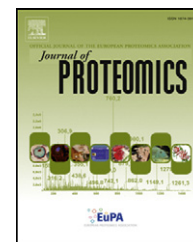


Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot

Identification of plasma Complement C3 as a potential biomarker for neuroblastoma using a quantitative proteomic approach

Patrick Y. Kim^a, Owen Tan^a, Sonya M. Diakiw^a, Daniel Carter^a, Eric O. Sekerye^a, Valerie C. Wasinger^b, Tao Liu^a, Maria Kavallaris^a, Murray D. Norris^a, Michelle Haber^a, Lou Chesler^c, Alla Dolnikov^a, Toby N. Trahair^{a,d}, Nai-Kong Cheung^e, Glenn M. Marshall^{a,d,*}, Belamy B. Cheung^{a,**}

^aChildren's Cancer Institute Australia for Medical Research, Randwick, NSW 2031, Australia

^bBioanalytical Mass Spectrometry Facility, Mark Wainwright Analytical Centre, University of New South Wales, Sydney, NSW, Australia

^cDivision of Cancer Biology, Institute for Cancer Research, Sutton, Surrey, UK

^dKids Cancer Centre, Sydney Children's Hospital, Randwick, NSW 2031, Australia

^eDepartment of Pediatrics, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY, United States

ARTICLE INFO

Article history:

Received 19 February 2013

Accepted 22 October 2013

Available online 4 November 2013

Keywords:

Neuroblastoma

Biomarker

Zinc-alpha2-glycoprotein

Complement C3

TH-MYCN^{+/+}

Complement system

ABSTRACT

The majority of patients diagnosed with neuroblastoma present with aggressive disease. Improved detection of neuroblastoma cancer cells following initial therapy may help in stratifying patient outcome and monitoring for relapse. To identify potential plasma biomarkers, we utilised a liquid chromatography–tandem mass spectrometry-based proteomics approach to detect differentially-expressed proteins in serum from TH-MYCN mice. TH-MYCN mice carry multiple copies of the human MYCN oncogene in the germline and homozygous mice for the transgene develop neuroblastoma in a manner resembling the human disease. The abundance of plasma proteins was measured over the course of disease initiation and progression. A list of 86 candidate plasma biomarkers was generated. Pathway analysis identified significant association of these proteins with genes involved in the complement system. One candidate, complement C3 protein, was significantly enriched in the plasma of TH-MYCN^{+/+} mice at both 4 and 6 weeks of age, and was found to be elevated in a cohort of human neuroblastoma plasma samples, compared to healthy subjects. In conclusion, we have demonstrated the suitability of the TH-MYCN^{+/+} mouse model of neuroblastoma for identification of novel disease biomarkers in humans, and have identified Complement C3 as a candidate plasma biomarker for measuring disease state in neuroblastoma patients.

Biological significance

This study has utilised a unique murine model which develops neuroblastoma tumours that are biologically indistinguishable from human neuroblastoma. This animal model has

* Correspondence to: G. M. Marshall, Children's Cancer Institute Australia for Medical Research, Randwick, NSW 2031, Australia. Tel.: +61 2 93821721; fax: +61 2 93821789.

** Corresponding author. Tel.: +61 2 9385 2450; fax: +61 2 9662 6584.

E-mail addresses: g.marshall@unsw.edu.au (G.M. Marshall), bcheung@ccia.unsw.edu.au (B.B. Cheung).

effectively allowed the identification of plasma proteins which may serve as potential biomarkers of neuroblastoma. Furthermore, the label-free ion count quantitation technique which was used displays significant benefits as it is less labour intensive, feasible and accurate. We have been able to successfully validate this approach by confirming the differential abundance of two different plasma proteins. In addition, we have been able to confirm that the candidate biomarker Complement C3, is more abundant in the plasma of human neuroblastoma patient plasma samples when compared to healthy counterparts. Overall we have demonstrated that this approach can be potentially useful in the identification of biomarker candidates, and that further validation of the candidates may lead to the discovery of novel, clinically useful diagnostic tools in the detection of sub-clinical neuroblastoma.

© 2013 The Authors. Published by Elsevier B.V. Open access under [CC BY license](#).

1. Introduction

Neuroblastoma is an embryonal tumour of the sympatho-adrenal tissues, and is one of the most common tumours in early childhood [1]. More than half of children with neuroblastoma present clinically beyond 18 months of age with aggressive disease resulting in high mortality rates. Most patients enter clinical remission with intensive chemo-radiotherapy and surgery, only to relapse later from a point of minimal residual disease (MRD). Accurate detection of disease activity following initial therapy, at a stage of MRD, would better predict relapse risk and aid in directing the choice of further treatment. Furthermore, the identification of sensitive biomarkers may offer the hope of a pre-clinical diagnosis strategy in infants prior to the initial clinical presentation at an earlier disease stage. mRNA expression levels of neuroblastoma-specific genes, such as paired-like homeobox 2B (PHOX2B) and tyrosine hydroxylase (TH), in circulating cells from blood samples taken during therapy have been developed for the detection of MRD [2]. However, heterogeneous expression of these biomarkers in tumour samples decreases the sensitivity of this strategy. Since no sensitive plasma protein biomarkers are currently available for risk estimation at diagnosis or during therapy, there is immense clinical value for the identification of novel biomarkers aimed at detection of neuroblastoma [3].

The MYCN oncogene is amplified in tumour tissue from 20–25% of children with neuroblastoma and predicts a poor prognosis [4,5]. The *TH-MYCN*^{+/+} mice used as a neuroblastoma disease model in this study express the human MYCN gene under the control of a rat TH promoter in neuroectodermal cells [6]. Large, paravertebral thoracic or abdominal tumours develop in 100% of homozygous *TH-MYCN*^{+/+} mice at 6 weeks of age and 33% in the hemizygous mice at 13 weeks of age [7]. Histologic, immunohistochemical and cytogenetic studies have shown that these tumours are indistinguishable from human neuroblastoma tissue [6]. Tumour formation in these mice follows delayed and incomplete regression of neuroblast hyperplasia in a manner representative of human childhood cancer initiation, which is believed to develop as a consequence of abnormal persistence of embryonal rest cells by largely unknown mechanisms. This model is ideal to use for the characterisation of differences in protein profiles from plasma samples of *TH-MYCN*^{+/+} and wild-type (WT) littermate mice, as they relate to different stages of tumourigenesis.

Label-free mass spectrometry allows for the identification and characterisation of individual proteins and peptides

within complex biological samples using ion count or spectral count techniques [8–10]. In combination with protein fractionation techniques, such as in-solution electrophoresis, this technique reduces the complexity of proteomic samples, and allows for increased sensitivity which is particularly important for detecting biologically significant proteins expressed at low levels in plasma where the total concentration range of proteins can exceed ten orders of magnitude [11]. Mass spectrometry techniques have led to the characterisation of several novel biomarkers which are currently being developed for the diagnosis of bladder [12], ovarian [13] and lung cancers [14]. The ion count label-free approach determines the abundance of a 'peptide' based on the normalised ion count of peptide features between sample groups, and has been successfully used to identify deregulated and abnormally expressed proteins in human disease [15,16].

In this study, plasma was harvested from *TH-MYCN*^{+/+} and WT control mice at 2, 4, and 6 week time-points, representing different stages in the initiation and progression of tumourigenesis. Using in-solution electrophoretic enrichment of the lower mass component, liquid chromatography–tandem mass spectrometry (LC–MS/MS) and label-free quantitation, approximately 200 differentially abundant proteins from each time-point were identified. The quantitative proteomics approach was validated using 2D gel electrophoresis and ELISA, and a final list of 86 protein candidates was generated which demonstrated significant differential expression at more than one time point. One candidate, Complement C3, was quantified in the plasma of a cohort of human neuroblastoma patients at diagnosis and found to be elevated compared to healthy controls, suggesting that this protein should be further studied as a plasma biomarker for assessing disease activity in neuroblastoma patients.

2. Materials and methods

2.1. Sample collection

Blood from ten *TH-MYCN*^{+/+} transgenic and ten WT mice were collected by cardiac puncture into vacutainers (5 ml EDTA tubes with a purple top), from 2, 4 and 6 week old mice groups (n = 60). Ethical approval for this study was obtained from the Animal Care and Ethics Committee (Animal ethics approval number 11/110A). Blood plasma was separated via centrifugation (3000 g for 5 min) and stored at –80 °C for further use.

Sympathetic ganglia or tumours were dissected from age matched WT and *TH-MYCN*^{+/+} mice respectively for later analysis by ELISA. Upon dissection, the isolated tissue was washed and cleaned of non-specific tissue in Hanks Balanced Salt Solution and snap-frozen for storage. The sympathetic ganglia (pooled superior cervical ganglia and coeliac ganglia) were isolated from WT mice for comparison with early and advanced-stage tumours from 4 week old and 6 week old *TH-MYCN*^{+/+} mice respectively. For each age/genotype, ganglia or tumours were isolated from 4 different mice and later purified as protein lysates for ELISA.

Human plasma samples from healthy volunteers were obtained from the Sydney Red Cross (N = 4) and Sydney Cord Blood Bank (N = 3) under the ethics approval (#08/148) from the South Eastern Sydney Illawarra. The samples obtained from the Red Cross were from healthy adults and the samples from the Cord Blood Bank were from healthy newborns. These samples were used as a normal healthy control for this project. The plasma samples from neuroblastoma patients were collected by Dr Nai-Kong Cheung from the Memorial Sloan-Kettering Cancer Centre (New York, USA), under the clinical trial study, Molecular Characterisation of Neuroblastic Tumour: Correlation with Clinical Outcome (Clinical Trials identifier: NCT00588068) and with the IRB protocol number #00–109. These samples were from patients with MYCN-amplified tumours (n = 9), which were all stage VI and MYCN-non-amplified tumours (n = 6), where four of these were from stage I and one each from stages II and III. The average age of the patients from which the samples were collected from was 3.7 years and all of these samples were collected at time of diagnosis.

2.2. Sample, fractionation, purification and analysis by LC–MS/MS

Aliquots (200 µg plasma proteins for each sample) were pooled and an appropriate volume of protease inhibitor was added (Complete™, mini, EDTA-free, protease inhibitor tablet from Roche Diagnostics, Mannheim, Germany) as per manufacturer recommendations. An in-solution electrophoretic device called ProteomeSep (NuSep, NSW, Australia) was used to analyse the plasma from each genotype and each time-point. Individual plasma samples were also stored for further analysis with 2D gel electrophoresis and ELISA. The plasma proteins were partitioned into five fractions based on protein mass, pI and surface charge using the ProteomeSep, a preparative electrophoresis instrument which causes charged molecules to migrate across hydro gel membranes into different molecular weight fractions when a voltage is applied. ProteomeSep fractionations were performed using 5-chamber assembly, where chamber boundaries were defined by 2.7 × 2.2 cm polyacrylamide membranes of different restrictive pore sizes. All buffers described for sample preparations were also used as circulating buffers. Plasma was diluted 1:1 to give a final concentration of 14 µg/ml plasma in 45 mM Tris/5 mM ε-aminocaproic acid (EACA) buffer in 2 M urea, pH 10.2. All fractionations were performed at 50 V for 30 min and then 250 V for 2 h at 15 °C. The fractions were recovered from the chambers using gel-loading tips. The fractions collected were; 1–5 kDa, 5–25 kDa, 25–50 kDa, 50–75 kDa and > 75 kDa. The two low molecular mass fractions, 1–5 kDa and

5–25 kDa fractions were desalted using C18 Stage Tips (Proxeon, Odense, Denmark) as per manufacturer's instructions.

The peptides were resuspended in 5 µl of buffer A 2% (v/v) acetic acid, 0.1% (v/v) FA and separated by nano-LC using an Ultimate 3000 HPLC and autosampler system (Dionex, Netherlands) and then concentrated and desalted onto a micro C18 pre-column (500 µm × 2 mm, Michrom Bioresources, USA) with 0.05% (v/v) heptafluorobutyric acid (HFBA) at 20 µl/min. After a 4 min washing, the pre-column was automatically switched (Valco 10 port valve, Houston, Texas, USA) into line with a fritless nano column, as previously published [17]. The LTQ-FT (Thermo Scientific, San Jose, CA, USA) was operated as described previously [17]. Peak lists were generated using 'Mascot Daemon/extract_msn' (Matrix Science), default parameters and then submitted to the protein database search programme Mascot (Matrix Science, Boston, USA).

2.3. Label-free quantification of identified peptides using ion count

All LC–MS/MS spectra were searched against the mammalian non-redundant NCBI database (2010) using Mascot (www.matrixscience.com). In solution digests analysed on the LTQ-FT Ultra were searched with the following criteria: trypsin digestion, precursor and product ion tolerances ±6 ppm and ±0.6 Da, respectively; variable modifications of methionine oxidation, phosphorylation (S, Y, T), carbamidomethyl and acrylamide; and one missed cleavage allowed. Identifications were accepted based on the MOWSE scores with a score > 44 indicating significant homology (p < 0.05).

Acquired spectra run in triplicate for each time point and in each fraction files were loaded into the Progenesis LC–MS/MS software (version 2.5, Nonlinear) for label-free quantification. The datasets generated by the MS analysis of plasma proteins from 2, 4 and 6 week old *TH-MYCN*^{+/+} and WT mice were transformed into peak lists representing a 2D map using retention time and m/z. Each size fraction from each time-point was analysed separately and aligned according to retention time to create a 2D LC feature map. Features selected for inclusion were charge state inclusive of MH2⁺ to MH3⁺, retention time from 13 min, and non-matching features were filtered out from further analysis. The samples were then allocated to their experimental group, including positive group (*TH-MYCN*^{+/+}) and negative group (WT). Statistical parameters (p < 0.05 and fold-change > 5) were added to the sample analysis criteria to identify significant differences in peptide quantification. Peptides differing in either *TH-MYCN*^{+/+} or WT mice were followed closely during the progression of disease. Parent ion information that was used to quantitate the peptides, and then the daughter ions were used to search Mascot, after which the identifications imported and linked back to each parent ion "feature".

2.4. Database search and protein identification

All LC–MS/MS spectra were searched using Mascot against the mammalian non-redundant NCBI database with the following criteria: the precursor tolerance and the product ion tolerances were at 6 ppm and ±0.6 Da respectively; variable modification of methionine oxidation and trypsin digest. Mascot generated Dat.files were also then accumulated in

Progenesis LC-MS/MS software to compare and assess for differential abundance based on ion counts.

2.5. Removal of plasma abundant proteins for validation assays

Mouse abundant plasma proteins were depleted from each sample using the ProteoSpin™ abundant plasma protein depletion kit (Norgen Biotek Corp., Ontario, Canada), according to the manufacturer's protocol.

Eluted protein samples were measured for protein concentration using the BCA protein assay (Thermo Scientific) and then stored at -20°C for further downstream experiments.

2.6. Two dimensional analysis of protein candidates

Abundant protein depleted plasma samples from 6 week old TH-MYCN^{+/+} transgenic and WT mice were separated and analysed on 2D gels. IPG strips (11 cm, pH 4–7, Bio-Rad) were rehydrated with 250 μl rehydration buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 2% (v/v) carrier ampholytes pH 3–10, 10 mM Tris, 43 mM DTT and 0.01% (v/v) bromophenol blue) for 16 h using a gel reswelling tray (GE Healthcare, Australia). Depleted plasma samples (200 μg) were mixed with rehydration buffer to a final volume of 200 μl and incubated at room temperature for 30 min with orbital mixing (1000 rpm), to solubilise and denature the plasma proteins. For, IEF, each sample was loaded onto the rehydrated IPG strips using sample cups, which were placed at the acidic end of the strip. IEF was then performed at room temperature for 120 kVh using the IPGphor electrophoresis unit (GE Healthcare). After IEF, the IPG strips were equilibrated (50 mM Tris-HCl pH 8.8, 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, 0.01% (v/v) bromophenol blue with trace amounts of DTT) for 20 min on an orbital shaker at low speed. Each of the IPG strips were laid on top of 4–20% Criterion gels (Bio-Rad) and sealed using a 0.5% (w/v) agarose sealing solution (Tris/Glycine/SDS buffer, 0.5% (w/v) agarose and 0.01% (v/v) bromophenol blue) for electrophoresis in the second dimension. Electrophoresis was performed on the gel using the buffer system of Laemilli [18] at 100 V until the bromophenol blue dye front migrated off the gel. After electrophoresis, the proteins on the 2D gel were transferred on to nitrocellulose membrane (GE Healthcare) by wet Western transfer. ZAG protein was detected on the membrane with rabbit polyclonal anti-ZAG antibody (Sigma-Aldrich, MO, USA); using a working dilution of 1:500 in an appropriate buffer.

2.7. ELISA quantitation of protein candidates

Sandwich enzyme-linked immuno sorbent assay (ELISA) kits were used for the quantification of ZAG (Uscn Life Science Inc. Wuhan, China), complement C3, C-reactive protein (CRP) and Serum Amyloid A (SAA) (Genway Biotech Inc. CA). Concentrations of ZAG and complement C3, proteins in the plasma of TH-MYCN^{+/+} transgenic and WT mice as well as the human plasma samples were determined according to the manufacturer's protocol. In brief, mouse plasma samples were diluted with dilution buffer, 1:1000 for ZAG and 1:50,000 for complement C3. Human plasma samples were diluted with dilution buffer, 1:30 for ZAG and 1:50,000 for complement C3.

Plasma concentration levels of CRP and SAA were determined in the human samples using the aforementioned ELISA kits and by diluting the samples 1:200 for CRP and 1:50 for SAA. C3 protein extracted from mouse tissue was diluted at 1:500 and protein concentration was determined according to the manufacturer's protocol.

2.8. Immunohistochemical analysis of TH-MYCN^{+/+} tissue samples

Paraffin-embedded section of TH-MYCN^{+/+} tissue at 6 weeks of age was harvested and fixed with formalin. Paraffin-embedded sections were dried in a 60°C oven for 1 h. The tissue slides were then immersed in xylene to deparaffinise the section. Slides were then re-hydrated with 100%, 90%, and 70% ethanol sequentially. Antigen retrieval was done by immersing the slides in 0.01 M tri-sodium citrate buffer with 0.05% Tween-20, pH6, at 104°C for 15 min. The endogenous peroxidases were inactivated by immersing slides in 3% hydrogen peroxide. 10% goat serum was used to block non-specific binding of immunoglobulin. The slides were incubated with rabbit anti-ZAG (Sigma Aldrich) at 1:150 dilution overnight at 4°C . Rabbit IgG prepared to the same concentration was used as negative control. Secondary antibody used was goat anti-rabbit immunoglobulin/biotinylated (Dako, E0432) at 1 in 500 dilution for 1 h, room temperature. The biotinylated antibody was then labelled with streptavidin-HRP (Dako, K1016) for 45 min at room temperature. The sections were developed with 3, 3'-diaminobenzidine tetrahydrochloride (Dako, K3468) and counterstained with haematoxylin.

2.9. Bioinformatics, pathway analysis and statistical analysis

For relative protein expression analysis, fold change was calculated in TH-MYCN^{+/+} mice compared to WT littermates and \log_2 transformed for normalisation. Unsupervised hierarchical cluster analysis was performed using Multi-Experiment Viewer (MEV-TM4 microarray analysis software suite) to generate evolutionary gene trees (Euclidean distance metric selection and average linkage clustering parameters) [19]. To identify functional protein groups and pathways the 86 candidate proteins were assessed using Gene Ontology (GO) terms via the Memorial Sloan-Kettering Cancer Centre database (<http://cbio.mskcc.org/CancerGenes>) and also the Broad Institute Molecular Signatures database (MSigDB — <http://www.broadinstitute.org/gea/msigdb>). Statistical significance was evaluated using the Mann-Whitney, unpaired, nonparametric t-test in GraphPad Prism 5 (p -values < 0.05 were considered statistically significant).

3. Results

3.1. Identification and quantification of differentially abundant plasma proteins from TH-MYCN^{+/+} and WT mice

The plasma collected from 10 TH-MYCN^{+/+} and 10 WT mice at 2, 4 and 6 weeks of age was pooled and prefractionated using the electrophoretic ProteomeSep device, to reduce the

complexity of the plasma proteome, prior to LC-MS/MS identification and quantification. Fig. 1A outlines the mass spectrometry and label-free approach designed for identification of differentially abundant candidate plasma proteins in *TH-MYCN^{+/+}* and WT mice and generation of the final list of candidate biomarkers for further analysis. Fig. 1B shows a representation of the peptides identified by LC-MS/MS separated by retention time and *m/z*. The datasets generated into peak lists and analysed by Progenesis LC-MS/MS identified > 200 proteins per size fraction and time-point to be differentially abundant in the plasma of *TH-MYCN^{+/+}* and WT mice (± 5 -fold relative peptide abundance and $p < 0.05$).

3.2. Profiling differentially abundant plasma proteins in *TH-MYCN^{+/+}* and WT mice

To validate the label free LC-MS/MS method we selected zinc-alpha2-glycoprotein (ZAG) as it displayed one of the highest concentrations in the plasma of *TH-MYCN^{+/+}* mice compared to controls at 6 weeks of age (408-fold higher concentration). For 2D Western immunoblot analysis, pooled plasma samples from ten *TH-MYCN^{+/+}* and ten WT mice were analysed following abundant plasma protein depletion. Three distinct isoforms of ZAG were detected in both the *TH-MYCN^{+/+}* and WT

pooled samples (Fig. 2A), however, all three isoforms in the *TH-MYCN^{+/+}* sample were more intense than the isoforms in the WT sample. A similar finding was observed by 2D Western immunoblot analysis of plasma from 3 individual *TH-MYCN^{+/+}* and 3 individual WT mice (Fig. 2B). Importantly, plasma from 2 of the WT controls (B5767 and B5850) only displayed 2 of the 3 isoforms of ZAG, which may reflect different functional contributions of ZAG isoforms in the development of neuroblastoma [20]. Consistent with the label-free analysis and 2D Western immunoblot data, ZAG concentrations were also found to be significantly higher in the plasma of *TH-MYCN^{+/+}* mice using ELISA. ELISA quantitation of pooled plasma 6 week old mice (three pooled plasma sets, with each set containing ten mice), showed that the *TH-MYCN^{+/+}* mice had an average ZAG concentration of approximately 500 pg/ml compared to 120 pg/ml in the plasma of 10 WT mice (p -value: 0.0034) (Fig. 2C). Quantitation of ZAG in the sera from 4 individual mice showed median plasma concentrations of ZAG in *TH-MYCN^{+/+}* samples were approximately 300 pg/ml compared with 180 pg/ml for WT mice (p -value: 0.0003) (Fig. 2D). When plasma levels of ZAG were measured in human plasma samples, no significant difference could be detected between the average values of ZAG concentration in 7 normal healthy plasma samples compared with 14 neuroblastoma patient

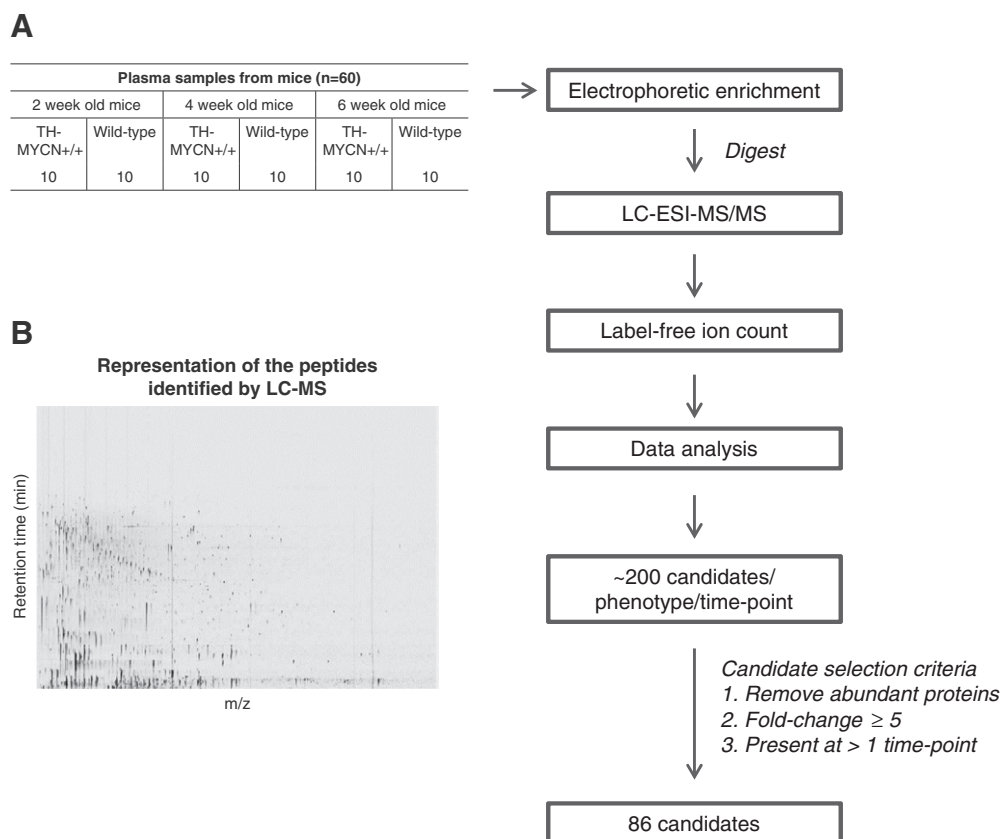


Fig. 1 – Overall study design implemented for neuroblastoma plasma protein biomarker identification, refinement and validation. (A) Plasma harvested from *TH-MYCN^{+/+}* and WT mice were pre-fractionated into five size fractions prior to LC-MS/MS and subsequent peptide quantification analysis was performed with ion count. A final list of 86 candidate proteins was generated based on differential abundance at more than one time point, and a fold change of at least 5 at one time-point. (B) A representative of a reference runs selection for the peptides, which was separated by retention time and *m/z* ratio.

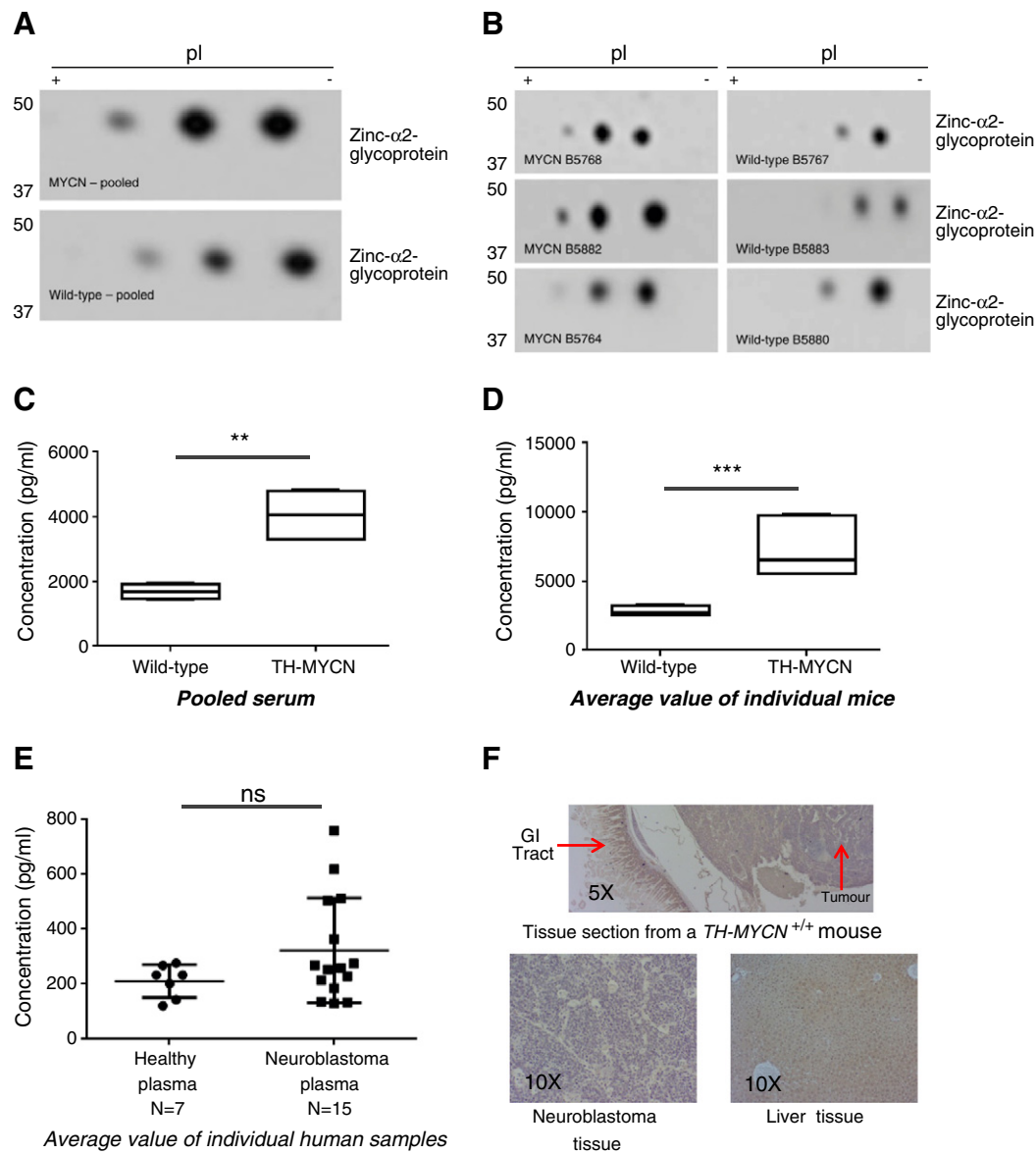


Fig. 2 – Validation of the proteomics approach using 2D Western immunoblotting, ELISA and IHC. (A) Pooled plasma samples from 6 week old TH-MYCN^{+/+} and WT mice were analysed by 2D electrophoresis and probed with an anti-ZAG specific polyclonal antibody. (B) Plasma collected from three independent TH-MYCN^{+/+} and three WT mice were analysed with 2D Western immunoblots and an anti-ZAG specific polyclonal antibody. (C) Three different sets of pooled plasma samples from 6 week old TH-MYCN^{+/+} and WT mice were analysed using the ZAG ELISA kit. Asterisks (**) represent the p -value = 0.0034 obtained using the Student's t-test. (D) Plasma concentrations of ZAG was measured from four individual TH-MYCN^{+/+} and WT mice with the ZAG ELISA kit, and average values plotted to determine the concentration of ZAG. Asterisks (***) represent the p -value = 0.0003. (E) Concentrations of ZAG was measured from the plasma of 7 healthy individuals (3 cord blood and 4 peripheral blood samples) and 15 neuroblastoma patient samples using the ZAG ELISA kit, (NS, non-specific). (F) Immunohistochemical staining for ZAG, using an anti-ZAG specific polyclonal antibody on tissue samples obtained from TH-MYCN^{+/+} mice. Organs indicated by the red arrows include the gastrointestinal (GI) tract and neuroblastoma tissue.

plasma samples (Fig. 2E). However, it should be noted that the ZAG concentration range in the healthy individual samples were in range of 120–276 pg/ml, where as some of the neuroblastoma patient plasma samples exhibited a ZAG plasma concentration of more than 760 pg/ml. All concentrations were calculated using data extrapolated from a standard

curve. To investigate whether ZAG was expressed in the neuroblastoma tumour, we performed immunohistochemical (IHC) analysis on tissue samples obtained from TH-MYCN^{+/+} mice (Fig. 2F). Synthesis and secretion of ZAG has been reported to occur in both the gastrointestinal (GI) tract as well as the liver [21]. As indicated by the IHC staining of the

mouse tissue samples, the staining of ZAG can be observed in both the GI tract and the liver. However, a much lower level of ZAG expression could be detected in the neuroblastoma tissue samples, compared with the GI tract and the liver. This suggested that the greater abundance of ZAG in the plasma of the *TH-MYCN*^{+/+} mice is not directly secreted from the tumour, but could be from another source that is affected by the initiation or the progression of neuroblastoma. Further studies will be required to determine the origin of ZAG in *TH-MYCN*^{+/+} mice bearing neuroblastoma tumours.

3.3. Assessment of the top differentially abundant plasma proteins in *TH-MYCN*^{+/+} and WT mice

From our initial analysis, we removed abundant proteins including albumin, immunoglobulins, transferrin, alpha-1-antitrypsin and haptoglobin. The list was further refined to include only proteins which were differentially expressed at more than one time point to generate a final list of 86 candidates. The top candidates were clustered into 4 main groups based on expression patterns over the time-course of the experiment in *TH-MYCN*^{+/+} mice relative to WT mice (Fig. 3).

Two groups of proteins (group 2 and group 3) demonstrated higher average expression in plasma from *TH-MYCN*^{+/+} mice compared to WT mice at the early 2 week time-point during tumour initiation. It is likely that these groups will contain suitable biomarkers for early detection of neuroblastoma at diagnosis or relapse. The average expression of group 1 by contrast was higher in *TH-MYCN*^{+/+} mice at 4 weeks during early tumour formation, and expression of group 4 was high at both 4 weeks and 6 weeks suggesting proteins associated with early to advanced tumour formation. Interestingly, the majority of proteins identified have not previously been associated with neuroblastoma. We found a high proportion of proteins which are secreted extracellularly (12 biomarkers from the 86 total biomarkers), several of which are associated with other solid tumour types such as Serpin6 (marker of resistance to therapy in breast cancer — [22]) and FN1 (protects against apoptosis in lung tumourigenesis — [23]).

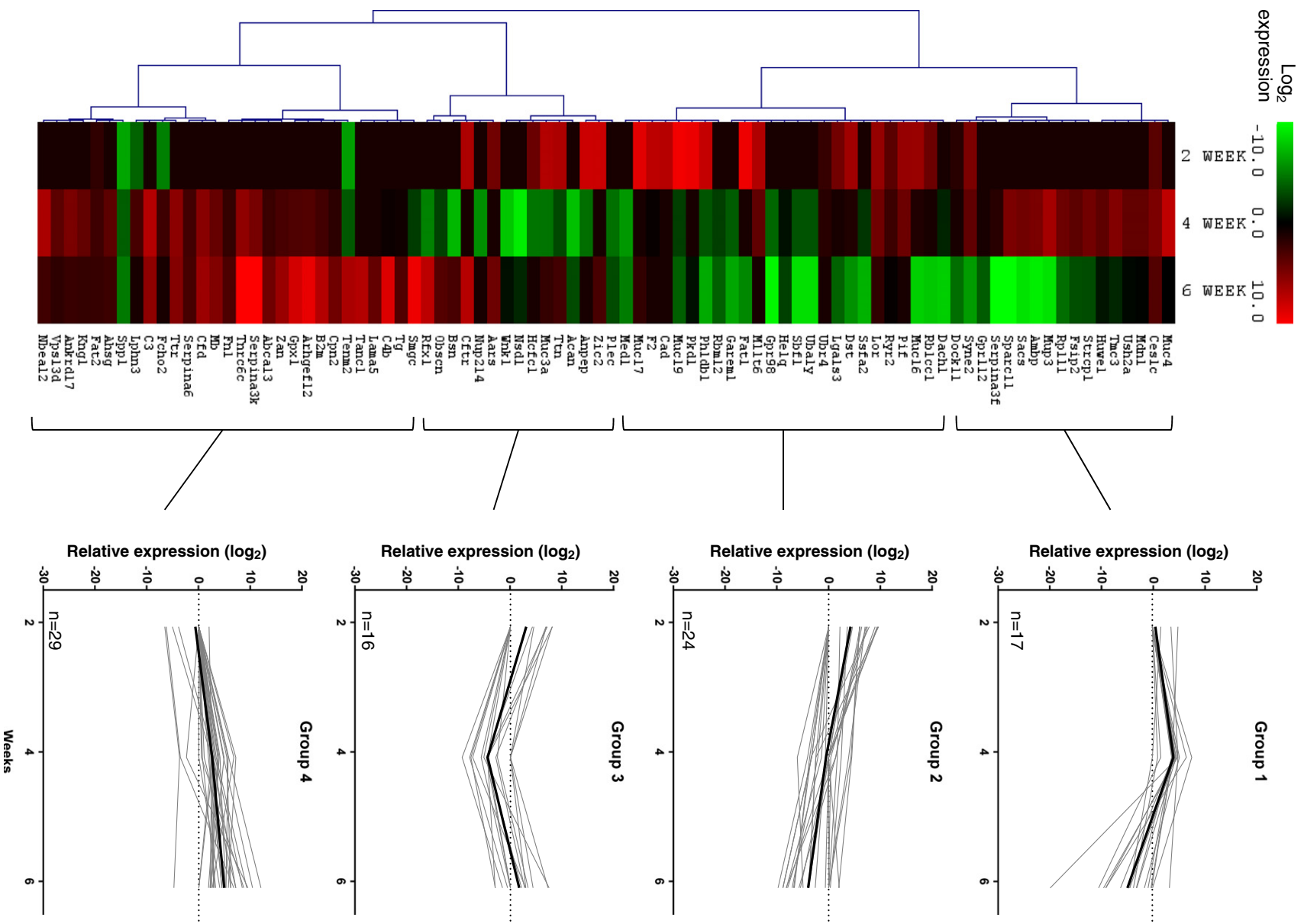
Pathway analysis was performed on the 86 candidate protein set using MSigDB and the top 10 hits ranked on *p*-value are shown in Fig. 4A. The top pathway identified involves protein metabolism, specifically O-glycosylation, a post-translational modification which has known roles in the regulation of cellular nutrient metabolism, protein stability and ligand presentation for functional recognition [24]. Two additional pathways related to glycosylation and metabolism were identified in the top 50 pathway hits (ranked #10 and #16), and four of the 5 proteins in these data sets belonged to expression group 2 with high expression in *TH-MYCN*^{+/+} mice at the early 2 week time-point (Fig. 4B). The second ranked pathway to be identified involves the complement and coagulation cascades, two linked proteolytic cascades with roles in inflammation, immunosuppression, and tumour development/metastasis [25]. A total of 5 protein sets were identified in the top 50 pathways relevant to the complement system, and 4 of the 5 proteins from our analysis in these data sets clustered in group 4 with enhanced abundance at 4 and 6 weeks in the *TH-MYCN*^{+/+} (Fig. 4C). Interestingly, we found a total of 6 data sets related to liver development, hepatoblastoma, or liver cancer in our top 50

pathways (including two ranked within the top 10). These datasets comprised a total of 23 of our candidate proteins, with 10 of these clustered in expression group 4 (shown in Fig. 4D). The liver is a common site of metastasis for neuroblastoma [26], however our finding could also suggest some overlaps in the molecular mechanisms or the pathological processes of these two cancer types.

3.4. Complement C3 protein is more abundant in plasma samples from human neuroblastoma patients than healthy volunteers

Complement C3 protein (C3) was significantly elevated in *TH-MYCN*^{+/+} mice compared to WT counterparts at both 4 and 6 weeks of age (140-fold and 50-fold higher respectively). The pathway analysis identified C3 as part of both the complement cascade and the liver cancer groups in our analysis (Fig. 4), and as C3 levels are elevated in other cancer types we chose to analyse it further in our mouse model and in human plasma samples. The proteomics data was validated using ELISA performed on pooled plasma from 10 mice of each genotype per time-point, which confirmed significantly higher C3 levels in *TH-MYCN*^{+/+} mice at both 4 and 6 weeks of age (7-fold and 2-fold higher than in WT mice respectively, Fig. 5A). The pattern of expression of C3 was similar to the initial proteomics analysis, although the magnitude of difference in plasma protein concentration was reduced when detected using ELISA. To determine whether elevated plasma C3 was a consequence of higher C3 expression in tumour tissues, ELISA was used to detect C3 levels in ganglia from WT mice and in tumours from *TH-MYCN*^{+/+} mice. C3 protein was significantly elevated in tumours from *TH-MYCN*^{+/+} mice compared to WT counterparts at 6 weeks of age (Fig. 5B), however no difference was observed at the earlier 4 week time-point. Whilst this observation may reflect a reduced limit of detection for C3 protein using ELISA, it could alternatively indicate that accumulation of C3 in tumours is secondary to increased levels in plasma.

To determine if these experimental observations would translate to a similar finding in human disease, plasma samples obtained from healthy adults (*N* = 4) and newborns (*N* = 3) were compared with plasma from 15 patients with neuroblastoma at diagnosis for expression of C3 protein. The mean plasma level of C3 in healthy individuals was 20 ng/ml (range: 8–37 ng/ml) compared to 110 ng/ml (range: 73–176 ng/ml) in neuroblastoma patients (median age 3.7 years) as determined by ELISA (Fig. 5C). As C3 is an important component of the immune system, we also assessed levels of two other inflammation-associated proteins – acute phase response protein C-reactive protein (CRP) and Serum Amyloid A (SAA) – to determine if C3 levels in neuroblastoma plasma samples were non-specifically elevated due to an inflammatory response. We did not observe significant differences in plasma concentration of CRP or SAA between healthy individuals and neuroblastoma patients using ELISA (Supplementary Fig. 1), suggesting that elevated C3 may potentially be a tumour-specific biomarker. Further validation will be required in larger patient populations as well as age-matched controls to confirm the validity of C3 as a plasma biomarker of neuroblastoma. This study also demonstrates the feasibility of using the *TH-MYCN*^{+/+} mouse model to identify



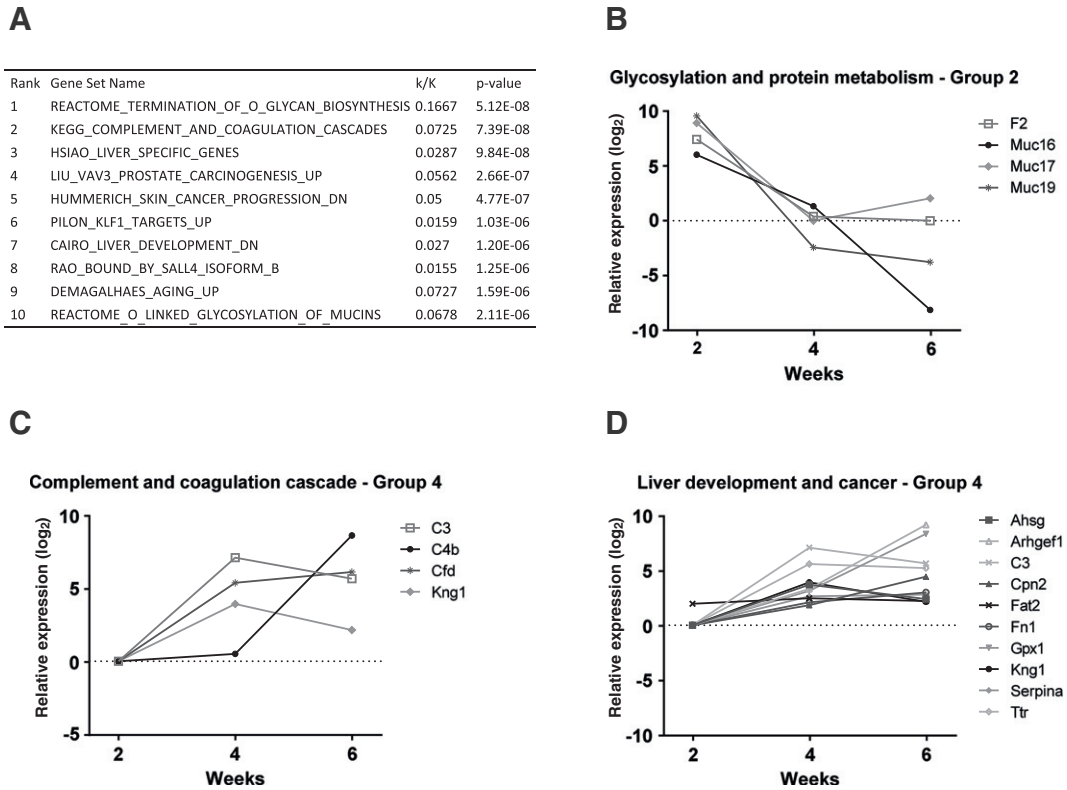


Fig. 4 – Identification of functional groups of protein biomarkers using pathway analysis. (A) Top 10 significantly associated pathways identified using MSigDB. Pathways were ranked by *p*-value (k/K indicates proportion of protein candidates which overlap with the indicated data-set). (B–D) Expression patterns of individual proteins in *TH-MYCN*^{+/+} mice relative to WT controls forming part of the top three pathways identified in part A.

novel neuroblastoma biomarkers with relevance to human disease.

4. Discussion

Initiation of neuroblastoma is executed by a series of molecular events that is mediated by dysfunctional genes and proteins [1]. Postnatally persistent embryonal cells with tumorigenic capacity possess the capacity to resist cell death imposed during the final stages of neurodevelopment and undergo secondary changes that characterize later tumour promotion and progression [27]. At 2 weeks of age, *TH-MYCN*^{+/+} mice show the first signs of neuroblastoma tumour initiation with the appearance of neuroblast hyperplasia, which persists until tumour formation and progression [7]. Studies have shown that this aberrant survival is due to a diminished p53 response to conditions of cellular stress, such as serum starvation or withdrawal of trophic factor [28,29]. We hypothesised that identification of proteins

differentially expressed at these early time-points, prior to the onset of overt neuroblastoma, may be clinically useful as indicators to predict the initiation and onset of neuroblastoma.

In this study, we employed a label-free approach based on a comparison of peptide intensities, and quantitative analysis by ion count, to assess differences in plasma protein concentrations from *TH-MYCN*^{+/+} and WT mice for the identification of candidate plasma protein biomarkers of neuroblastoma. This methodology has several advantages over traditional proteomic analysis techniques utilising 2D polyacrylamide gel electrophoresis (PAGE), including improved detection of low mass and basic proteins and enhanced reproducibility [30]. Using a combination of the label-free proteomics approach and a transgenic *MYCN* mouse model of neuroblastoma, we identified a list of 86 candidate protein biomarkers present in plasma with altered levels compared to WT mice. A total of 27 proteins from the candidate list were significantly higher in plasma of the *TH-MYCN*^{+/+} mice than in WT littermates at the early 2 week time-point, and 49 were higher at the 4 week

Fig. 3 – Time-course expression patterns of 86 candidate protein biomarkers in *TH-MYCN*^{+/+} mice relative to WT controls. Relative expression of proteins was plotted at 2, 4 and 6 weeks of age and hierarchical cluster analysis performed to generate 4 groups of proteins with similar expression trends. The evolutionary tree is depicted on the left, and expression patterns of proteins within each of the 4 groups is shown on the right (grey lines represent expression of individual proteins, black lines represent mean expression pattern for each group).

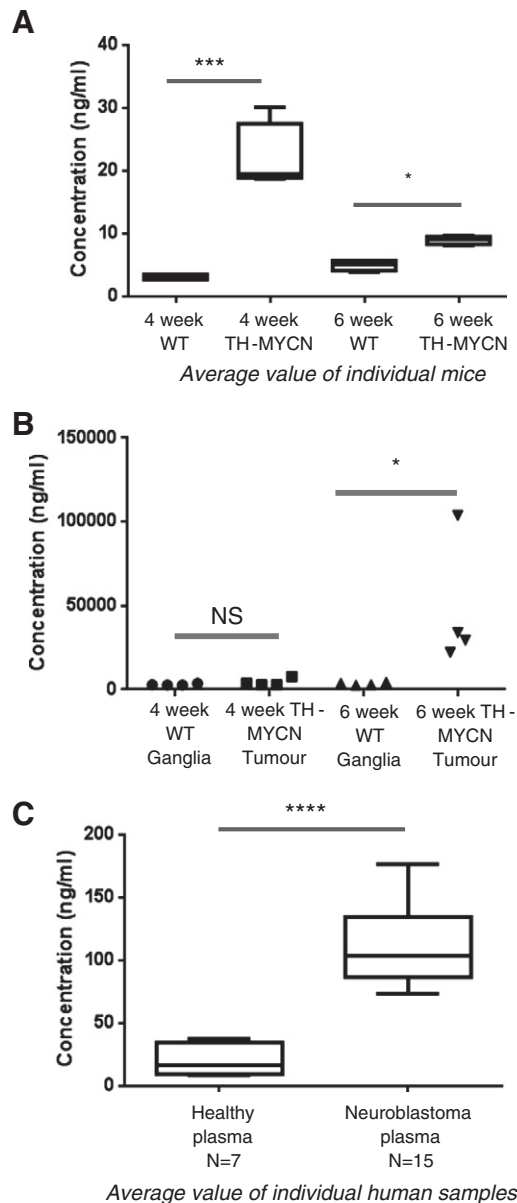


Fig. 5 – Plasma concentration levels of complement C3 in the *TH-MYCN*^{+/+} mouse model and in human neuroblastoma patient plasma samples. (A) ELISA assays demonstrate a significant difference in the concentration of complement C3 in the pooled plasma of 4 week old *TH-MYCN*^{+/+} and WT mice. Asterisks (*) represent *p*-value < 0.001. (B) Proteins extracted from the neuroblastoma tumour and ganglia tissue from 4 and 6 week old *TH-MYCN*^{+/+} (N = 4/genotype/time-point) and WT mice were analysed with a complement C3 ELISA. Asterisk (*) represents *p*-value < 0.05. (C) ELISA assays were performed to detect C3 protein levels of plasma in a cohort of human neuroblastoma plasma samples (N = 15) and in plasma from healthy individuals (N = 3 from cord blood and N = 4 from peripheral blood). Asterisks (****) represent *p*-value < 0.0001.**

time-point, suggesting that these candidates should be considered as markers for early detection of neuroblastoma.

Cluster and pathway analyses identified a group of proteins involved in post-translational protein modification, specifically glycosylation, which demonstrated significantly higher expression in *TH-MYCN*^{+/+} mice at 2 weeks of age before steadily declining during disease progression. Aberrations in the glycosylation of proteins have been associated with almost all forms of human cancers [31]. Studies have also suggested that aberrations in protein glycosylation can result in the initiation of oncogenesis, or that oncogenic transformation can cause alterations in protein glycosylation [31]. Other studies have shown similarities in the glycosylation pattern of proteins in different human neuroblastoma. The glycopeptides extracted from neuroblastoma tumours exhibited a low percentage of fucose-containing glycosylation [32]. Thus, it appears that the potential biomarkers identified by our cluster and pathway analysis may have significant roles in neuroblastoma biology. Albeit, further functional studies of each candidate will be required to confirm this hypothesis.

Pathway analysis also identified the complement system as being significantly associated with our candidate protein list. From this data, C3 was further characterised as a plasma biomarker candidate for neuroblastoma in the *TH-MYCN*^{+/+} murine model and in primary human plasma samples from neuroblastoma patients. Activation of C3 is central to the three complement pathways, which collectively result in the inflammatory response and the elimination of self- and non-self-antigenic targets [33]. However, using assays that detect CRP and SAA plasma levels we have shown that elevated C3 levels are not due acute phase responses and are most likely linked to neuroblastoma tumour formation. C3 is elevated in sera from pancreatic adenocarcinoma patients compared to normal individuals [34], and high levels are also observed in serum from breast, colorectal and lung cancer patients [30]. According to The Royal College of Pathologists of Australasia manual, the normal range of C3 in human plasma is recorded to be 0.9–1.8 g/l. Complement proteins have a central role in innate immunity, where they destroy antibody-coated targets such as apoptotic cells [35]. Deregulation of these molecules has also been associated with autoimmune and non-immune diseases [36,37]. In the context of carcinogenesis, the complement C3 protein has been implicated in the production of VEGF and extracellular matrix reorganisation and disintegration which contributes to the tumorigenic hallmarks of angiogenesis, invasion and migration, respectively [38].

Complement C4-B and complement factor D, which are other members of the complement family were identified by ion count to be more abundant in the plasma of *TH-MYCN*^{+/+} compared to the WT mice at 4 and 6 weeks of age. Therefore, it is possible that as an immune response to the initiation and development of neuroblastoma, several members of the complement protein family have been found to be more abundant in the plasma of *TH-MYCN*^{+/+} mice, compared to their WT counterparts. It also should be noted that a recent study has suggested that complement protein members C3, C4 and C5a may aid tumour growth by a mechanism of immunosuppression [38,39]. Thus, the abundance of these complement proteins in the plasma of the *TH-MYCN*^{+/+} mice

may either be a marker of the immune response to the cancer or aid in the development and growth of neuroblastoma.

5. Conclusions

Early detection and disease recurrence monitoring are critical areas in cancer treatment in which specific biomarker panels are likely to be very important in these key areas. The MYCN transgene of the mouse model used in our studies accurately reflects the more clinically aggressive forms of neuroblastoma. The validity of our data is strengthened by the finding that ZAG and C3, chosen for validation, have been previously published as serum markers in cancer. Further characterisation of the candidate biomarkers in our neuroblastoma animal model, and in human plasma patient samples, will likely result in a panel of biomarkers, that could be utilised for early detection and prediction of relapse risk in childhood neuroblastoma.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jprot.2013.10.032>.

Acknowledgements

This research was supported by Australian Postgraduate Research Scholarship, University of NSW, Program Grants from the NHMRC Australia, Cancer Institute NSW, and Cancer Council NSW. The Children's Cancer Institute Australia for Medical Research is affiliated with the University of NSW and Sydney Children's Hospital.

REFERENCES

- [1] Brodeur GM. Neuroblastoma: biological insights into a clinical enigma. *Nat Rev Cancer* 2003;3:203–16.
- [2] Kagedal B. Detecting minimal residual disease in neuroblastoma: still a ways to go. *Clin Chem* 2009;55:1268–70.
- [3] Escobar MA, Hoelz DJ, Sandoval JA, Hickey RJ, Grosfeld JL, Malkas LH. Profiling of nuclear extract proteins from human neuroblastoma cell lines: the search for fingerprints. *J Pediatr Surg* 2005;40:349–58.
- [4] Beierle EA. MYCN, neuroblastoma and focal adhesion kinase (FAK). *Front Biosci (Elite Ed)* 2011;3:421–33.
- [5] Park JR, Eggert A, Caron H. Neuroblastoma: biology, prognosis, and treatment. *Hematol Oncol Clin North Am* 2010;24:65–86.
- [6] Weiss WA, Aldape K, Mohapatra G, Feuerstein BG, Bishop JM. Targeted expression of MYCN causes neuroblastoma in transgenic mice. *EMBO J* 1997;16:2985–95.
- [7] Hansford LM, Thomas WD, Keating JM, Burkhardt CA, Peaston AE, Norris MD, et al. Mechanisms of embryonal tumor initiation: distinct roles for MycN expression and MYCN amplification. *Proc Natl Acad Sci U S A* 2004;101:12664–9.
- [8] Sahab ZJ, Semaan SM, Sang QX. Methodology and applications of disease biomarker identification in human serum. *Biomark Insights* 2007;2:21–43.
- [9] Findeisen P, Neumaier M. Mass spectrometry based proteomics profiling as diagnostic tool in oncology: current status and future perspective. *Clin Chem Lab Med* 2009;47:666–84.
- [10] Graham DR, Elliott ST, Van Eyk JE. Broad-based proteomic strategies: a practical guide to proteomics and functional screening. *J Physiol* 2005;563:1–9.
- [11] Srinivas PR, Verma M, Zhao Y, Srivastava S. Proteomics for cancer biomarker discovery. *Clin Chem* 2002;48:1160–9.
- [12] Vlahou A, Schellhammer PF, Mendrinos S, Patel K, Kondylis FI, Gong L, et al. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine. *Am J Pathol* 2001;158:1491–502.
- [13] Wang L, Chen S, Zhang M, Li N, Chen Y, Su W, et al. Legumain: a biomarker for diagnosis and prognosis of human ovarian cancer. *J Cell Biochem* 2012;113:2679–86.
- [14] Tan F, Jiang Y, Sun N, Chen Z, Lv Y, Shao K, et al. Identification of isocitrate dehydrogenase 1 as a potential diagnostic and prognostic biomarker for non-small cell lung cancer by proteomic analysis. *Mol Cell Proteomics* 2012;11 [M111 008821].
- [15] Hauck SM, Dietter J, Kramer RL, Hofmaier F, Zipplies JK, Amann B, et al. Deciphering membrane-associated molecular processes in target tissue of autoimmune uveitis by label-free quantitative mass spectrometry. *Mol Cell Proteomics* 2010;9:2292–305.
- [16] Merl J, Ueffing M, Hauck SM, von Toerne C. Direct comparison of MS-based label-free and SILAC quantitative proteome profiling strategies in primary retinal Muller cells. *Proteomics* 2012:1902–11.
- [17] Ly L, Wasinger VC. Peptide enrichment and protein fractionation using selective electrophoresis. *Proteomics* 2008;8:4197–208.
- [18] Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;227:680–5.
- [19] Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, et al. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 2003;34:374–8.
- [20] Hassan MI, Waheed A, Yadav S, Singh TP, Ahmad F. Zinc alpha 2-glycoprotein: a multidisciplinary protein. *Mol Cancer Res* 2008;6:892–906.
- [21] Pelletier CC, Koppe L, Croze ML, Kalbacher E, Vella RE, Guebre-Egziabher F, et al. White adipose tissue overproduces the lipid-mobilizing factor zinc alpha2-glycoprotein in chronic kidney disease. *Kidney Int* 2013;83:878–86.
- [22] de Ronde JJ, Lips EH, Mulder L, Vincent AD, Wesseling J, Nieuwland M, et al. SERPINA6, BEX1, AGTR1, SLC26A3, and LAPTM4B are markers of resistance to neoadjuvant chemotherapy in HER2-negative breast cancer. *Breast Cancer Res Treat* 2013;137:213–23.
- [23] Han SW, Roman J. Fibronectin induces cell proliferation and inhibits apoptosis in human bronchial epithelial cells: pro-oncogenic effects mediated by PI3-kinase and NF-kappa B. *Oncogene* 2006;25:4341–9.
- [24] Van den Steen P, Rudd PM, Dwek RA, Opdenakker G. Concepts and principles of O-linked glycosylation. *Crit Rev Biochem Mol Biol* 1998;33:151–208.
- [25] Markiewski MM, Nilsson B, Ekdahl KN, Mollnes TE, Lambris JD. Complement and coagulation: strangers or partners in crime? *Trends Immunol* 2007;28:184–92.
- [26] DuBois SG, Kalika Y, Lukens JN, Brodeur GM, Seeger RC, Atkinson JB, et al. Metastatic sites in stage IV and IVS neuroblastoma correlate with age, tumor biology, and survival. *J Pediatr Hematol Oncol* 1999;21:181–9.
- [27] Arai T, Miklossy J, Klegeris A, Guo JP, McGeer PL. Thrombin and prothrombin are expressed by neurons and glial cells and accumulate in neurofibrillary tangles in Alzheimer disease brain. *J Neuropathol Exp Neurol* 2006; 65:19–25.
- [28] Kenzelmann Broz D, Attardi LD. In vivo analysis of p53 tumor suppressor function using genetically engineered mouse models. *Carcinogenesis* 2010;31:1311–8.

- [29] Suzuki K, Matsubara H. Recent advances in p53 research and cancer treatment. *J Biomed Biotechnol* 2011;2011:978312.
- [30] Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* 2003;2:43–50.
- [31] Hakomori S. Glycosylation defining cancer malignancy: new wine in an old bottle. *Proc Natl Acad Sci U S A* 2002;99:10231–3.
- [32] Woodbury RA, Santer UV, Elkins WL, Glick MC. Similarities in glycosylation of human neuroblastoma tumors and cell lines. *Cancer Res* 1986;46:3692–7.
- [33] Janssen BJ, Huizinga EG, Raaijmakers HC, Roos A, Daha MR, Nilsson-Ekdahl K, et al. Structures of complement component C3 provide insights into the function and evolution of immunity. *Nature* 2005;437:505–11.
- [34] Gongidi V, Ring C, Moody M, Brekken R, Sage EH, Rakic P, et al. SPARC-like 1 regulates the terminal phase of radial glia-guided migration in the cerebral cortex. *Neuron* 2004;41:57–69.
- [35] Rutkowski MJ, Sughrue ME, Kane AJ, Ahn BJ, Fang S, Parsa AT. The complement cascade as a mediator of tissue growth and regeneration. *Inflamm Res* 2010;59:897–905.
- [36] Walport MJ. Complement. First of two parts. *N Engl J Med* 2001;344:1058–66.
- [37] Walport MJ. Complement. Second of two parts. *N Engl J Med* 2001;344:1140–4.
- [38] Rutkowski MJ, Sughrue ME, Kane AJ, Mills SA, Parsa AT. Cancer and the complement cascade. *Mol Cancer Res* 2010;8:1453–65.
- [39] Markiewski MM, DeAngelis RA, Benencia F, Ricklin-Lichtsteiner SK, Koutoulaki A, Gerard C, et al. Modulation of the antitumor immune response by complement. *Nat Immunol* 2008;9:1225–35.