


Identification of putative biomarkers for leptomeningeal invasion in B-cell non-Hodgkin lymphoma by NMR metabolomics

Gonçalo Graça¹ · Joana Desterro² · Joana Sousa¹ · Carlos Fonseca² · Margarida Silveira² · Jacinta Serpa^{2,3} · Tânia Carvalho⁴ · Maria G. da Silva² · Luís G. Gonçalves¹ 

Received: 9 May 2017 / Accepted: 21 September 2017 / Published online: 6 October 2017
© Springer Science+Business Media, LLC 2017

Abstract

Introduction B-cell non-Hodgkin lymphoma (B-NHL) is the most common hematological malignancy and different genetic alterations are frequently detected in transformed B lymphocytes. Within this heterogeneous disease, certain aggressive subgroups have an increased risk of central nervous system (CNS) involvement at diagnosis and/or relapse, resulting in parenchymal or leptomeningeal infiltration (LI) in 5–15% of cases. The current sensitivity limitations of cerebrospinal fluid (CSF) cytology and contrast-enhanced MRI for CNS involvement, mainly at early stages, motivates the search for alternative diagnostic methods.

Objectives Here we aim at using untargeted ¹H-NMR metabolomics to identify putative biomarkers for LI in B-NHL patients.

Methods CSF and peripheral blood samples were obtained from B-NHL patients with a positive (n=7, LI group) or negative LI diagnostic (n=13, control group). For seven

patients, CSF samples were collected during the course of intrathecal chemotherapy, making it possible to assess the patient's response to treatment. ¹H-NMR spectra were acquired and statistical multivariate and univariate analysis were performed to identify significant alterations.

Results Significant metabolite differences were found between LI and control groups in CSF, but not in serum. A predictive PLS-DA cross-validated model identified significant pool changes in glycine, alanine, pyruvate, acetylcarnitine, carnitine, and phenylalanine. Additionally, increments in protein signals were detected in the LI group. Significantly, the PLS-DA model predicted correctly all samples obtained from the group of patients in remission during LI treatment.

Conclusions The results show that the CSF NMR-metabolomics approach is a promising complementary method in clinical diagnosis and treatment follow-up of LI in B-NHL patients.

Electronic supplementary material The online version of this article (doi:10.1007/s11306-017-1269-9) contains supplementary material, which is available to authorized users.

✉ Luís G. Gonçalves
lgafeira@itqb.unl.pt

¹ ITQB NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av da República, 2780-157 Oeiras, Portugal

² IPOLFG, Instituto Português de Oncologia de Lisboa Francisco Gentil, Lisbon, Portugal

³ CEDOC, Centro de Investigação em Doenças Crónicas, NOVA Medical School, Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisbon, Portugal

⁴ IMM, Instituto de Medicina Molecular, Universidade de Lisboa, Lisbon, Portugal

Keywords NMR metabolomics · B cell non-Hodgkin lymphoma · Leptomeningeal infiltration · Cerebrospinal fluid · Serum

1 Introduction

B-cell non-Hodgkin lymphoma (B-NHL) arises from aberrant DNA mutations taking place during the development of B-lymphocytes in the lymphoid system (Nogai et al. 2011). Systemic chemotherapeutic treatment enables survival of more than half of the patients within this heterogeneous entity, with about 40% aggressive cases showing relapse and refractory disease (Stenson et al. 2016). Since the current standard chemotherapy agents fail to penetrate the blood–brain barrier, some patients affected by the

malignancy ultimately develop parenchymal and leptomeningeal infiltration (LI), both associated with poor prognosis. Central Nervous System (CNS) relapse may result from either occult disease present at time of initial diagnosis or it may develop due to a later acquisition of CNS-penetrating subtypes of malignant clones and occurs in 5–15% of all aggressive, non Burkitt B-NHL cases (Chamberlain 2008; Demopoulos 2004). Several risk factors for the development of LI in B-NHL patients have been identified, such as high International Prognostic Index (IPI) (lymphoma subtype and stage), extent and type of extranodal disease, site-specific disease, such as testes, breast, paranasal or parameningeal, age, performance status and elevated serum lactate dehydrogenase (LDH) levels (Nolan and Abrey 2003; Schmitz et al. 2016).

The final diagnosis of LI usually involves the cytological evaluation of cerebrospinal fluid (CSF) and flow-cytometry to detect the presence of lymphoma cells, in parallel with gadolinium-enhanced MRI scans, and the search for CSF biochemical markers (lactate and glucose), and tumor-associated protein markers (Cavaliere et al. 2012; Nolan and Abrey 2003). Although essential to the establishment of a definitive diagnosis, all these procedures have drawbacks. For instance, the rate of positive results in the first CSF cytology can be as low as 71%, thus requiring a series of invasive collections of CSF until a conclusive diagnosis is attained (Lee 2015). MRI can also lead to 30–70% false negative results (Cavaliere et al. 2012). Therefore, more sensitive and less-invasive diagnostic procedures are necessary so that more informed decisions can be made and interventions can be performed at the early stages of the disease.

Metabolomics has been proposed as a robust methodology to look for non-cellular biomarkers of CNS invasion, owing to its untargeted nature and better sensitivity compared to methods such as flow cytometry (Huang and Ouyang 2013; Weston et al. 2011). In particular, $^1\text{H-NMR}$ metabolomics has been extensively used in the search for disease biomarkers due to the simplicity of sample preparation, non-destructive nature, and ability to detect and quantify hundreds of metabolites simultaneously in complex samples like blood serum, plasma, urine and CSF (Bollard et al. 2005; Lindon et al. 1999).

CSF is the biofluid of choice to probe the CNS condition despite the invasive procedure associated with its collection. Indeed, several CSF $^1\text{H-NMR}$ metabolomics studies have addressed a variety of neurological diseases such as multiple sclerosis (Lutz and Cozzone 2011; Sinclair et al. 2010), neuron-motor diseases (Blasco et al. 2010), Parkinson's disease (Öhman and Forsgren 2015), Alzheimer's disease (Kork et al. 2009), complex regional pain (Meissner et al. 2013), several infections including meningitis (Subramanian et al. 2005), and rabies (O'Sullivan et al. 2013). Moreover, NMR metabolomics has been used to characterize biopsies

of CNS metastasis from humans as well as CSF samples from a LI-rat model (Cho et al. 2012; Sjøbakk et al. 2013). Recently, An et al. 2015 reported the use of CSF $^1\text{H-NMR}$ metabolomics to the diagnosis of LI in lung adenocarcinoma patients by comparing the metabolite composition of CSF from this group of patients with CSF from matched non-cancer patients (An et al. 2015). However, an improved experimental design should include as controls cancer patients without LI and, to our knowledge such a study has not been undertaken.

Blood serum/plasma $^1\text{H-NMR}$ metabolomics analysis has been extensively applied to study several hematologic malignancies, in an effort to put forward further diagnostic and prognostic biomarkers for diseases such as acute myeloid leukemia (AML) (Carrabba et al. 2016; Wang et al. 2013), pediatric acute lymphoblastic leukemia (ALL) (Tiziani et al. 2013), multiple-myeloma (Puchades-Carrasco et al. 2013; Lodi et al. 2013), chronic lymphocytic leukaemia (CLL) (MacIntyre et al. 2010) and diffuse large B-cell lymphoma (Stenson et al. 2016). In contrast, the use blood serum or plasma, are largely under-explored as source of biomarkers for LI in aggressive B-NHL lymphomas.

In this study we used $^1\text{H-NMR}$ metabolomics to characterize CSF and serum from aggressive B-NHL patients in order to identify biomarkers for LI by analyzing the CSF metabolite profiles of cancer patients with and without LI.

2 Methods

2.1 Ethics statement

This study was reviewed and approved by the ethical committee of the Portuguese Oncology Institute Francisco Gentil, Lisbon (Approval Number: GIC/733 + UIC/660) and performed in accordance with the 1964 Helsinki declaration and its later amendments.

2.2 Patient information

Patients diagnosed with aggressive non-Burkitt B-cell non-Hodgkin lymphoma (diffuse large B-cell and “double hit” cases), with medical indication to collect CSF for LI diagnosis were included in this study. The cohort included adult male and female patients (Table 1), all at an advanced stage of the disease (stage IV), with involvement of more than one extranodal site and under systemic therapy. LI positive cases were diagnosed by the presence of malignant cells in CSF by flow-cytometry. Cancer patients who were negative for CSF cell invasion (LI negative) were used as controls. At the time of collection all patients were under treatment with the current standard therapy based on immune-chemotherapy (RCHOP).

Table 1 Patient characteristics and respective CSF clinical biochemistry data

	Control	LI	<i>p</i> value
N	13	7	–
Age	23–86	32–68	<i>NS</i> *
Gender (M/F)	5/8	4/3	–
Samples			
CSF	13	5	–
Serum	12	6	–
CSF and Serum (matched)	11	4	–
CSF biochemistry median (range)			
Glucose (mg/dL)	67 (52–122)	58 (21–59)	0.009
LDH (Unit/dL)	18 (11–40)	163 (20–295)	0.026
Total protein (mg/dL)	37 (25–62)	212 (65–603)	0.002

**NS* not-significant (*p* value > 0.05)

In total, 20 patients were included, 13 without (control group), and 7 with lymphoma cells identified in CSF (LI group). The ¹H-NMR analysis of CSF was possible only for 5 positive LI cases due to limitations in the sample volumes of two patients. Serum samples were not available for one LI-positive case and two LI-negative (control) cases. CSF follow-up samples, i.e., samples collected from the same patient over the course of intrathecal chemotherapy with methotrexate and corticosteroids either as prophylactic (5 LI-negative cancer patients considered high risk for CNS infiltration), or therapeutic treatment against LI (2 LI-positive patients), were obtained. Two to 7 samples were collected from each subject during the course of treatment.

2.3 Sample collection

The CSF samples were collected by lumbar puncture from non-fasting patients. After collection, the CSF samples were centrifuged to separate the cells for cytological evaluation. The volume remaining after routine biochemical analysis was stored at –80 °C until NMR analysis. Samples contaminated with blood and those containing less than 300 μL were discarded. Blood samples were collected in the same day as the CSF samples. After collection, the blood serum tubes were centrifuged and the serum stored at –80 °C until analysis.

2.4 NMR spectral acquisition, processing, and metabolite identification

Prior to NMR analysis, the samples were thawed at room temperature. Three hundred microliter of CSF were mixed with 300 μL of potassium phosphate buffer 70 mM at pH 7.0 in 99.8% D₂O containing 2 mM NaN₃; then 560 μL were transferred to 5 mm NMR tubes. All experiments were

performed at 25 °C in a Bruker Avance II+ spectrometer, operating at a frequency of 800.33 MHz for ¹H, equipped with a 5 mm three channel probe (TXI-Z H/C/N/-D). For each CSF sample, one-dimensional ¹H-NMR spectra were acquired using the NOESY1D pulse sequence with optimized water presaturation (noesygp1d), as a sum of 128 free induction decays, with 128 k complex points, using a spectral window of 20 ppm (16025.64 Hz), 4 s relaxation delay and 10 ms mixing time.

Serum samples were prepared by mixing 300 μL of each sample with 300 μL of a sodium phosphate buffer solution prepared in 80% H₂O and 20% D₂O; then 560 μL were transferred to a 5 mm NMR tube. Serum samples were analyzed in a Bruker Avance III spectrometer operating at the frequency of 600.10 MHz for ¹H, equipped with a 5 mm four channel cryo-probe (QCI-Z H/C/P/N/-D) and a refrigerated autosampler at 25 °C. For each sample a 1D ¹H-NMR spectrum was acquired using a CPMG pulse sequence with water presaturation (cpmgpr1d) as a sum of 32 free induction decays, with 64 k complex points, using a spectral window of 20 ppm (12019.23 Hz), 4 s relaxation delay, an echo time of 0.3 ms and 126 repetitions of the echo time per scan.

Spectral processing was carried out using the TopSpin software version 3.2 (Bruker BioSpin, Rheinstetten, Germany). Briefly, each free induction decay was multiplied by an exponential function to produce a line broadening of 0.3 Hz prior to Fourier transformation. The resulting spectra were manually phased and baseline corrected. The chemical shifts were referenced internally to the anomeric proton signal of α-D-glucose at 5.22 ppm. Two-dimensional NMR spectroscopy experiments and selective one-dimensional experiments were acquired to assist metabolite identification (spectral assignments): 2D ¹H *J*-resolved spectra were acquired for each sample and homonuclear (¹H-¹H) total correlation spectroscopy (TOCSY), (¹H-¹³C) heteronuclear single quantum correlation spectroscopy (HSQC) and 1D-selective-TOCSY experiments were carried out for selected samples. For confirmation purposes the identified peaks were compared with NMR spectra from pure standards available in the Human Metabolome Database (HMDB) (Wishart et al. 2013) or compared with spectra from selected CSF samples spiked with pure compounds.

2.5 Statistical data analysis

The spectral intensities of the 1D NMR spectra were converted to rectangular matrices where each line corresponded to one CSF or serum spectrum and each column to a chemical shift. 1D spectra were subjected to inter-sample chemical shift alignment using the CluPA algorithm (Vu et al. 2011), particularly in the regions corresponding to protons near charged groups of organic acids and amino acids, for which the chemical shift position is more variable. The regions

affected by water presaturation (the HOD signal, at around 4.70 ppm and urea, at around 5.76 ppm), and the regions containing noise only (-5 to -1 ppm and 15 to 10 ppm) were excluded from the analysis. The dataset was log-transformed to place high and low intensity peaks on a comparable scale. Unsupervised analysis was carried out using Principal Component Analysis (PCA). Supervised analysis of controls and LI spectra was performed by Partial Least Squares-Discriminant Analysis (PLS-DA). The predictive ability of the PLS-DA models was assessed by predicting 10 subsets of the dataset on the model created with the remaining data (tenfold cross-validation). The goodness of prediction (Q^2) value of the resulting models was used to evaluate the cross-validation procedure. A permutation analysis was performed in order to further confirm the validity of the PLS-DA model (Westerhuis et al. 2008). The permutation analysis consisted in the repetition of the PLS-DA model 100 times by randomly assigning the group labels (control or LI), for each new model and calculating the Q^2 . One hundred non-permuted models (true models) were also obtained by randomly changing the sample order in the original spectra matrix (without changing the group labels), to exclude any possibility of over-fitting. The spectral variables responsible for the differences between control and LI groups were identified from the loading weights and variable importance

to the projection (VIP) values of the cross-validated PLS-DA models.

The significance of the metabolite differences were further evaluated by performing Mann-Whitney-Wilcoxon non-parametric tests on the spectral areas of the corresponding metabolite peaks after local baseline correction; the obtained p -values were false discovery rate (FDR) adjusted (q -values) (Benjamini and Hochberg 1995). All statistical analyses and data processing were performed using in-house scripts and the PLS package (Mevik and Wehrens 2007) in R statistical software version 3.1.3 (<http://www.r-project.org>).

3 Results

3.1 Characterization of the differences between control and LI-positive groups

Typical $^1\text{H-NMR}$ spectra, and representative assignments, of CSF and serum from B-NHL patients are shown in Fig. 1a, b, respectively. The qualitative analysis of the spectra from both biofluids enabled the identification of 56 metabolites and characteristic protein signals in CSF and about 40 metabolites and signals characteristic of lipids from lipoproteins and glycoproteins in serum. The

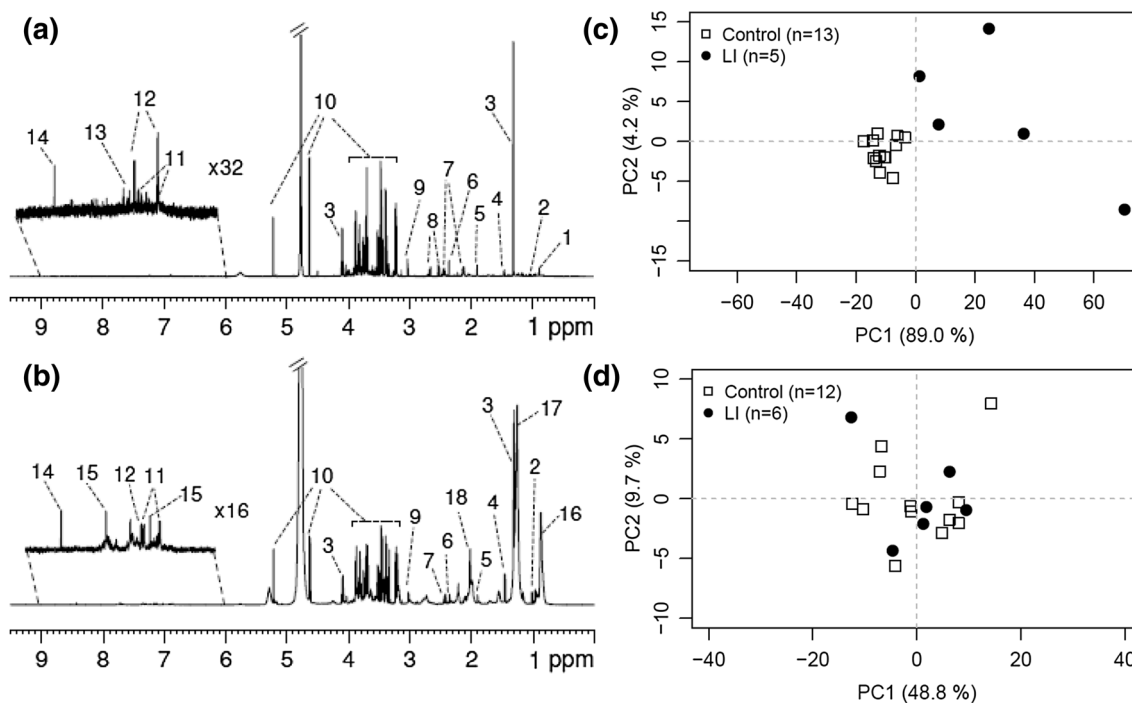


Fig. 1 Analysis of $^1\text{H-NMR}$ spectra from B-NHL patients with and without LI (control). **a** Representative CSF noesygppr1d spectrum from a CSF sample; **b** representative cpmgpr1d spectrum from a serum sample. Legend-1: 2-hydroxybutyrate, 2: isoleucine, leucine, valine, 3: lactate, 4: alanine, 5: acetate, 6: pyruvate, 7: glutamine, 8:

citrate, 9: creatine, creatinine, 10: glucose, 11: tyrosine, 12: acetaminophen, 13: phenylalanine, 14: formate, 15: histidine, 16: lipoprotein- CH_3 , 17: lipoprotein- CH_2 , 18: glycoprotein N -acetyl groups; **c** PCA scores scatter plot of CSF spectra; **d** PCA scores scatter plot of serum spectra

majority of the assigned metabolites comprised amino and organic acids, which were present in both fluids. A comprehensive list of assignments is provided in Supplementary Table S1.

To evaluate the differences between control and LI groups, both the CSF and serum $^1\text{H-NMR}$ spectra were analyzed through PCA (Fig. 1). A good separation between control and LI groups was observed in the PCA scores plot (Fig. 1c), thus indicating significant metabolite differences in CSF samples. On the other hand, the serum samples led to an overlap in the corresponding scores plot (Fig. 1d), therefore reflecting no significant differences in terms of serum composition between control and LI groups.

The same set of LI and control CSF spectra was further examined by means of a supervised multivariate statistical method, PLS-DA. This method was chosen in order to allow the construction of a predictive multivariate model based on CSF $^1\text{H-NMR}$ spectra while simultaneously enabling the identification of the metabolites with greatest impact on the discrimination between control and LI groups. The robustness of the PLS-DA model was evaluated by cross-validation. By using tenfold cross-validation procedure the model presented a goodness-of-prediction parameter (Q^2) of 0.63, already indicating a good predictive ability. Further validation was performed by permutation analysis. The distributions of Q^2 for the permuted and true PLS-DA models (Supplementary Fig. S1) clearly showed that the permuted models span a wide range of negative Q^2 values with a median value of -0.67 , whereas the true models Q^2 values result in a skewed distribution of only positive Q^2 values with a median of 0.63.

The scores plot of the resulting PLS-DA model (Fig. 2a), shows a similar separation between groups when compared with the corresponding PCA analysis. However, the validated PLS-DA model enabled the identification of the metabolite resonances which are responsible for the separation of control and LI groups through the inspection of the model weight loadings and VIP parameter. Only the signals corresponding to VIP values > 1 were considered as having a significant impact on group separation. This criterion was fulfilled for 2-hydroxyisovalerate, alanine, acetone, acetylcarnitine, betaine, carnitine, formate, glycine, histidine, isoleucine, lactate, leucine, lysine, methionine, phenylalanine, proline, pyruvate, tyrosine, valine, 2-hydroxybutyrate, 2-oxoisovalerate and an unassigned metabolite, as well as broad protein signals (Fig. 2b; Table 2). Moreover, the corresponding loading weights revealed that all metabolites were augmented in the LI group as compared to controls, except for 2-hydroxybutyrate, 2-oxoisovalerate and the unknown metabolite whose levels were decreased. The loading weights of glucose signals indicate a decrease of this metabolite, in accordance with clinical biochemical data (Table 1),

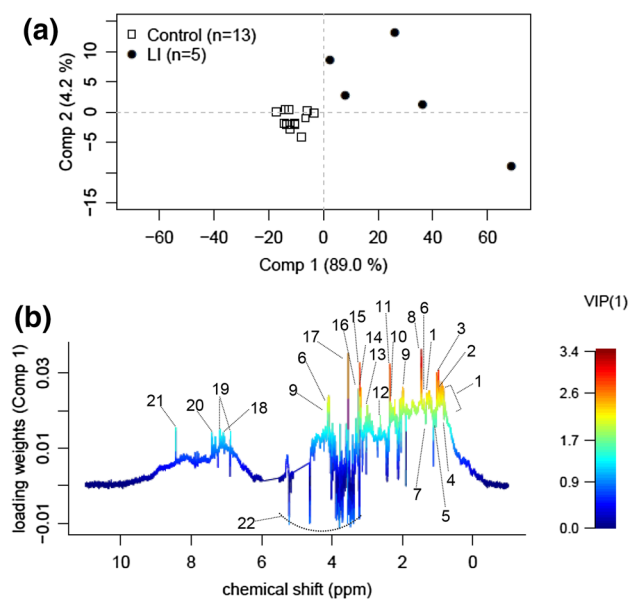


Fig. 2 PLS-DA model of log-transformed $^1\text{H-NMR}$ spectra of CSF from B-NHL patients. **a** Scores scatter plot; **b** Loading weights line plot of Component 1 color-coded with the corresponding variable importance to the model (VIP value). Legend-1: $-\text{CH}_3$ signals assigned to proteins, 2: 2-hydroxyisovalerate, 3: isoleucine, leucine and valine, 4: 2-hydroxybutyrate, 5: 2-oxoisovalerate, 6: lactate, 7: unassigned, 8: alanine, 9: proline, 10: acetone, 11: pyruvate, 12: methionine, 13: lysine, 14: acetylcarnitine, 15: carnitine, 16: betaine, 17: glycine, 18: histidine, 19: tyrosine, 20: phenylalanine, 21: formate, 22: glucose

however it has negligible impact on the separation apparent in the PLS-DA model, i.e., $\text{VIP} < 1$ (Fig. 2b; Table 2).

The statistical significance of the metabolite alterations was further assessed by comparing the corresponding spectral areas using non-parametric univariate tests adjusted for multiple testing (Table 2). Glycine, alanine, pyruvate, acetylcarnitine, carnitine, and phenylalanine as well as the protein methyl signals in the CSF spectra (all increased in the LI group) were significantly altered between control and LI-positive samples (Table 2). In brief, this analysis led to the following final set of significantly altered metabolites: glycine, alanine, pyruvate, carnitine, acetylcarnitine, and phenylalanine.

3.2 Evaluation of CSF metabolic differences during the course of treatment for LI

As part of the clinical procedure, all B-NHL patients for whom there is suspicion at diagnosis of CNS invasion are subjected to intrathecal chemotherapy in addition to concomitant systemic chemotherapy. Some of the CSF samples collected over the course of treatment were not analysed by NMR due to volume limitation ($< 300 \mu\text{L}$), however it was

Table 2 CSF metabolites relevant for the discrimination between control and LI-positive groups

Metabolite	Signals (ppm)	VIP (Comp 1) ^a	Integrals (mean ± SD)		Fold change	<i>q</i> -value ^{b,c}
			Control (n = 13)	LI (n = 5)		
Alanine	1.47	3.4	1020 ± 284	5687 ± 4162	6	0.005
Glycine	3.55	3.4	87 ± 37	1147 ± 766	13	0.005
Carnitine	3.22	3.0	80 ± 19	1053 ± 957	13	0.037
Pyruvate	2.36	3.0	789 ± 560	2832 ± 1362	4	0.021
Valine	0.98	2.8	393 ± 100	1902 ± 1312	5	NS
2-Hydroxybutyrate	0.89	2.6	158 ± 112	401 ± 282	3	NS
Acetylcarnitine	3.18	2.5	48 ± 14	347 ± 270	7	0.021
Betaine	3.25	2.5	26 ± 11	183 ± 171	7	NS
Isoleucine	0.93, 1.00	2.5	156 ± 43	553 ± 418	4	NS
Protein (–CH ₃)	0.85	2.4	1307 ± 360	7379 ± 5774	6	0.005
Leucine	0.95	2.4	199 ± 142	524 ± 392	3	NS
Lactate	1.32	2.4	9794 ± 2390	26,191 ± 15,873	3	NS
Acetone	2.21	2.1	109 ± 56	370 ± 475	3	NS
Proline	4.12	2.1	21 ± 5	124 ± 115	6	NS
Lysine	3.01	2.0	149 ± 37	419 ± 259	3	NS
2-Hydroxyisovalerate	0.82, 0.95	1.8	191 ± 268	145 ± 55	–1	NS
Methionine	2.63	1.8	33 ± 8	68 ± 36	2	NS
Unassigned	1.35	1.6	50 ± 40	43 ± 21	–1	NS
2-Oxoisovalerate	1.06	1.5	308 ± 123	434 ± 194	1	NS
Formate	8.45	1.4	107 ± 36	255 ± 114	3	NS
Tyrosine	6.89, 7.18	1.4	219 ± 33	505 ± 317	2	NS
Histidine	7.08	1.3	211 ± 35	366 ± 161	2	NS
Phenylalanine	7.32, 7.36, 7.42	1.3	290 ± 59	625 ± 233	2	0.021

^aVIP (Comp 1), variable importance for discrimination of separating component 1

^b*q*-value from the statistical comparison between LI and control groups (FDR adjusted *p*-value)

^cNS not-significant if adjusted *p*-value > 0.05

possible to analyse at least two data points for 5 controls and 2 LI-positive patients.

In order to evaluate the changes in the CSF profile that occurred during treatment, the spectra from one patient without LI (patient 1, with 3 time points over treatment), and one patient diagnosed with LI (patient 2, with 6 time point collections), were projected in the previously obtained PLS-DA model (Fig. 3). The CSF samples collected from patients 1 and 2, immediately before starting intrathecal chemotherapy, are indicated with arrows (Fig. 3). Interestingly, the CSF samples from patient 2 over the course of treatment are projected near the cloud of controls, and this result correlates with the absence of malignant cells in the cytological analysis of the respective CSF samples.

Further evaluation of the CSF composition over the course of treatment was performed focusing on the metabolite and protein signals with highest significance for group discrimination. For the sake of clarity, only the data relative to patients 1 and 2 are shown in Fig. 4, but similar trends were observed for the other patients followed-up during the

course of treatment (Supplementary Fig. S2). For the LI-positive subject (patient 2), alanine, glycine, pyruvate, carnitine, acetylcarnitine and the protein signals have a sharp decrease from diagnosis time (week 0) to the 3rd week of treatment. Phenylalanine showed a slower decrease over the same time period (Fig. 4). Thereafter, metabolite and protein levels remained nearly constant up to the 13th week of treatment. For the LI-negative subject (patient 1), the metabolite and protein levels remained unchanged from diagnosis up to the 11th week of treatment (Fig. 4).

4 Discussion

During the last decade ¹H-NMR-based metabolomics developed to a powerful tool for the identification of biochemical markers for a variety of human disorders and lifestyles (Dona et al. 2014; Gowda and Raftery 2015). Herein, we explored the use of this technique to find metabolic markers of CNS invasion in aggressive non-Burkitt B-NHL patients.

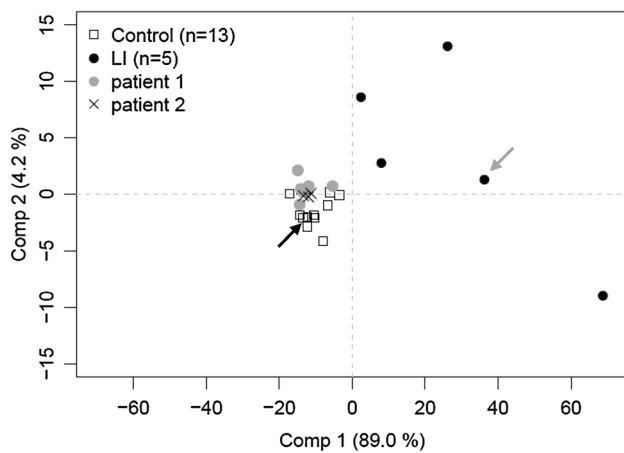


Fig. 3 Projection of CSF samples from B-NHL patients into the scores plot of the PLS-DA of $^1\text{H-NMR}$ spectra from control and LI patients. The scores from patient 1 represent subsequent CSF collections from a control patient and the scores from patient 2 represent CSF samples collected from an LI patient during the course of intrathecal treatment. The first sample from patient 1 is indicated with a black arrow; the first sample from patient 2 (before treatment) is indicated with a grey arrow

Serum and CSF samples were analyzed to probe the impact of LI on extra- and intra-CNS compartments, respectively. An unsupervised multivariate comparison between data from control and LI groups revealed that only the latter biofluid was detectably altered by the presence of LI, implying the significant contribution of malignant cell metabolism to the CSF metabolite profile (Fig. 1b). On the other hand, the lack of evidence for alterations in serum metabolites indicates the absence of detectable systemic effects connected with invasion of the CNS, in agreement with the “sanctuary role” of this location (Basu et al. 2014).

4.1 Significant metabolite alterations associated with LI

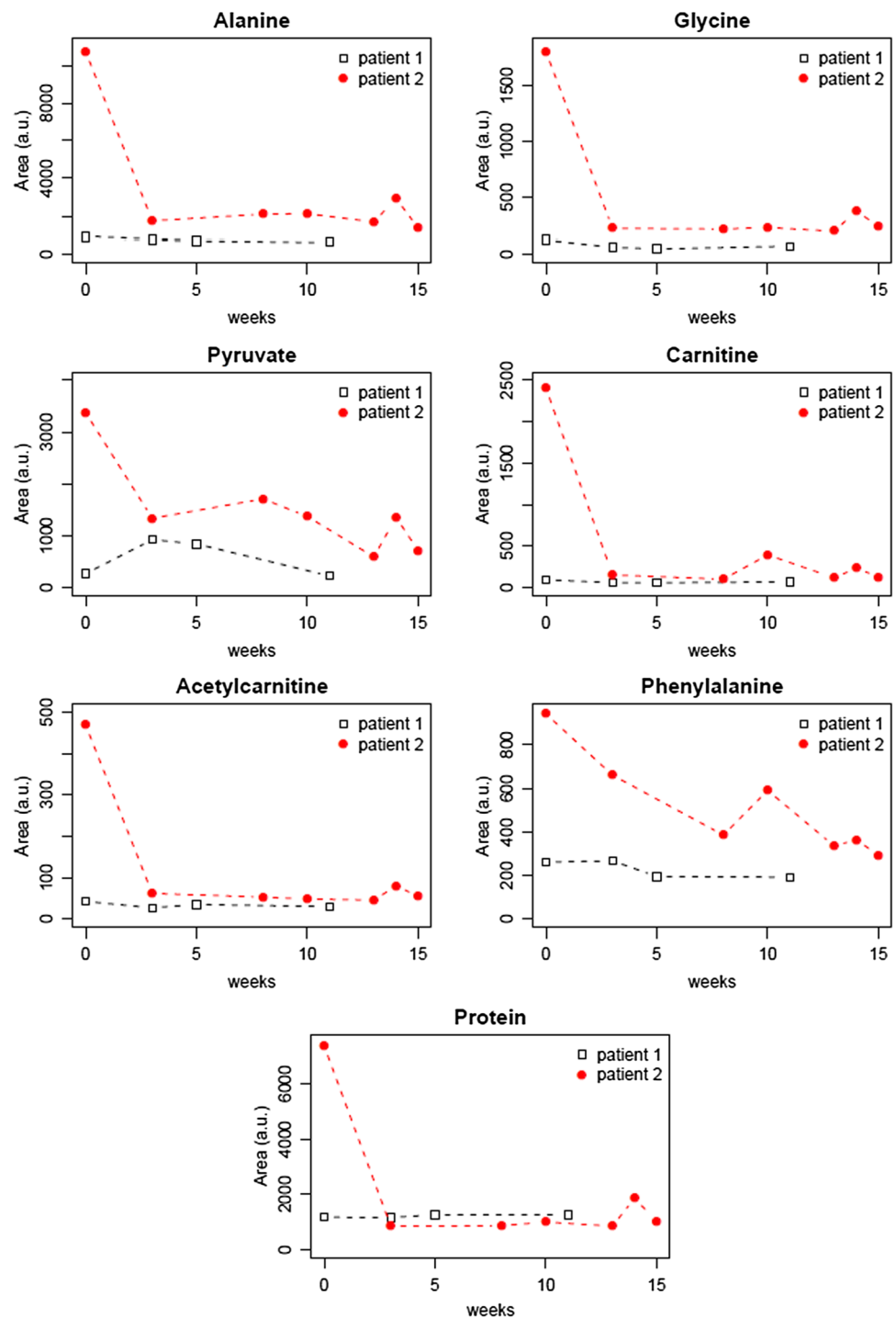
The analysis of the PLS-DA loadings and further validation by univariate analysis enabled the identification of the metabolite and protein signals which were significantly altered in the CSF samples of the LI group. Among these metabolites are several amino acids (alanine, phenylalanine and glycine), pyruvate, carnitine and acetylcarnitine, which increased in comparison with the control group (Table 2). These alterations reflect to some extent the metabolic activity of lymphoma cells in the CSF and/or in the leptomeninges of the subjects with CNS lymphoma infiltration. The increase in the pool of pyruvate is ascribed to the Warburg effect, commonly observed in cancer cells and tumors, which implies an enhanced aerobic glycolytic flux (Griffin and Shockcor 2004; Vermeersch and Styczynski 2013). However, it is worth noting that the alterations

in the pools of glucose and lactate depicted in Fig. 2 are not significant and these metabolites had no impact on the separation depicted in the PLS-DA model (Fig. 2; Table 2). On the other hand, the higher level of alanine observed in LI-diagnosed patients can be ascribed to a shift in the flux from pyruvate to the production of alanine via pyruvate transamination (Griffin and Shockcor 2004).

The amino acids glycine, alanine, and phenylalanine, as well as the protein signals, were significantly increased in the LI group. This alteration may denote disruption in the blood–brain barrier permeability, a common feature during tumor CNS infiltration (Chamberlain et al. 2014). Elevated protein levels were also detected in the routine biochemical analysis of CSF samples from LI-diagnosed patients (Table 1). Special notice should be given to glycine which is among the metabolites with greatest influence on the discrimination between LI and control groups. Glycine is derived from serine, which in turn is synthesized from 3-phosphoglycerate, a glycolytic metabolite. Glycine and serine are the primary donors of one-carbon units for the folate and methionine cycles that generate building blocks for the synthesis of lipids, proteins, nucleotides and glutathione, components essential for cell growth and proliferation (Li and Zhang 2015). It is presently recognized that serine and glycine metabolism play a central role in cancer pathogenesis (Locasale 2013 and references therein). Therefore, we propose that the significantly incremented glycine levels in the LI group represent an important putative diagnostic and prognostic biomarker. In the brain, glycine is an inhibitory neurotransmitter and elevated levels have been associated with brain tumors (Kinoshita and Yokota 1997). Furthermore, increased glycine levels have been correlated with tumor aggressiveness and progression in triple negative breast cancer (Cao et al. 2014), in locally advanced rectal cancer (Redalen et al. 2016) or aggressive prostate cancer (McDunn et al. 2013).

The increased levels of carnitine and acetylcarnitine in the LI group are probably associated with disturbed lipid metabolism, which is one of the hallmarks of cancer cell metabolism (Vermeersch and Styczynski 2013). For example, lipid accumulation was observed in the cytoplasm of metastatic cancer cells from solid brain tumors (Sjøbakk et al. 2013); and fatty acid synthesis was up-regulated in an in-vitro model of a B-cell type lymphoma (primary effusion lymphoma) (Bhatt et al. 2012). Carnitine acts as a transporter of acetyl and acyl groups between the cytosol and the mitochondria. Acetyl groups are transported via acetylcarnitine from the mitochondria to the cytosol where they can be used for fatty acid synthesis (Bhatt et al. 2012; Pacilli et al. 2013). We speculate that the higher levels of carnitine and acetylcarnitine are related with enhanced lipid biosynthesis in malignant cells.

Fig. 4 Variation of CSF metabolite levels at diagnosis and during intrathecal treatment of a control (patient 1) and an LI patient (patient 2)



4.2 CSF alterations during the course of LI treatment

The availability of several samples from the same patient collected over the course of intrathecal therapy enabled the evaluation of the metabolite changes connected with the chemotherapeutic intervention. After 3 weeks of treatment there were differences in the CSF metabolic profile of LI positive patients, but not in the profile of control patients.

In fact, the levels of the metabolites with greatest discriminatory impact, i.e., glycine, alanine, pyruvate, acetylcarnitine, carnitine, and phenylalanine, decreased after treatment initiation in those cases. These changes correlate with the absence of malignant cells in CSF of those patients after treatment, as evaluated by cytological examination.

In spite of the pathophysiology differences between non-Hodgkin lymphoma and other hematologic malignancies, it

is interesting to note that the levels of some of the metabolites decreasing during LI remission, seem to correlate with remission in AML, ALL and multiple myeloma patients. Indeed, decreases in the levels of alanine and phenylalanine have been observed in AML patients at remission (Carrabba et al. 2016) and decreases in the levels of carnitine and acetylcarnitine have been observed in multiple myeloma (Lodi et al. 2013). Therefore, the proposed biomarkers may similarly serve as useful indicators of the treatment efficacy.

4.3 Robustness of the predictive model and study limitations

One of the most important advantages of the untargeted $^1\text{H-NMR}$ approach derives from the high density of information available in a single spectrum, reflecting the signature of hundreds of metabolites in a complex mixture. The combination of $^1\text{H-NMR}$ and supervised multivariate statistical tools, such as PLS-DA, provides a powerful toolbox for biomarker identification and further translation to a reliable diagnostic model (Lindon and Nicholson 2008; Xia et al. 2013). In this work, $^1\text{H-NMR}$ spectra of CSF samples were used to construct a predictive model for LI diagnosis based on a supervised analysis (PLS-DA). Our PLS-DA model has good predictive ability as assessed by the cross-validation and permutation tests. Importantly, the goodness of the model was further attested by the correct classification of the CSF samples from patients over the course of LI therapy (Fig. 3).

We are aware of the major limitation related with a small cohort consisting of 13 control and 5 LI positive patients. This is due to the low percentage of aggressive B-NHL cases developing LI, well recognized in the literature, but also to the insufficient volume of some CSF samples after the entire routine clinical laboratory analysis was performed. Nevertheless, this study provides proof of concept of the application of $^1\text{H-NMR}$ metabolomics in the diagnosis of LI from aggressive B-NHL patients, where early markers for CNS infiltration may have a significant impact on therapy and clinical outcome. Before our findings can be translated to a clinical diagnostic test, further studies involving larger sample groups are necessary. Ideally, a CSF set of samples should include an equivalent proportion of control and LI-positive samples to enable the appropriate testing and validation of the proposed multivariate models and biomarker panel.

5 Conclusions

Leptomeningeal tumor infiltration still presents diagnostic challenges, particularly in hematologic malignancies such as the aggressive lymphomas and acute leukemia, often requiring cytological and phenotypic analysis of a series of CSF

samples collected over time for diagnostic confirmation. In this work, an NMR metabolomics strategy was applied to a small cohort of samples from aggressive B-NHL patients with and without CNS infiltration diagnosis. A multivariate model based on $^1\text{H-NMR}$ spectra of CSF was obtained which allowed the identification of a set of putative biomarkers for LI: glycine, alanine, pyruvate, carnitine, acetylcarnitine, and phenylalanine. The reliability of these metabolite biomarkers is strengthened by the observation of complete reversal of the LI metabolic traits upon intrathecal treatment. Despite the limited sample size, the results are definitely encouraging and stimulate further investigation with larger cohorts, which are essential for biomarker validation and translation into clinical application.

Acknowledgements The authors want to acknowledge Prof. Helena Santos for her support, involvement and contribution to the project. The NMR data was acquired at CERMAX (Centro de Ressonância Magnética António Xavier) and at CICS-UBI which are members of the Portuguese NMR network.

Funding This work was supported by project PTDC/BIM-ONC/1242/2012 from Fundação para a Ciência e a Tecnologia (FCT), Portugal; project LISBOA-01-0145-FEDER-007660 (Microbiologia Molecular, Estrutural e Celular) and iNOVA4Health—UID/Multi/04462/2013 funded by FEDER through COMPETE2020—POCI and by national funds through FCT. GG and LGG were recipients of post-doc Grants, SFRH/BPD/93752/2013 and SFRH/BPD/111100/2015, awarded by FCT.

Author contributions GG, LGG, and MGS wrote the manuscript. GG, assisted by JSo, prepared samples and acquired the NMR spectra. GG with assistance of LGG performed the analysis of spectra, produced the statistical models and interpreted results. JD, CF, MGS and MS collected CSF and blood samples, supervised the biochemical analyses, and gathered the clinical data. GG, LGG, TC, JSe and MGS designed the study. All authors contributed to the revision of the manuscript.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflicts of interest.

Ethical approval This study was reviewed and approved by the ethical committee of the Portuguese Oncology Institute Francisco Gentil, Lisbon (Approval Number: GIC/733 + UIC/660) and performed in accordance with the 1964 Helsinki declaration and its later amendments.

Informed consent Serum and CSF samples were collected for routine clinical procedures and analyzed retrospectively; therefore in this study a formal consent is not required.

References

An, Y. J., Cho, H. R., Kim, T. M., Keam, B., Kim, J. W., Wen, H., et al. (2015). An NMR metabolomics approach for the diagnosis

- of leptomeningeal carcinomatosis in lung adenocarcinoma cancer patients. *International Journal of Cancer*, *136*, 162–171.
- Basu, S. K., Remick, S. C., Monga, M., & Gibson, L. F. (2014). Breaking and entering into the CNS: Clues from solid tumor and non-malignant models with relevance to hematopoietic malignancies. *Clinical and Experimental Metastasis*, *31*, 257–267.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*, 289–300.
- Bhatt, A. P., Jacobs, S. R., Freermerman, A. J., Makowski, L., Rathmell, J. C., Dittmer, D. P., & Damania, B. (2012). Dysregulation of fatty acid synthesis and glycolysis in non-Hodgkin lymphoma. *Proceedings of the National Academy of Sciences United States of America*, *109*, 11818–11823.
- Blasco, H., Corcia, P., Moreau, C., Veau, S., Fournier, C., Vour'h, P., et al. (2010). ¹H-NMR-based metabolomic profiling of CSF in early amyotrophic lateral sclerosis. *PLoS ONE*, *5*, e13223. doi:10.1371/journal.pone.0013223.
- Bollard, M. E., Stanley, E. G., Lindon, J. C., Nicholson, J. K., & Holmes, E. (2005). NMR-based metabolomic approaches for evaluating physiological influences on biofluid composition. *NMR in Biomedicine*, *18*, 143–162.
- Cao, M. D., Lamichhane, S., Lundgren, S., Bofin, A., Fjøsne, H., Giskeødegård, G. F., & Bathen, T. F. (2014). Metabolic characterization of triple negative breast cancer. *BMC cancer*, *14*, 941.
- Carrabba, M. G., Tavel, L., Oliveira, G., Forcina, A., Quilici, G., Nardelli, F., et al. (2016). Integrating a prospective pilot trial and patient-derived xenografts to trace metabolic changes associated with acute myeloid leukemia. *Journal of Hematology and Oncology*, *9*, 115. doi:10.1186/s13045-016-0346-2.
- Cavaliere, R., Savani, A., Schiff, D., & Wen, P. (2012). Nervous system metastases. In R. B. Daroff, G. M. Fenichel, J. Jankovic & J. C. Mazziotta (Eds.), *Bradley's Neurology in Clinical Practice* (6th edn., pp. 1182–1199). Philadelphia: Elsevier Saunders.
- Chamberlain, M., Soffiotti, R., Raizer, J., Rudà, R., Brandsma, D., Boogerd, W., et al. (2014). Leptomeningeal metastasis: a response assessment in neuro-oncology critical review of endpoints and response criteria of published randomized clinical trials. *Neuro-Oncology*, *16*, 1176–1185.
- Chamberlain, M. C. (2008). Neoplastic meningitis. *The Oncologist*, *13*, 967–977.
- Cho, H. R., Wen, H., Ryu, Y. J., An, Y. J., Kim, H. C., Moon, W. K., et al. (2012). An NMR metabolomics approach for the diagnosis of leptomeningeal carcinomatosis. *Cancer Research*, *72*, 5179–5187.
- Demopoulos, A. (2004). Leptomeningeal metastases. *Current Neurology and Neuroscience Reports*, *4*, 196–204.
- Dona, A. C., Jiménez, B., Schäfer, H., Humpfer, E., Spraul, M., Lewis, M. R., et al. (2014). Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Analytical Chemistry*, *86*, 9887–9894.
- Gowda, G. A., & Raftery, D. (2015). Can NMR solve some significant challenges in metabolomics? *Journal of Magnetic Resonance*, *260*, 144–160.
- Griffin, J. L., & Shockcor, J. P. (2004). Metabolic profiles of cancer cells. *Nature Reviews Cancer*, *4*, 551–561.
- Huang, Q., & Ouyang, X. (2013). Predictive biochemical-markers for the development of brain metastases from lung cancer: Clinical evidence and future directions. *Cancer Epidemiology*, *37*, 703–707.
- Kinoshita, Y., & Yokota, A. (1997). Absolute concentrations of metabolites in human brain tumors using in vitro proton magnetic resonance spectroscopy. *NMR in Biomedicine*, *10*(1), 2–12.
- Kork, F., Holthues, J., Hellweg, R., Jankowski, V., Tepel, M., Ohring, R., et al. (2009). A possible new diagnostic biomarker in early diagnosis of Alzheimers disease. *Current Alzheimer Research*, *6*, 519–524.
- Lee, E. Q. (2015). Nervous system metastases from systemic cancer. *Continuum Lifelong Learning in Neurology*, *21*, 415–428.
- Li, Z., & Zhang, H. (2015). Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. *Cellular and Molecular Life Sciences*, *73*, 377–392.
- Lindon, J. C., & Nicholson, J. K. (2008). Spectroscopic and statistical techniques for information recovery in metabolomics and metabolomics. *Annual Review of Analytical Chemistry*, *1*, 45–69.
- Lindon, J. C., Nicholson, J. K., & Everett, J. R. (1999). NMR spectroscopy of biofluids. *Annual Reports on NMR Spectroscopy*, *38*, 1–88.
- Locasale, J. W. (2013). Serine, glycine and the one-carbon cycle: cancer metabolism in full circle. *Nature Reviews Cancer*, *13*, 572–583.
- Lodi, A., Tiziani, S., Khanim, F. L., Günther, U. L., Viant, M. R., et al. (2013). Proton NMR-based metabolite analyses of archived serial paired serum and urine samples from myeloma patients at different stages of disease activity identifies acetylcarnitine as a novel marker of active disease. *PLoS ONE*, *8*(2), e56422. doi:10.1371/journal.pone.0056422.
- Lutz, N. W., & Cozzone, P. J. (2011). Metabolic profiling in multiple sclerosis and other disorders by quantitative analysis of cerebrospinal fluid using nuclear magnetic resonance spectroscopy. *Current Pharmaceutical Biotechnology*, *12*, 1016–1025.
- MacIntyre, D. A., Jiménez, B., Lewintre, E. J., Reinoso Martín, C., Schäfer, H., García Ballesteros, C., et al. (2010). Serum metabolome analysis by ¹H-NMR reveals differences between chronic lymphocytic leukaemia molecular subgroups. *Leukemia*, *24*, 788–797.
- McDunn, J. E., Li, Z., Adam, K. P., Neri, B. P., Wolfert, R. L., Milburn, M. V., et al. (2013). Metabolomic signatures of aggressive prostate cancer. *The Prostate*, *73*, 1547–1560.
- Meissner, A., van der Plas, A. a, van Dasselaar, N. T., Deelder, A. M., van Hilten, J. J., & Mayboroda, O. a. (2013). ¹H-NMR metabolic profiling of cerebrospinal fluid in patients with complex regional pain syndrome-related dystonia. *Pain*, *155*, 1–7.
- Mevik, B.-H., & Wehrens, R. (2007). The pls package: principal component and partial least squares regression in R. *Journal of Statistical Software*, *18*, 1–23.
- Nogai, H., Dörken, B., & Lenz, G. (2011). Pathogenesis of non-Hodgkin's lymphoma. *Journal of Clinical Oncology*, *29*, 1803–1811.
- Nolan, C. P., & Abrey, L. E. (2003). Leptomeningeal metastases from leukemias and lymphomas. In L. Abrey, M. Chamberlain & H. Engelhardt (Eds.), *Neurological disorders: Course and treatment* (pp. 897–909). New York: Springer.
- O'Sullivan, A., Willoughby, R. E., Mishchuk, D., Alcarraz, B., Cabezas-Sanchez, C., Condori, R. E., et al. (2013). Metabolomics of cerebrospinal fluid from humans treated for rabies. *Journal of Proteome Research*, *12*, 481–490.
- Öhman, A., & Forsgren, L. (2015). NMR metabolomics of cerebrospinal fluid distinguishes between Parkinson's disease and controls. *Neuroscience Letters*, *594*, 36–39.
- Pacilli, A., Calienni, M., Margarucci, S., D'Apolito, M., Petillo, O., Rocchi, L., et al. (2013). Carnitine-acyltransferase system inhibition, cancer cell death, and prevention of myc-induced lymphomagenesis. *Journal of the National Cancer Institute*, *105*, 489–498.
- Puchades-Carrasco, L., Lecumberri, R., Martínez-López, J., Lahuerta, J.-J., Mateos, M.-V., Prósper, F., et al. (2013). Multiple myeloma patients have a specific serum metabolomic profile that changes after achieving complete remission. *Clinical Cancer Research*, *19*, 4770–4779.
- Redalen, K. R., Sitter, B., Bathen, T. F., Grøholt, K. K., Hole, K. H., Dueland, S., et al. (2016). High tumor glycine concentration is

- an adverse prognostic factor in locally advanced rectal cancer. *Radiotherapy and Oncology*, 118, 393–398.
- Schmitz, N., Zeynalova, S., Nickelsen, M., Kansara, R., Villa, D., Sehn, L. H., et al. (2016). CNS International prognostic index: A risk model for CNS relapse in patients with diffuse large B-cell lymphoma treated with R-CHOP. *Journal of Clinical Oncology*, 34, 3150–3156.
- Sinclair, A. J., Viant, M. R., Ball, A. K., Burdon, M. A., Walker, E. A., Stewart, P. M., et al. (2010). NMR-based metabolomic analysis of cerebrospinal fluid and serum in neurological diseases—A diagnostic tool? *NMR in Biomedicine*, 23, 123–132.
- Sjøbakk, T. E., Vettukattil, R., Gulati, M., Gulati, S., Lundgren, S., Gribbestad, I. S., et al. (2013). Metabolic profiles of brain metastases. *International Journal of Molecular Sciences*, 14, 2104–2118.
- Stenson, M., Pedersen, A., Hasselblom, S., Nilsson-Ehle, H., Karlsson, B. G., Pinto, R., & Andersson, P.-O. (2016). Serum nuclear magnetic resonance-based metabolomics and outcome in diffuse large B-cell lymphoma patients—A pilot study. *Leukemia and Lymphoma*, 57, 1814–1822.
- Subramanian, A., Gupta, A., Saxena, S., Gupta, A., Kumar, R., Nigam, A., et al. (2005). Proton MR CSF analysis and a new software as predictors for the differentiation of meningitis in children. *NMR in Biomedicine*, 18, 213–225.
- Tiziani, S., Kang, Y., Harjanto, R., Axelrod, J., Piermarocchi, C., et al. (2013). Metabolomics of the tumor microenvironment in pediatric acute lymphoblastic leukemia. *PLoS ONE*, 8(12), e82859. doi:10.1371/journal.pone.0082859.
- Vermeersch, K. a, & Styczynski, M. P. (2013). Applications of metabolomics in cancer research. *Journal of Carcinogenesis*, 12, 9.
- Vu, T. N., Valkenburg, D., Smets, K., Verwaest, K. A., Dommissie, R., Lemièr, F., et al. (2011). An integrated workflow for robust alignment and simplified quantitative analysis of NMR spectrometry data. *BMC Bioinformatics*, 12, 405.
- Wang, Y., Zhang, L., Chen, W.-L., Wang, J.-H., Li, N., Li, J.-M., et al. (2013). Rapid diagnosis and prognosis of de novo acute myeloid leukemia by serum metabolomic analysis. *Journal of Proteome Research*, 12, 4393–4401.
- Westerhuis, J., Hoefsloot, H., Smit, S., Vis, D., Smilde, A., van Velzen, E., et al. (2008). Assessment of PLS-DA cross validation. *Metabolomics*, 4, 81–89.
- Weston, C. L., Glantz, M. J., & Connor, J. R. (2011). Detection of cancer cells in the cerebrospinal fluid: current methods and future directions. *Fluids and Barriers of the CNS*, 8, 14. doi:10.1186/2045-8118-8-14.
- Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., et al. (2013). HMDB 3.0—The human metabolome database in 2013. *Nucleic Acids Research*, 41(Database issue), D801-7.
- Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics*, 9, 280–299.