University Hospital. Additionally, the profiling of colorectal and pancreatic cancer sera (including matched control sera) before and after treatment in collaboration with Dr. Elin Kure, Oslo University Hospital is also continuing as is the profiling of lung cancer sera carried out in collaboration with Dr. Vilde Haakensen, Oslo University Hospital.

This research has in turn laid the foundations for the development of HTP methods to identify potential biomarkers in numerous diseases including different cancers, Galactosemia, alpha-1 antitrypsin deficiency, rheumatoid arthritis and congenital disorders of glycosylation.

Partner 9 (UMONS) carried out three studies under the WP8 : on (i) naso-sinusal diseases, (ii) thyroid lesions and (iii) head and neck cancers (oral cavity, pharynx, larynx). These studies led to different results that may have a socio-ecomonic impact. They were perfomed with anti-galectin antibodies and biotinylated-galectins provides by **Partner 3** (LMUM).

The first study on naso-sinusal diseases showed that galectin-9 is overexpressed in naso-sinusal carcinomas compared to chronic rhinosinusitis, nasal polyposis and inverted papillomas. This could be advantageous for the treatment of squamous cell carcinomas. The expression of galectins-1, -3, -7 and -8 binding sites found during the second study on thyroid lesions allowed adenomas to be distinguished from papillary carcinomas. A biotinylated galectin-1 or -7 immuno-staining corresponded to a papillary carcinoma and a biotinylated galectin-3 or -8 immuno-staining

correspond to an adenoma. This allows a differential diagnosis and avoids unnecessary surgery as 90% of surgeries are performed on benign tumours. Finally, the study conducted on head and neck cancers (oral cavity, pharynx and larynx) showed expression of galectin-2 in these cancers, increasing with tumour progression. This series will be soon expanded to assess the possible correlation between the galectin-2 expression and the prognosis. The role of this galecti-2 is not yet known, but if it is related to prognosis, it would be interesting to study in detail to define its precise role in tumour progression and eventually in recurrence or treatment resistance.

1.4.9 Publications and Dissemination Activites

The innovations achieved by the consortium described in this report have fuelled a comprehensive dissemination programme.

Website and On-line

The project website has been an important part of our dissemination plan. The website has been kept up to date throughout the period. In particular the news page of the website was frequently updated with items and events of interest.

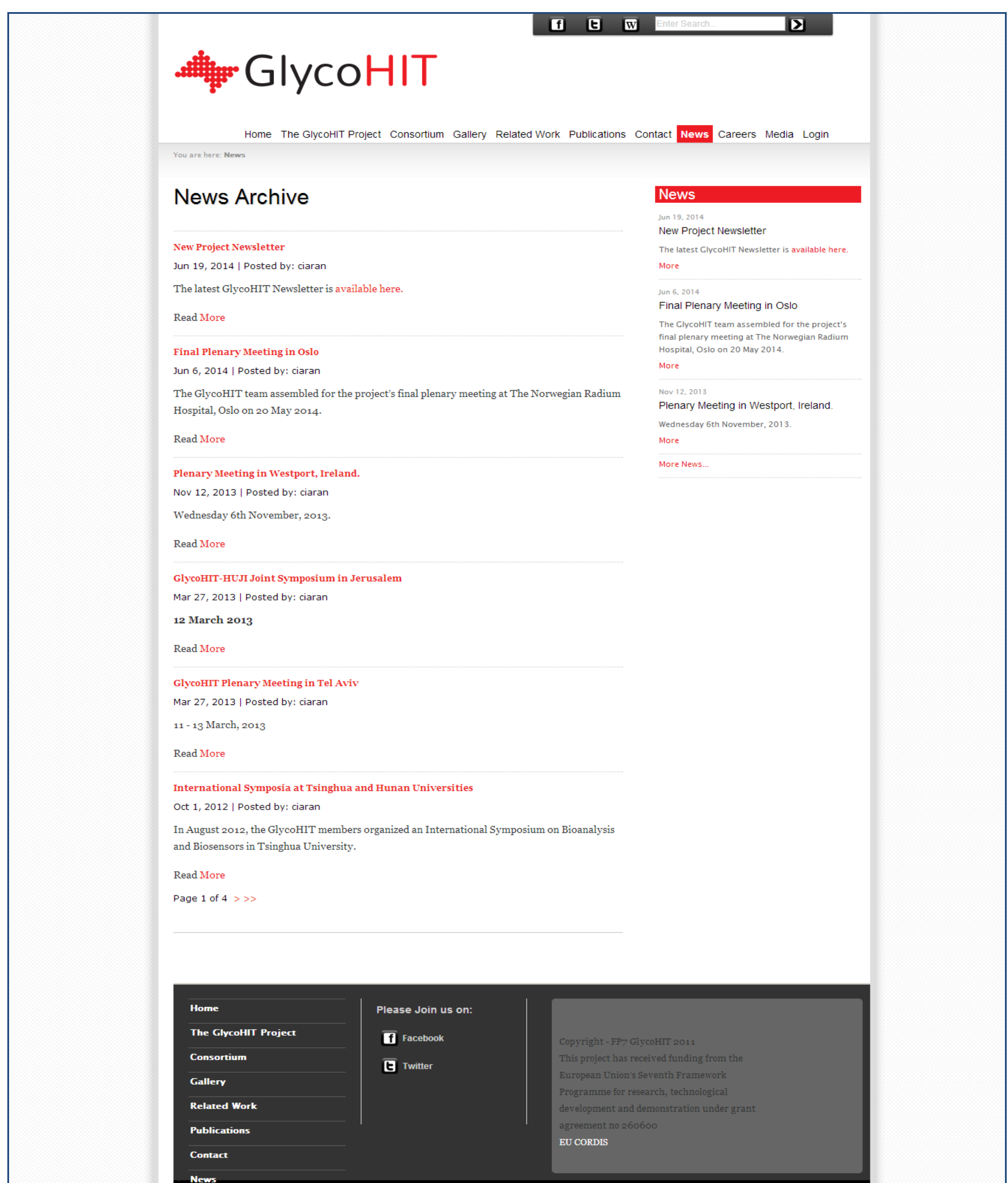


Figure 15 Screenshot extract from GlycoHIT website News Page

We recently added an on-line newsletter to the website. The newsletter is designed to capture project highlights, results, collaborations and project publications achieved and is available for download.

1

GlycoHIT

Welcome to the GlycoHIT on-line newsletter!

Glycohit at 'Research in Action' exhibit in Dublin's EU Commission house, NUIG researchers with EU Commissioner for Research, Innovation & Science Máire Geoghegan-Quinn

This edition includes the following:

- **GlycoHIT - Glycobiomarkers for Better Cancer Diagnosis**: a summary of what the project's all about.
- **GlycoHIT Collaboration with Other Researchers:** examples of GlycoHIT's collaboration with researchers based outside of the EU.
- **Key publications**: a selection of some significant GlycoHIT publications achieved by project partners.
- **News update:** recent project news of interest.
- **What's Next?:** next steps for the project and the partners.
- **GlycoHIT partners**: details of the GlycoHIT consortium.

Visit us at **www.GlycoHIT.eu**

GlycoHIT - Glycobiomarkers for Better Cancer Diagnosis

GlycoHIT aims to improve our understanding of cancer biomarkers and will develop rapid analysis technologies for diagnosing cancer using these biomarkers. The biomarkers that are under study are glycobiomarkers, arising from a process called protein glycosylation. Most of the known cancer-associated glycobiomarkers have been discovered using time-consuming, labour-intensive analysis methods. There is a critical need for innovative, rapid, and high-throughput (HTP) technologies to identify new glycobiomarkers for cancer, as well as to better diagnose cancer in patients. Rapid, HTP technologies will make sure that cancer-associated glycobiomarkers discovered as part of basic research programmes are used in clinics.

The **GlycoHIT** consortium brings together a highly experienced, innovative and interdisciplinary team of researchers from Europe, China and USA representing academia, industry and clinical fields to significantly enhance some of the existing glycoanalytical technologies and to advance novel HTP glycoanalytical technologies beyond the current state of the art.

Ultimate aims of the project include:

- To deliver new technologies that will enable the **rapid analysis of patient samples.**
- To **better understand glycobiomarkers associated with cancer** by performing glycan-profiling of cancer glycobiomarkers from serum, in a high-throughput manner.
- To **develop better diagnostic tests for cancer** by developing novel reagents and binding agents using human lectins as well as lectin mimics mined from antibody fragment, peptide and oligonucleotide libraries.
- To **identify new cancer glycobiomarkers** by improving sample preparation techniques and the interpretation of complex data emerging from glycoanalytical techniques.
- To **validate the new technologies** developed using samples from cancer patients.

GlycoHIT This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 260600.
© 2014 GlycoHIT.

Figure 16 Screenshot of GlycoHIT newsletter (page 1 of 6)

Project materials and flyer, etc. were produced for dissemination purposes.

In addition to the website, our on-line presence over the period has also included our Facebook page and Twitter account.

Publications and presentations

The partners have been particularly active in securing significant project publications in peer-reviewed journals and making presentations at conferences and events. By the end of the project, there were 137 dissemination activities listed on the project's dissemination activities register (available on the EU portal). In addition, there were 141 GlycoHIT papers published in scientific journals/chapters in edited books (as noted on the publications register on the EU portal). A selection of key GlycoHIT publications are described below.

1. Saldova, et al., Association of medication with the human plasma *N*-glycome. Journal of Proteome Research 2012 11 (3) pp 1821-1831. *Note: Details the importance of influence of medication on potential glycan biomarkers.*
2. McCarthy, et al., Increased outer arm and core fucose residues on the *N*-glycans of mutated alpha-1 antitrypsin protein from alpha-1 antitrypsin deficient individuals. J Proteome Research 2014, 13 (2), 596-605. *Note: Application of UPLC technology in screening of glycosylation in AAT.*
3. Saldova, et al., Increase in sialylation and branching in the mouse serum N-glycome correlates with inflammation and ovarian tumour progression. PLoS One, 2013, 8, e71159. *Note: Details the importance of inflammation in carcinogenesis and correlation with glycosylation and is therefore important for potential glycobiomarkers.*
4. Saldova, et al., Association of *N*-Glycosylation with Breast Carcinoma and Systemic Features Using High-Resolution Quantitative UPLC. J Proteome Research 2014, 13, 2314-27. *Note: Details the UPLC technology and application to breast cancer samples (result of collaboration of **Partners 4, 2 and 7**).*
5. Branca, et al. HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics. Nature Methods, 2014, 11(1), 59-62. *Note: Detailing the Hi-RIEF LC-MS technology.*
6. Amano, et al., Tumor suppressor p16(INK) (4a) : anoikis-favoring decrease in N/O-glycan/cell surface sialylation by down-regulation of enzymes in sialic acid biosynthesis in tandem in a pancreatic carcinoma model. FEBS J. 2012, 279(21), 4062-80. *Note: Details the orchestrated modulation of N/O-glycosylation and galectin expression to induce anoikis in human Capan-1 pancreatic carcinoma cells.*
7. Johansson, et al., Retinoic acid receptor alpha has potential predictive value in tamoxifen treated breast cancer patients. Nature Commun. 2013, 4, 2175. *Note: The nuclear receptor retinoic acid receptor alpha was identified as a marker of tamoxifen resistance and may be a novel therapeutic target.*
8. Nejman, et al., Molecular rules governing de novo methylation in cancer. Cancer Res. 2014, 74(5):1475-83. *Note: A universal set of tumour targets were identified each with its own 'coefficient of methylation. Methylation undergoes expansion early in tumourigenesis and it may act as an inhibitor of development-associated gene activation.*
9. Tahiri, et al., Deregulation of cancer-related miRNAs is a common event in both benign and malignant human breast tumors. Carcinogenesis (2014) 35 (1): 76-85. *Note: MicroRNA (miRNA) was examined in benign and malignant tumours lead to the identification of deregulated expression of oncomirs and tumor suppressor miRNAs in benign breast tumours.*
10. Wang, et al., A Functional Glycoprotein Competitive Recognition and Signal Amplification Strategy for Carbohydrate-Protein Interaction Profiling and Cell Surface Carbohydrate Expression Evaluation, Nanoscale, 2013, 5, 7349-7355. *Note: A sensitive biosensor for carbohydrate-lectin profiling and in situ cell surface carbohydrate expression was designed by taking advantage of a functional glycoprotein of glucose oxidase acting as both a multivalent recognition unit and a signal amplification probe.*
11. Chen, et al., Dynamic Evaluation of Cell Surface *N*-Glycan Expression via Electrogenerated Chemiluminescence Cytosensor Integrating Concanavalin A and Gold Nanoparticles Modified Ru(bpy)32+-Doped Silica Nanoprobe, Anal. Chem., 2013, 85, 4431−4438. *Note: The biosensor was employed to dynamically profile cell surface N-glycan expression at different phases of cell growth* in vitro*.*

12. Remmelink, et al., Quantitative immunohistochemical fingerprinting of adhesion/growth-regulatory galectins in salivary gland tumors: divergent profiles with diagnostic potential. Histopathology, 2011, 58, 543-556. *Note: Details the proof-of-principle study for diagnostic potential.*

13. Gabius, et al., From lectin structure to functional glycomics: principles of the sugar code. Trends Biochem Sci., 2011, 36, 298-313. *Note: Seminal review on the sugar code (cited more than 100 times).*

14. Ruiz, et al., Natural single amino-acid polymorpism (F19Y) in human galectin-8: detection of structural alterations and increased growth-regulatory activity on tumor cells. FEBS J. 2014 Volume 281, Issue 5, pages 1446–1464. *Note: Pilot study on combined structural and cell biological analyses of a clinically relevant variant from single amino-acid polymorphism for human galectin-8 detecting short- and long-range effects up to alteration of growth regulation.*

15. Canales, et al., Breaking pseudo-symmetry in multiantennary complex N-glycans using lanthanide-binding tags and NMR pseudo-contact shifts. Angew Chem Int Ed Engl., 2013, 52, 13789-93. *Note: First example of the derivation of the independent conformational behaviour of the two branches of a complex-type N-linked glycan and their interaction with lectins. Considered as VIP paper by Angew Chem.*

16. Vidal, et al., Conformational selection in glycomimeics. Human galectin-1 only recognizes syn/ψ type conformations of β1,3-linked lactose and its C-glycosyl derivative. Chem. Eur. J., 2013, 19, 14581-14590. *Note: Combined analysis of glycan binding to a human galectin, with relevance for in vivo recognition and drug design.*

17. Toegel, et al., Glycophenotyping of osterarthritic cartilage and chondrocytes by RT-PCR, mass spectrometry, histochemistry with plant/human lectins and lectin localization with a glycoprotein. Arthrit. Res. Ther., 2013, 15, R147. *Note: Glycan biomarker detection in an autoimmune disease of socioeconomic impact, i.e. arthritis.*

18. Cunningham, et al., Development of a convenient competitive ELISA for the detection of the free and protein-bound non-human galactosyl-α-(1,3)-galactose epitope based on highly specific chicken single-chain antibody variable-region fragments. Anal. Chem., 2013, 85, 949-955. *Note: Detailing the generation and characterisation of specific anti-Gal-�-(1,3)-Gal scFvs and development of a competitive quantitative ELISA suitable for glycoprotein therapeutics.*

19. Gerlach, et al., Microarray evaluation of the effects of lectin and glycoprotein orientation and data filtering on glycoform discrimination. Anal. Methods, 2014, 6, 440-449. *Note: Use and validation of lectin microarray platform for glycosignaturing of different glycoprotein glycoforms.*

20. Wang, et al., Cross-platform comparison of glycan microarray formats. Glycobiology, 2014, 24, 507-517. *Note: Technical comparison of lectin recognition of glycan ligands across microarray platforms and the effect of surface chemistry, presentation and linkers.*

21. Svarovsky, et al., Cancer glycan biomarkers and their detection – past, present and future. Anal. Methods, 2014, 6, 3918-3936. *Note: Timely review of cancer glycobiomarkers and the current status and future developments of glycoanalytics.*

Attendance and presentation at specialised conferences is another important facet of dissemination. A schedule of GlycoHIT dissemination activities (including conferences organised by GlycoHIT partners and presentations by members of the consortium) is included on the EU portal. The following conferences are of particular impact or importance to the project:

1. 17th European Carbohydrate Symposium (EuroCarb17), Tel-Aviv, Israel, July 7-11th 2013. *Note: A number of GlycoHIT partners presented at this conference including:*
    - Steinfeld, et al., Serum Glycomics and Integrated Data Analysis in Breast Cancer.
    - Kilcoyne, et al., Deciphering host-pathogen interactions: Construction and use of a mammalian lectin microarray. *Note: This talk shared the work on the mammalian lectin microarray platform using bacterial interactions as a model. This conference was attended by leading European scientists with research interests in glycobiology and industry representatives.*
    - Rudd, et al., Linking the genome and the glycome: State of the art glycoanalytics for systems glycobiology, biomarker discovery and the pharmaceutical industry.
2. Yakhini, Advanced Measurement Techniques with Applications to Glycobiology, GlycoHIT symposium at CN-HUPO at Chong-Qing, China, September 2013.
3. 6th Annual Glycoscience Ireland Meeting at Westport, Ireland, November 2013. *Note: This conference was co-organised by **Partners 1 and 4** and took place following the GlycoHIT plenary meeting in Westport, Ireland. It gave the GlycoHIT partners the opportunity to network, discuss and disseminate their findings to Irish national and international industrial and academic researchers. Partner presentations included the following:*
    - Yakhini, Z. Using synthetic long oligonucleotides in protein optimization and in studying the role of glycosylation in breast cancer.
    - Rudd, PM./ Stockmann, H. When genomics meets glycomics a new picture emerges.
    - Rudd, PM./ Hickey, R. Oligosaccharide analysis/ Milk oligosaccharide function and analysis.
    - Haakensen, V. Breast cancer systems biology and glycan-related alterations.
4. Joshi, L. Glycobiology of pathogenic and commensal bacteria and gut cells, Microbial Glycobiology FASEB, Itasca, IL, USA, June 8-13th, 2014. *Note: This talk shared the work on glycomics array platforms with leading members of the international microbial glycobiology community.*
5. **Partner 4** co-organised a meeting held in Split on Glycobioinformatics in systems biology, involving the European Bioinformatics Institute in the University of Cambridge, U.K. and the Swiss Institute of Bioinformatics, June 26-27th, 2014.
6. Rudd, P.M. Glycomics and genomics, IBD and oncology meeting Alberta, Canada, June 6th 2014. *Note: This talk highlighted importance of glycosylation in cancer and link to genomics.*
7. Haakensen, V. Glycan gene expression in normal and malignant breast tissue, Nordic proteoglycan workshop, Oslo, Norway, May 14th, 2013. *Note: This talk presented our findings on glycan gene expression changes in breast cancer which was highly relevant to the GlycoHIT project outputs.*
8. An important GlycoHIT paper entitled "*Optimizing analytical depth and cost efficiency of IEF-LC/MS proteomics" (Kifer, Branca, Xu, Ben-Dor, Lehtio, Yakhini)* was accepted to BIBM 2014. BIBM is the leading bioinformatics conference for IEEE and our paper was accepted as a full paper and selected to be presented in a full talk. Ilona Kifer from GlycoHIT Partner Agilent will present the paper at the conference in Belfast (2-5 November 2014).

Engagement with other researchers, projects and initiatives

In addition to reaching a worldwide audience though our extensive record of scientific publications, the team has been involved in a number of targeted initiatives to disseminate our findings and further collaboration with other researchers and initiatives.

Tsinghua University and Hunan University are valued members of the GlycoHIT team. Building on this collaboration, GlycoHIT is now closely linked to a sister project in China (funded under the "973" Programme and International Collaboration Programme from the Chinese Ministry of Science and Technology) which aims to develop technologies for quantitative proteomic studies of protein post-translational modification in liver disease. Engagement with our colleagues in China is building potential for future collaboration and common application in diagnostic and predictive medicine for both cancer and liver disease.

Several meetings have been arranged between the GlycoHIT researchers and their Chinese counterparts over the course of the project. The 1st China-Europe Symposium between the projects was held in Hangzhou, China during the 7th CNHUPO (China Human Proteome Organization) conference in 2011. In August 2012, another International Symposium on Bioanalysis and Biosensors was arranged in Tsinghua University and Hunan University also hosted the second joint meeting between GlycoHIT and its sister Chinese consortium in September 2012. Most recently, the 3rd China-Europe symposium on glycoproteomics was held as part of the 8th CNHUPO meeting, on September 9/10, 2013 in Chongqing, China.

Figure 17 Members of the GlycoHIT team, China-Europe Symposium, 2012

Other examples of engagement with related projects and initiatives during this reporting period include the following:

- On 12 March 2013 the Israel Institute of Advanced Studies hosted a joint GlycoHIT - Hebrew University of Jerusalem (HUJI) symposium on Glyco-Proteomics: Technologies and the Role in Human Health and Disease. Members of the GlycoHIT consortium and the Hebrew University of Jerusalem presented their work and perspectives at the meeting.
- A joint GlycoHIT-AGRC meeting was held in Westport Ireland in November 2013. The AGRC (the Alimentary Glycosciences Research Cluster) is an inter-institutional, multi-disciplinary consortium of academic and industrial researchers funded by Science Foundation Ireland.
- GlycoHIT researchers attended the GlycoScience Ireland (GSI) Annual Conference in Westport Ireland in November 2013.

- Institute seminars presented by GlycoHIT Principal Investigators were organized at the Institute for Cancer Research at the Norwegian Radium Hospital on 22 May 2014.
- GlycoHIT researchers presented at the inaugural session of the Glycomics Symposium, titled "Increasing the Impact of Glycoscience through New Tools and Technologies", held July 19-21 2012 in San Sebastián. The conference, organized by Prof. Niels-Christian Reichardt of CIC biomaGUNE, Guipúzcoa, featured a broad-range of carbohydrate-related biological, computational, chemical, and business topics presented through over fifty talks and thirty-two posters. Several members of the Glycoscience Group from NUI Galway, including GlycoHIT-funded researchers, presented posters and gave oral presentations at the meeting.
- GlycoHIT exhibited at the 'Research in Action' exhibit in Dublin's EU Commission house, July 9th - 20th 2012.

1.5 *Website and contact details*

Website and on-line

The following images are extracts from the GlycoHIT website and other online resources and illustrate how the project is communicating and disseminating its results online. For further information please see the project website at **http://www.glycohit.eu/**

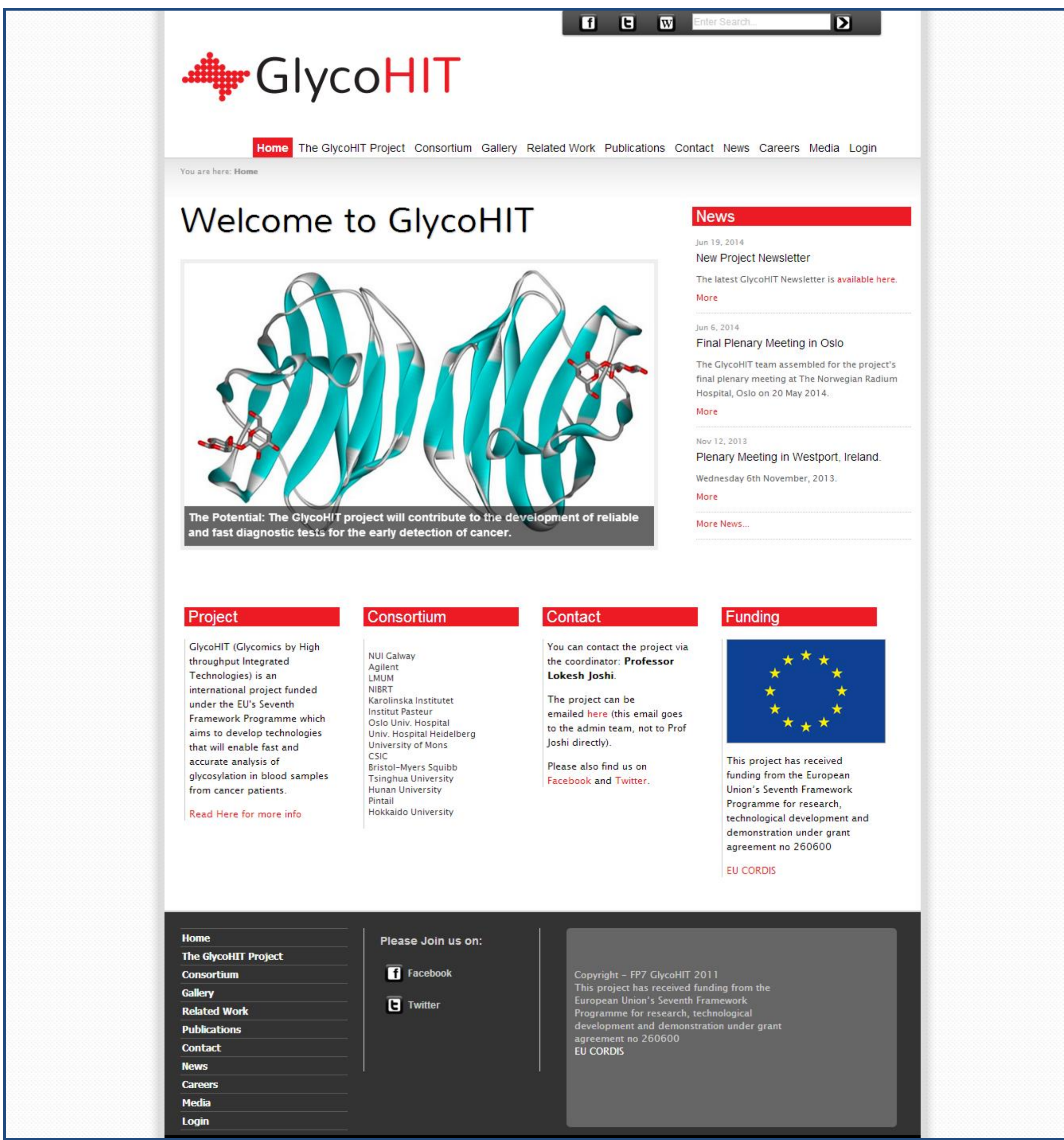


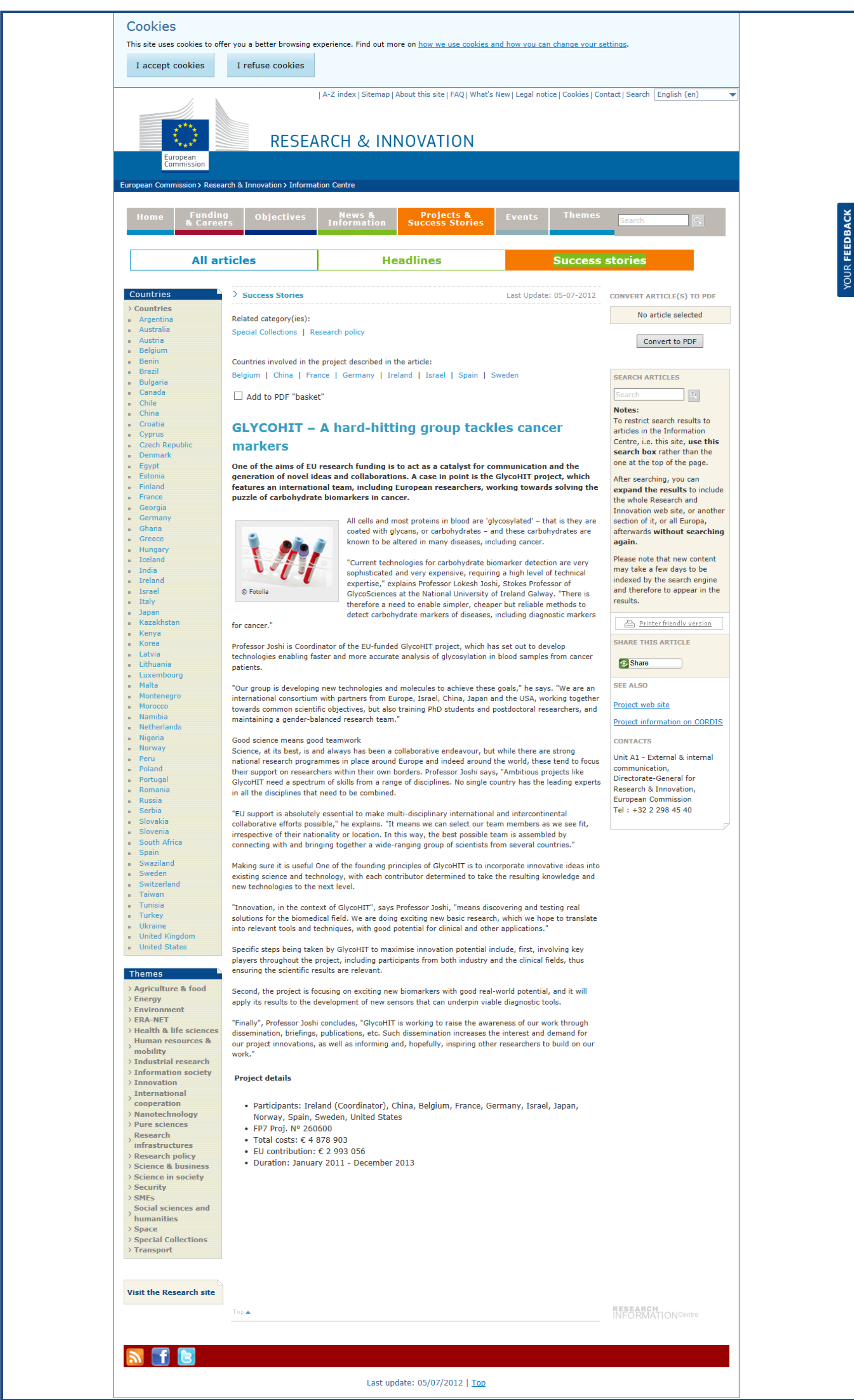
Figure 18 Screenshot of GlycoHIT Homepage http://www.glycohit.eu/

You can also follow us on twitter and/or on Facebook.

Figure 19 Extract from Facebook page

GlycoHIT was featured on the ec.europa.eu website:

http://ec.europa.eu/research/infocentre/article_en.cfm?id=/research/star/index_en.cfm?p=ss-glycohit&calledby=infocentre&item=Countries&artid=26054&caller=SuccessStories



RESEARCH & INNOVATION

Success Stories — Last Update: 05-07-2012

Related category(ies):
Special Collections | Research policy

Countries involved in the project described in the article:
Belgium | China | France | Germany | Ireland | Israel | Spain | Sweden

GLYCOHIT – A hard-hitting group tackles cancer markers

One of the aims of EU research funding is to act as a catalyst for communication and the generation of novel ideas and collaborations. A case in point is the GlycoHIT project, which features an international team, including European researchers, working towards solving the puzzle of carbohydrate biomarkers in cancer.

© Fotolia

All cells and most proteins in blood are 'glycosylated' – that is they are coated with glycans, or carbohydrates – and these carbohydrates are known to be altered in many diseases, including cancer.

"Current technologies for carbohydrate biomarker detection are very sophisticated and very expensive, requiring a high level of technical expertise," explains Professor Lokesh Joshi, Stokes Professor of GlycoSciences at the National University of Ireland Galway. "There is therefore a need to enable simpler, cheaper but reliable methods to detect carbohydrate markers of diseases, including diagnostic markers for cancer."

Professor Joshi is Coordinator of the EU-funded GlycoHIT project, which has set out to develop technologies enabling faster and more accurate analysis of glycosylation in blood samples from cancer patients.

"Our group is developing new technologies and molecules to achieve these goals," he says. "We are an international consortium with partners from Europe, Israel, China, Japan and the USA, working together towards common scientific objectives, but also training PhD students and postdoctoral researchers, and maintaining a gender-balanced research team."

Good science means good teamwork
Science, at its best, is and always has been a collaborative endeavour, but while there are strong national research programmes in place around Europe and indeed around the world, these tend to focus their support on researchers within their own borders. Professor Joshi says, "Ambitious projects like GlycoHIT need a spectrum of skills from a range of disciplines. No single country has the leading experts in all the disciplines that need to be combined.

"EU support is absolutely essential to make multi-disciplinary international and intercontinental collaborative efforts possible," he explains. "It means we can select our team members as we see fit, irrespective of their nationality or location. In this way, the best possible team is assembled by connecting with and bringing together a wide-ranging group of scientists from several countries."

Making sure it is useful One of the founding principles of GlycoHIT is to incorporate innovative ideas into existing science and technology, with each contributor determined to take the resulting knowledge and new technologies to the next level.

"Innovation, in the context of GlycoHIT", says Professor Joshi, "means discovering and testing real solutions for the biomedical field. We are doing exciting new basic research, which we hope to translate into relevant tools and techniques, with good potential for clinical and other applications."

Specific steps being taken by GlycoHIT to maximise innovation potential include, first, involving key players throughout the project, including participants from both industry and the clinical fields, thus ensuring the scientific results are relevant.

Second, the project is focusing on exciting new biomarkers with good real-world potential, and it will apply its results to the development of new sensors that can underpin viable diagnostic tools.

"Finally", Professor Joshi concludes, "GlycoHIT is working to raise the awareness of our work through dissemination, briefings, publications, etc. Such dissemination increases the interest and demand for our project innovations, as well as informing and, hopefully, inspiring other researchers to build on our work."

Project details

- Participants: Ireland (Coordinator), China, Belgium, France, Germany, Israel, Japan, Norway, Spain, Sweden, United States
- FP7 Proj. N° 260600
- Total costs: € 4 878 903
- EU contribution: € 2 993 056
- Duration: January 2011 - December 2013

Last update: 05/07/2012 | Top

Figure 20 Screenshot from GlycoHIT feature on ec.europa.eu website

GlycoHIT Partners

GlycoHIT is composed of 15 partners. The consortium includes partners not only from across Europe, but also from China, Japan, Israel and the United States. The multi-disciplinary researchers in the project are supported by an experienced project management team.

Figure 21 GlycoHIT at 'Research in Action' exhibit in Dublin's EU Commission House.
NUI Galway researchers with EU Commissioner for Research,
Innovation & Science Máire Geoghegan-Quinn

Figure 22 Members of the GlycoHIT team at the consortium meeting in Oslo, May 2014