Carcinogenesis vol.36 no.1 pp.68–75, 2015
doi:10.1093/carcin/bgu226
Advance Access publication November 3, 2014

NMR metabolomics of human lung tumours reveals distinct metabolic signatures for adenocarcinoma and squamous cell carcinoma

Cláudia M.Rocha, António S.Barros[1], Brian J.Goodfellow, Isabel M.Carreira[2,3], Ana Gomes[3,4], Vitor Sousa[3,4,5], João Bernardo[3,5,6], Lina Carvalho[3,4,5], Ana M.Gil and Iola F.Duarte*

CICECO, Department of Chemistry, University of Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal, [1]QOPNA Research Unit, Department of Chemistry, University of Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal, [2]Laboratory of Cytogenetics and Genomics, Faculty of Medicine, University of Coimbra, Azinhaga de Santa Comba, 3000–548 Coimbra, Portugal, [3]Centre for Investigation in Environment, Genetics and Oncobiology (CIMAGO), Faculty of Medicine, University of Coimbra, 3004–504 Coimbra, Portugal, [4]Department of Pathological Anatomy, University Hospitals of Coimbra, 3000–075 Coimbra, Portugal, [5]Institute of Pathological Anatomy, Faculty of Medicine, University of Coimbra, Rua Larga, 3004–504 Coimbra, Portugal and [6]Cardiothoracic Surgery, University Hospitals of Coimbra, 3000-075 Coimbra, Portugal

*To whom correspondence should be addressed. Tel: +351 234401424; Fax: +351 234401470;
Email: ioladuarte@ua.pt

Lung tumour subtyping, particularly the distinction between adenocarcinoma (AdC) and squamous cell carcinoma (SqCC), is a critical diagnostic requirement. In this work, the metabolic signatures of lung carcinomas were investigated through ^{1}H NMR metabolomics, with a view to provide additional criteria for improved diagnosis and treatment planning. High Resolution Magic Angle Spinning Nuclear Magnetic Resonance (NMR) spectroscopy was used to analyse matched tumour and adjacent control tissues from 56 patients undergoing surgical excision of primary lung carcinomas. Multivariate modeling allowed tumour and control tissues to be discriminated with high accuracy (97% classification rate), mainly due to significant differences in the levels of 13 metabolites. Notably, the magnitude of those differences were clearly distinct for AdC and SqCC: major alterations in AdC were related to phospholipid metabolism (increased phosphocholine, glycerophosphocholine and phosphoethanolamine, together with decreased acetate) and protein catabolism (increased peptide moieties), whereas SqCC had stronger glycolytic and glutaminolytic profiles (negatively correlated variations in glucose and lactate and positively correlated increases in glutamate and alanine). Other tumour metabolic features were increased creatine, glutathione, taurine and uridine nucleotides, the first two being especially prominent in SqCC and the latter in AdC. Furthermore, multivariate analysis of AdC and SqCC profiles allowed their discrimination with a 94% classification rate, thus showing great potential for aiding lung tumours subtyping. Overall, this study has provided new, clear evidence of distinct metabolic signatures for lung AdC and SqCC, which can potentially impact on diagnosis and provide important leads for future research on novel therapeutic targets or imaging tracers.

Introduction

Lung cancer is a complex, heterogeneous disease with high incidence and mortality worldwide, being responsible for 22.5 and 13.8% of cancer-related deaths in men and women, respectively (1). Besides prevention, namely through smoking cessation, one of the greatest hopes to improve lung cancer management relies on personalized medicine, i.e. the selection of the safest and most efficacious treatment for each patient, based on the specific tumour pathological characteristics (2). According to the World Health Organization (WHO), malignant epithelial lung tumours are classified into seven major categories, the most prevalent histological types being adenocarcinoma (AdC) and squamous cell carcinoma (SqCC [3]). Several clinical trials have demonstrated differing efficacy and/or toxicity of particular treatments depending on tumour histological type, thus making accurate subtyping of lung tumours, particularly the distinction between AdC and SqCC, a mandatory diagnostic requirement (4). Tumour subtyping is normally performed by morphological assessment through light microscopy, whereas in cases of poorly differentiated tumours or small biopsy specimens, immunohistochemical markers are needed to achieve reliable subtyping (5).

Metabolic reprogramming of tumour cells to support continuous growth and proliferation is recognized as an emergent hallmark in cancer development (6). Indeed, there is cumulative evidence that many solid tumours show altered metabolic pathways, such as enhanced glucose uptake and glycolytic activity, increased *de novo* biosynthesis of nucleotides, increased glutaminolysis or a shift in citrate metabolism from oxidation to lipogenesis (7,8). Therefore, new insights into disease molecular aetiology, diagnostic markers or potential therapeutic targets will probably arise from the in-depth study of tumour cell metabolism. Metabolomic profiling of tumour tissues by ^{1}H High Resolution Magic Angle Spinning (HRMAS) nuclear magnetic resonance (NMR) spectroscopy offers a direct window to cancer altered metabolism, as tissue metabolite levels closely reflect the metabolic activity of tumour cells and their microenvironment. This approach has enabled general and specific metabolic markers of malignancy to be proposed for different cancer types and has great potential to assist in the interpretation of *in vivo* MRS data (9).

The metabolic features of lung carcinomas have been explored before by our group and others, either by direct HRMAS NMR spectroscopy (10–12) or by analysis of tissue extracts using NMR (13,14) or Mass Spectrometry (MS) methods (15,16). Although involving small numbers of patients (from 7 to 24), the previous studies have provided preliminary evidence for abnormal metabolism in malignant tissue, hence stimulating the more thorough examination of lung tumour metabolic signatures and their relation to histological and clinical parameters. This work represents the first global metabolomic study of a considerably larger sample set (over 100 tissue samples from 56 lung cancer patients), where direct tissue analysis by ^{1}H HRMAS NMR spectroscopy has been combined with multivariate modeling and validation procedures to ensure the robustness of findings. With this approach, besides characterizing the general metabolic features of malignant lung tumours, we aim to investigate the differences in the metabolic behaviour of adenocarcinoma and squamous cell carcinoma, as these differences may provide additional criteria for complementing diagnosis and improving clinical decisions.

Abbreviations: AdC, adenocarcinoma; GSH, glutathione; GPC, glycerophosphocholine; HRMAS, high resolution magic angle spinning; MCCV, Monte Carlo cross validation; NMR, nuclear magnetic resonance; PC, phosphocholine; PE, phosphoethanolamine; PLS-DA, Partial Least Squares Discriminant Analysis; ROC, receiver operating characteristic; SqCC, squamous cell carcinoma; UDP/UTP, uridine di/triphosphate; VIP, variable importance to the projection.

Materials and methods

Subjects

In total, 56 subjects diagnosed with primary lung cancer at surgical stage (40M/16F, 62 ± 10 average age) were included in this study, following informed consent and approval of the study protocol by the Ethics Committee of the University Hospitals of Coimbra. Supplementary Table 1, available at *Carcinogenesis* online, shows the information available for each subject. Final diagnosis was established after histopathological examination of surgical specimens, the tumours having been classified as 19 AdCs, 19 SqCCs, 6 sarcomatoid carcinomas, 4 large cell carcinomas, 2 adenosquamous carcinomas, 4 carcinoid tumours and 2 small cell carcinomas, according to the 2004 WHO Classification of Tumours (3). Based on the TNM staging system, most cases were classified as stage I (n 30) and stage II (n 18) and a few cases were of

stage III (n 6). In all cases, the surrounding parenchyma presented histological characteristics of 'smoking lung', mostly emphysema and bronchiolitis. The smoking history of the patients was not available.

Tissue sampling and preparation

Lung tumour and non-involved adjacent tissue (control) samples were retrieved from surgical specimens within a maximum of 30 min after surgery, immediately snap-frozen in liquid nitrogen and stored at −80°C until NMR analysis. The percentages of tumour cells and of necrotic tissue were determined by microscopic observation of mirror sections of samples collected for NMR and are presented in Supplementary Table 1, available at *Carcinogenesis* online. At the time of analysis, the frozen tissues were washed with a few drops of D_2O saline (0.9%) to remove excess blood, and *ca.* 40 mg of thawed cold tissue was packed into 50 µl HRMAS rotors, containing 10 µl of D_2O saline with 0.25% 3-(trimethylsilyl) propionate sodium salt (TSP)-*d*4 to provide a signal for lock (D_2O) and for shimming (TSP).

NMR measurements

NMR spectra were acquired on a Bruker Avance DRX-500 spectrometer operating at 500.13 MHz for ^1H observation, at 277 K, using a 4 mm HRMAS probe, in which the rotor containing the sample was spun at 4 kHz. Standard 1D ^1H spectra with water presaturation (pulse program 'noesypr1d', Bruker library) were acquired with a 6510 Hz spectral width, 32 K data points, a 4 s relaxation delay (d1), 100 ms mixing time (d8) and 256 scans. A T_2-edited spectrum (pulse program 'cpmgpr') and a diffusion-edited spectrum (pulse program 'ledbpgp2s1dpr') were also recorded for each sample, as described previously (10). All 1D spectra were processed with a line broadening of 0.3 (standard and T_2-edited) or 0.5 Hz (diffusion-edited), zero filling to 64 K data points, manual phasing and baseline correction. The chemical shifts were referenced internally to the alanine signal at δ 1.48 ppm. Spectral assignment was based on 2D total correlation spectroscopy (TOCSY) and heteronuclear single quantum coherence (HSQC) spectra (10), consultation of spectral databases, namely Bruker Biorefcode database and the human metabolome database (HMDB [17]) and further supported by statistical total correlation spectroscopy (STOCSY [18]).

Multivariate analysis and spectral integration

A total of 115 spectra corresponding to tumour and matched control tissues from 56 lung cancer patients (112 + 3 extra samples) were considered for multivariate analysis. Spectra were aligned using recursive segment-wise peak alignment (19) to minimize chemical shift variations, and normalized by probabilistic quotient normalization (PQN [20]), using the median of the control tissue spectra as reference, to account for dilution-independent effects on spectral area. Scaling to unit variance (UV) was then performed to highlight changes in low abundance metabolites. Data matrices were built in Amix-viewer (version 3.9.14, BrukerBiospin, Rheinstetten, Germany) using all intensity values in the δ 0.25–8.80 region (excluding the subregions containing suppressed water and contaminant signals from ethanol and polyethyleneglycol) or, in some instances, subsets of signal integrals or selected variables. Variable selection was based on the intersection method described by Diaz *et al.* (21), which accounted for the variable importance to the projection (VIP greater than 1) as well as VIP standard errors and b-coefficients. Principal component analysis, used to detect intrinsic grouping and/or outliers within the data set, was followed by Partial Least Squares Discriminant Analysis (PLS-DA) to maximize class discrimination (SIMCA-P 11.5, Umetrics, Umeå, Sweden). Model robustness was assessed by Monte Carlo Cross Validation (MCCV [22]) with 500 iterations and permutation testing (class membership randomly assigned). The distribution of Q^2 values (expressing the cross-validated prediction power) and the receiver operating characteristic (ROC) maps, representing the true positive rate (TPR or sensitivity) as a function of the false positive rate (FPR or 1-specificity), were recovered from MCCV applied to both original and permuted models. PLS1 regression was employed to assess the relationship between tissue metabolic profiles and different clinical and histological parameters (e.g. age of subjects, % of tumour cells, % necrosis).

To evaluate metabolite quantitative variations, selected signals in the standard 1D spectra were deconvoluted and integrated using Amix-Viewer (version 3.9.14, BrukerBiospin), as described previously (10), and normalized by the PQN quotient of each spectrum. Statistically significant differences in integral values between the groups compared were assessed using the *t*-test or the non-parametric analogue Wilcoxon rank-sum test with continuity correction. Statistical significance was assessed by Bonferroni-corrected *P* values (confidence level 95%). Moreover, effect sizes and corresponding confidence intervals were estimated to evaluate the magnitude of the differences, as described in Berben *et al.* (23). Additionally, Spearman correlation analysis was employed to evaluate the correlations within NMR variables and between these and some histological parameters, using a correlation coefficient cut-off of $|r| > 0.7$ and correlation significance of $P < 0.004$ (Bonferroni-corrected).

Results

General metabolic profile of lung tumours

Figure 1 shows the average spectral profiles of pulmonary control tissue and of the different lung tumour types analysed. Although there were several differences amongst the tumour profiles, which will be later addressed, some features (e.g. depleted glucose and increased lactate and phosphocholine) appeared to be common changes in relation to control tissue. Therefore, our first approach consisted of applying multivariate analysis to assess the discrimination between all tumours and control tissues, based on general metabolic features.

The resulting PLS-DA model, applied to standard 1D spectra, showed high explained variance and predictive power, the two tissue classes being well-separated along LV1 (Figure 2A). The robustness of this discrimination was then verified by MCCV and permutation testing, with prediction results being plotted in the ROC space (Figure 2B). While the original model (true classes assigned) showed 96.1% sensitivity, 97.8% specificity and an overall classification rate of 96.9% (Supplementary Table 2, available at *Carcinogenesis* online), random permutation of class membership resulted in a decrease of these parameters to values of no discrimination (approximately 50%). Moreover, the Q^2 values obtained for the original model showed a high median (0.83) and a distribution clearly distinct from that obtained for permutations (Figure 2C), therefore confirming the predictive power of the original model. PLS-DA and MCCV were also applied to T_2-edited and diffusion-edited spectra (Supplementary Figure 1, available at *Carcinogenesis* online). However, as shown by the resulting statistical parameters (Supplementary Table 3, available at *Carcinogenesis* online), the model built with the standard 1D spectra was found to be the most robust (higher predictive power, sensitivity and specificity); hence, further data analysis was based on the standard 1D experiment.

The thorough inspection of PLS-DA LV1 loading weights coloured as a function of VIP (Figure 2D) allowed the main compounds contributing to tissue discrimination to be identified. Lactate, alanine, glutamate, reduced glutathione (GSH), creatine, taurine, phosphocholine (PC), glycerophosphocholine (GPC), phosphoethanolamine (PE), uridine di/triphosphate (UDP/UTP) and peptide moieties (with negative loadings) were increased in tumours, whereas acetate and glucose (positive loadings) were decreased. Indeed, the integrals of all these metabolite signals showed statistically significant differences ($P < 0.004$, Bonferroni-corrected) and absolute effect sizes greater than 0.5, thus confirming their importance in the discrimination between tumour and control tissues. The percentages of variation for each discriminant metabolite are presented in Table I (left) and corresponding boxplot representations can be found in Supplementary Figure 2, available at *Carcinogenesis* online. Further confirmation of the discriminatory power of this set of metabolites was achieved by PLS-DA of their integrals, as model robustness and accuracy was maintained (Supplementary Table 2 and Supplementary Figure 3, available at *Carcinogenesis* online).

As the amount of tumour cells on mirror sections of NMR-analysed samples was found to vary between 5 and 99% (Supplementary Table 1, available at *Carcinogenesis* online), median and mode values being 50 and 90%, respectively; the possible influence of such variability on tumour metabolic composition was addressed by PLS1 regression, considering either the full spectral profile or the set of 13 discriminant metabolites. The resulting Q^2 values were very low (about 0.1 in both cases), thus indicating that the metabolic profiles of tumour tissues were not strongly determined by the varying proportions of stromal and tumour cells on each sample. The possible correlation between the NMR spectral profiles and the amount of necrotic tissue (Supplementary Table 1, available at *Carcinogenesis* online) were also investigated through PLS1 regression. The resulting model (n 53) showed a reasonable predictive power (Q^2 0.48 for 2 latent variables), with less necrotic samples located in negative LV1 and samples with higher % necrosis distributed towards positive LV1 (Supplementary Figure 4A, available at *Carcinogenesis* online). The

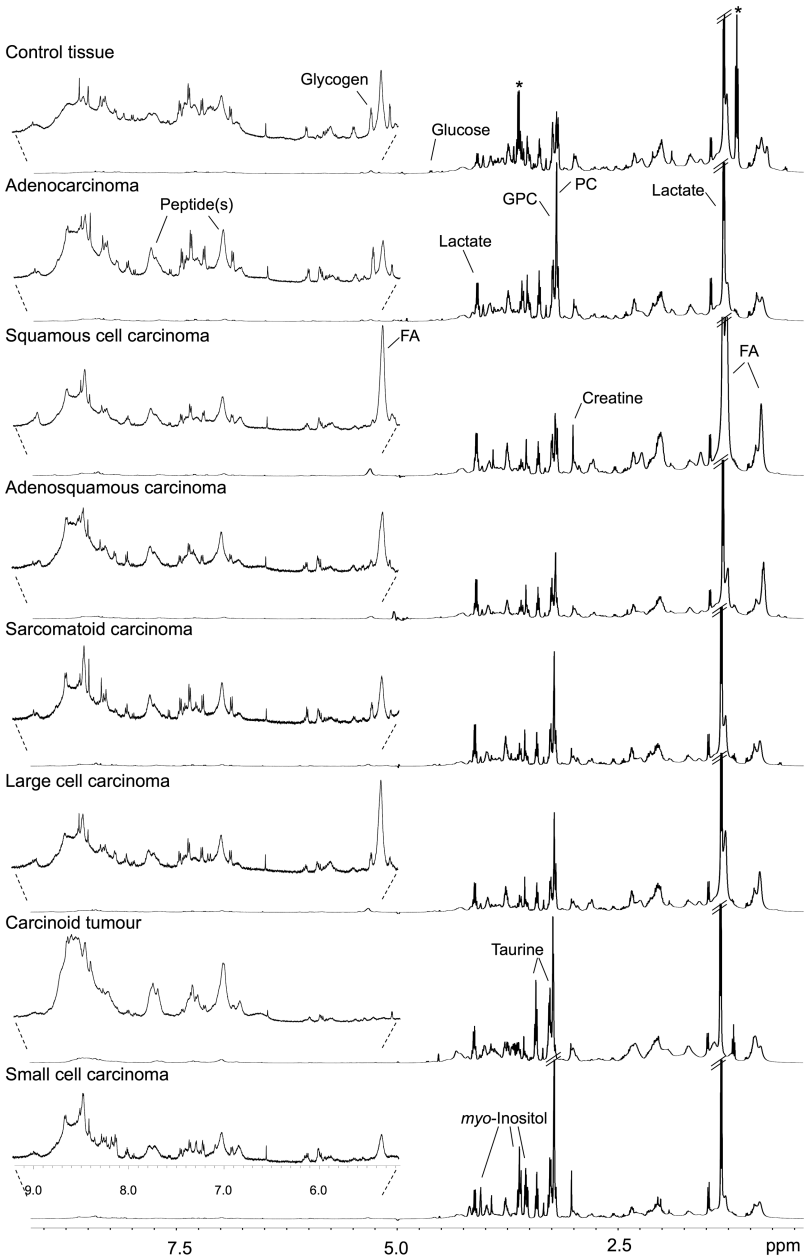

Fig. 1. Average standard 1D ^{1}H HRMAS NMR spectra of lung control and tumour tissues of different histological types. Some apparent metabolite differences are indicated. FA, fatty acyl chains in lipids. *Ethanol contamination.

corresponding VIP-coloured loadings (Supplementary Figure 4B, available at *Carcinogenesis* online) indicated that higher lipid levels characterized highly necrotic samples, which was confirmed by the significant correlations found between lipid signal areas and % necrosis.

Finally, the impact of stage on the tumour signature has been addressed by building separate PLS-DA models for each stage (I: 29 pairs, II: 19 pairs, III: 6 pairs). While the stage III model couldn't be properly validated, most probably due to the low number of samples available, the classification rates obtained for discriminating stage I or stage II tumours from their respective controls were similar to those achieved when considering the whole dataset (Supplementary Table 2, available at *Carcinogenesis* online). Moreover, inspection of the loadings profiles (not shown) revealed that the main discriminant features were the same for all stages, with the exception of taurine which did not appear important in the discrimination of stage III tumours from their controls. Concordantly, no valid discrimination between tumour tissues of different stages was obtained, thus indicating that, in the set of samples studied, stage did not have a major impact on the tumour metabolic profiles.

Metabolic behaviour of different tumour types: focus on AdC and SqCC

The differences between the tumour profiles shown in Figure 1 suggest that there may be a relationship between tumour histomorphology and metabolic composition. For instance, carcinoids were characterized by very low lipid levels and high taurine and peptide abundance, whereas small cell carcinomas showed prominently high levels of creatine and *myo*-inositol together with absence of glucose. There were also several apparent differences in the relative levels of PC, GPC and glycogen. However, given the low sample numbers representing most types (≤ 6 patients), only the profiles of AdC (n 19) and SqCC (n 19) were further examined in this study, with two main goals in mind: (i) to find out if these tumour types have distinct metabolic

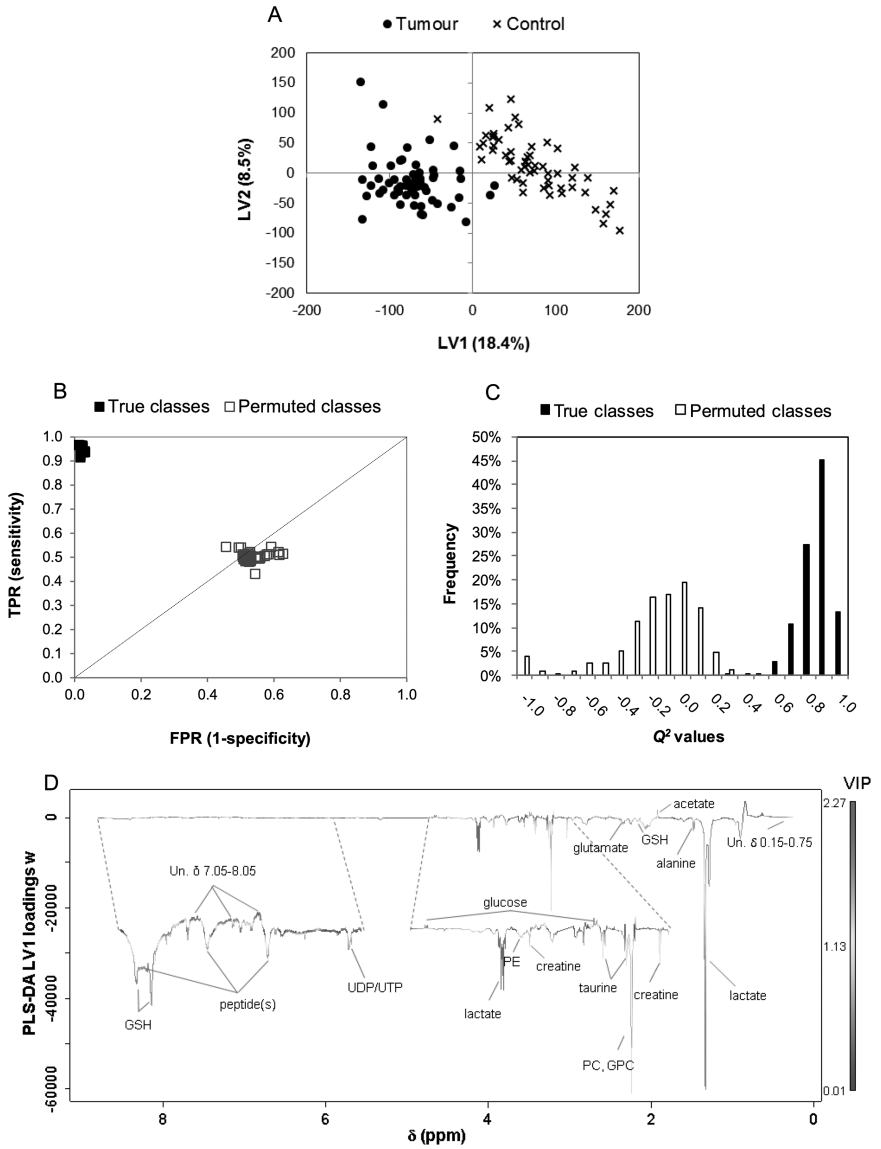

Fig. 2. Multivariate analysis of standard 1D ^{1}H NMR spectra from lung control (n 57) and tumour (n 58) tissues (**A**) PLS-DA scores scatter plot (*LV* 2, R^2X 0.27, R^2Y 0.85, Q^2 0.80), (**B**) ROC space and (**C**) Q^2 histogram obtained from MCCV of the PLS-DA model (500 iterations), (**D**) PLS-DA LV1 loading weights (w) coloured as a function of VIP, where most important metabolites for group discrimination are indicated.

behavior; (ii) to evaluate the potential of using NMR metabolomics as an adjunct tool in AdC versus SqCC differential classification.

To address the first goal, we have investigated the discrimination between each of these two tumour types and their respective control tissues. MCCV-validated PLS-DA models were obtained for both AdC and SqCC compared to controls, classification rates being 87.8 and 97.7%, respectively (Supplementary Table 2, available at *Carcinogenesis* online). The corresponding scores scatter plots, ROC maps, Q^2 distributions and loadings are shown in Supplementary Figure 5, available at *Carcinogenesis* online (A-D for AdC and E-H for SqCC). Notably, as confirmed by spectral integration, several metabolites showed differing importance in the discrimination between either AdC or SqCC tumours and their respective control tissues. For instance, AdC showed higher increases in PC, GPC, UDP/UTP and peptides, whereas SqCC was characterized by higher increases of lactate, glutamate, alanine, GSH and creatine, together with a more pronounced decrease of glucose (Table I, centre for AdC and right for SqCC).

Examination of dependencies between metabolite variations further complemented this analysis, the results being plotted as correlation heat maps colour-coded by the strength of Spearman correlation coefficients (Figure 3). A striking observation was that the intermetabolite correlation pattern was clearly different between AdC and SqCC. In AdC, significant correlations ($|r| > 0.7$, $P < 0.004$, Bonferroni-corrected) involved mainly phospholipid-related metabolites (PC, GPC and PE), while the SqCC pattern was characterized by correlations involving lactate, glucose, glutamate, alanine, GSH and creatine. A selected set of graphs showing the quantitative relationship between some metabolites found to be highly correlated in either AdC or SqCC can be found in Supplementary Figure 6, available at *Carcinogenesis* online.

For assessing the potential of metabolomics to discriminate between AdC and SqCC, multivariate analysis has been applied to their spectra, leaving out the control tissues. After excluding two clear outliers (AdC T13 and SqCC T37, characterized by very high lipids), the resulting PLS-DA model showed moderate accuracy and predictive power (classification rate 81.3% and median Q^2 0.47) (Supplementary Table 2, available at *Carcinogenesis* online). This discrimination could, however, be significantly improved by applying the variable selection method described in the experimental section, allowing a model with 94.3% classification rate and median Q^2 0.58 to be built. The corresponding PLS-DA scores and MCCV validation results are shown in Figure 4. Furthermore, as the AdC and SqCC groups compared were unbalanced in terms of gender (AdC: 7M/11F;

Table I. Metabolites showing statistically significant differences ($P < 0.004$, Bonferroni-corrected) between lung tumours and control tissues, considering all samples (2nd column) or each of the main histological types (3rd and 4th columns)

Metabolite (δ, multiplicity)[a]	All tumour types (56 pairs)			Adenocarcinoma (19 pairs)			Squamous cell (19 pairs)		
	% variation	Effect size	*P* value	% variation	Effect size	*P* value	% variation	Effect size	*P* value
Acetate (δ 1.92, s)	−57.0 ± 17.2	−0.88 ± 0.39	3.3×10^{-6}	−61.9 ± 22.2	−1.3 ± 0.70	1.3×10^{-4}	*n.s.*		
Alanine (δ 1.48, d)	37.1 ± 5.3	1.1 ± 0.40	1.2×10^{-6}	*n.s.*			46.9 ± 7.7	1.6 ± 0.73	1.6×10^{-4}
Creatine (δ 3.93, s)	63.9 ± 7.3	1.2 ± 0.40	1.2×10^{-8}	28.2 ± 6.7	1.2 ± 0.69	6.4×10^{-4}	126.4 ± 12.1	2.1 ± 0.79	3.8×10^{-6}
Glucose (δ 4.65, d)	−46.3 ± 6.8	−1.7 ± 0.43	7.9×10^{-10}	−29.3 ± 10.8	−1.0 ± 0.68	2.0×10^{-3}	−51.4 ± 11.4	−2.0 ± 0.77	3.8×10^{-6}
Glutamate (δ 2.35, m)	38.7 ± 3.8	1.6 ± 0.43	5.8×10^{-10}	24.2 ± 5.8	1.2 ± 0.69	4.2×10^{-4}	55.8 ± 6.1	2.3 ± 0.82	3.8×10^{-6}
GPC (δ 3.23, s)	118.4 ± 10.7	1.3 ± 0.41	5.0×10^{-8}	115.2 ± 15.8	1.5 ± 0.72	2.1×10^{-4}	*n.s.*		
GSH (δ 2.16, m)	42.8 ± 5.4	1.2 ± 0.40	2.3×10^{-9}	*n.s.*			72.5 ± 9.9	1.7 ± 0.75	3.8×10^{-6}
Lactate (δ 4.12, q)	95.6 ± 5.1	2.4 ± 0.49	8.2×10^{-11}	84.8 ± 7.4	2.6 ± 0.86	3.8×10^{-6}	121.8 ± 9.1	2.7 ± 0.88	3.8×10^{-6}
PC (δ 3.22, s)	177.4 ± 18.7	0.95 ± 0.39	6.5×10^{-9}	216.3 ± 21.5	1.6 ± 0.73	1.6×10^{-4}	37.5 ± 11.6	0.88 ± 0.67	2.0×10^{-3}
PE (δ 3.98, m)	30.9 ± 4.5	1.1 ± 0.40	2.0×10^{-7}	32.5 ± 5.8	1.6 ± 0.73	5.2×10^{-4}	*n.s.*		
Taurine (δ 3.42, t)	56.0 ± 7.3	1.1 ± 0.40	6.6×10^{-7}	*n.s.*			*n.s.*		
UDP/UTP (δ 5.96, m)	255.2 ± 17.4	1.2 ± 0.40	6.3×10^{-9}	422.2 ± 40.3	1.1 ± 0.68	8.0×10^{-4}	243.7 ± 2 5.3	1.4 ± 0.71	1.3×10^{-4}
Peptide (δ 7.72, br)	17.8 ± 5.8	0.54 ± 0.38	2.2×10^{-4}	28.0 ± 9.2	0.87 ± 0.67	2.0×10^{-3}	*n.s.*		

For each metabolite, the average percentage and coefficient of variation were obtained by spectral integration of selected signals. Effect sizes and *P* values are shown, indicating, respectively, the magnitude and statistical significance of the differences between tumour and control tissues.
n.s. not significant.
[a]s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad.

SqCC: 16M/2F), possible gender-related bias has been assessed. The principal component analysis scores plot, where a trend for separation of tumour types was clear, showed no gender-related clustering (Supplementary Figure 7A and B, available at *Carcinogenesis* online). Moreover, when considering only adenocarcinoma samples (7M/11F), no valid discrimination could be achieved between males and females, as shown by the superimposed Q^2 distributions and ROC spaces of true and permuted models (Supplementary Figure 7C and D, available at *Carcinogenesis* online). Therefore, gender could be excluded as an important confounding factor. In regard to age, the difference between groups was not statistically significant (median age 64 and 63.5 for AdC and SqCC groups, respectively), thus reducing the possibility of any age-related bias. Moreover, the subjects' age was found to have a negligible impact on the tumour tissue profile as expressed by the low Q^2 values of PLS1 regression models built for each of the two major tumour types (not shown).

The metabolic features discriminating AdC and SqCC tumour tissues were then inspected by integrating a number of signals within the variables selected. Those showing statistically significant differences ($P < 0.002$, Bonferroni-corrected) between AdC and SqCC are listed in Table II, along with the percentages of variation and effect sizes. Creatine and GSH were found to be relatively higher in SqCC, whereas PC, PE, *myo*-inositol and taurine were higher in AdC. The ability of this set of six signals to discriminate between AdC and SqCC was then further evaluated by PLS-DA of their integrals (Supplementary Figure 8, available at *Carcinogenesis* online). A valid model could still be obtained although with lower classification rate (82.8%), thus showing the importance of considering a more comprehensive profile.

Discussion

The results hereby presented provide clear evidence of a strong metabolic signature for lung tumours, defined by a set of thirteen metabolites, which corroborates some typical features of cancer metabolism. Moreover, although sharing most of this malignancy signature, the major tumour types studied, adenocarcinoma (AdC) and squamous cell carcinoma (SqCC), were found to be characterized by important differences in their metabolic behaviour. While all tumours generally showed significantly depleted glucose together with increased lactate, this variation was much more pronounced in SqCC than in AdC, glucose and lactate levels showing a significant negative correlation in the former. These results highlight the dependency of the known glycolytic phenotype of lung tumours, exploited in cancer diagnosis and staging through ^{18}fluorodeoxyglucose (FDG) positron emission tomography (PET) imaging (24), on the histological type. Concordantly, previous studies have shown that, compared to AdC, SqCC tumours were characterized by higher FDG uptake and increased overexpression of key membrane glucose transporters, namely GLUT1, necessary to support the high rate of glycolysis (25–27). In addition to lactate, glutamate and alanine were also significantly increased and positively correlated in SqCC tumours, whereas in AdC, glutamate registered a smaller increase and alanine did not differ significantly between tumour and control tissues. Glutamate may originate from hydrolysis of glutamine (glutaminolysis), which, in addition to glucose, is regarded as an important source of energy and intermediate building blocks for tumour cell growth (28,29). Subsequent transamination between glutamate and pyruvate, catalysed by alanine transaminase, produces alanine and α-ketoglutarate, the latter entering the tricarboxylic acid cycle. Glutamate may also leave the glutaminolytic pathway and be excreted, as an effective way to excrete hydrogen, or to act as immunosuppressive in the protection of tumour cells (7). Therefore, our results suggest higher glutaminolytic activity for SqCC tumours than for AdC. Another important difference in the metabolic profiles of these histological types regarded the levels of reduced glutathione (GSH, γ-Glu-Cys-Gly) which were significantly increased in SqCC but not in AdC. GSH is a key

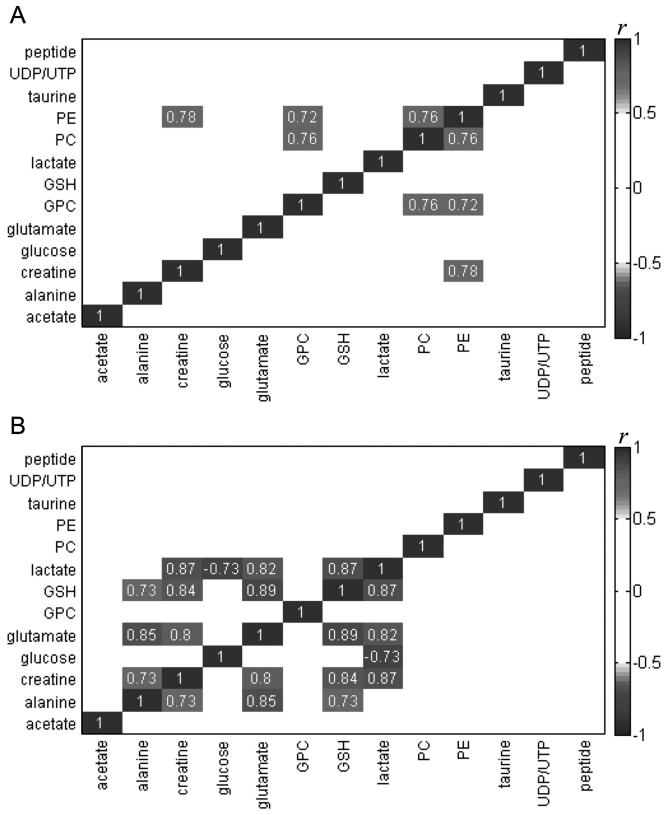

Fig. 3. Correlation heat maps colour-coded by the strength of Spearman correlation coefficients (r) between metabolites found to be important in tumour versus control discrimination: (**A**) AdC and (**B**) SqCC. Cut-off values of $|r| > 0.7$ and $P < 0.004$ (Bonferroni-corrected) have been used.

component of the cellular defence system for detoxification of oxidative stress-causing species or drugs. High intracellular GSH levels together with increased activity of GSH-related enzymes are a typical feature of several tumours (30), including lung tumours (31–33), and have been associated with chemoresistance and high metastatic potential (34). Regarding the GSH variation in lung tumours of different histology, our findings are consistent with an early report (35) but disagree with another study reporting higher GSH increase in AdC than in SqC compared to non-malignant adjacent tissue (33). It should however be noted that other patients characteristics, such as age and smoking habits, are expected to strongly influence the oxidative burden, hence the tissue GSH levels. Therefore, data on the subjects smoking history, which is lacking in our study, would be necessary to further interpret the observed discrepancy and achieve solid conclusions. Creatine was significantly increased in both tumour types but at much higher magnitude in SqCC than in AdC (126.4 versus 28.2% increase, respectively). Elevated creatine levels in tumour tissues have previously been related to altered energetic transfer processes (36,37) and may reflect decreased activity of creatine kinase (CK). Indeed, CK activity has been previously reported to be lower in both AdC and SqCC lung tumours, compared to normal adjacent tissues, although not showing a significant difference between tumour types (38).

In regard to phospholipid-related metabolites (PC, GPC and PE), the differences between tumour and control tissues were more marked in AdC than in SqCC. PC may arise from both anabolic and catabolic transformations of phophatidylcholine (PtdCho), the major cell membrane phospholipid. A key enzyme involved in this cycle is choline kinase (ChoK), which phosphorylates choline to produce PC, as the first step in PtdCho synthesis. ChoK has been found overexpressed and highly active in lung tumours and cell lines (39,40), and has been proposed as a prognostic factor for early-stage lung cancer (40). Also, ^{11}C-choline PET is being explored as an alternative to ^{18}FDG-PET in lung cancer staging, based on increased choline uptake and metabolism by rapidly proliferating tumour cells (41). On the other hand, phospholipid breakdown has also been linked to abnormal

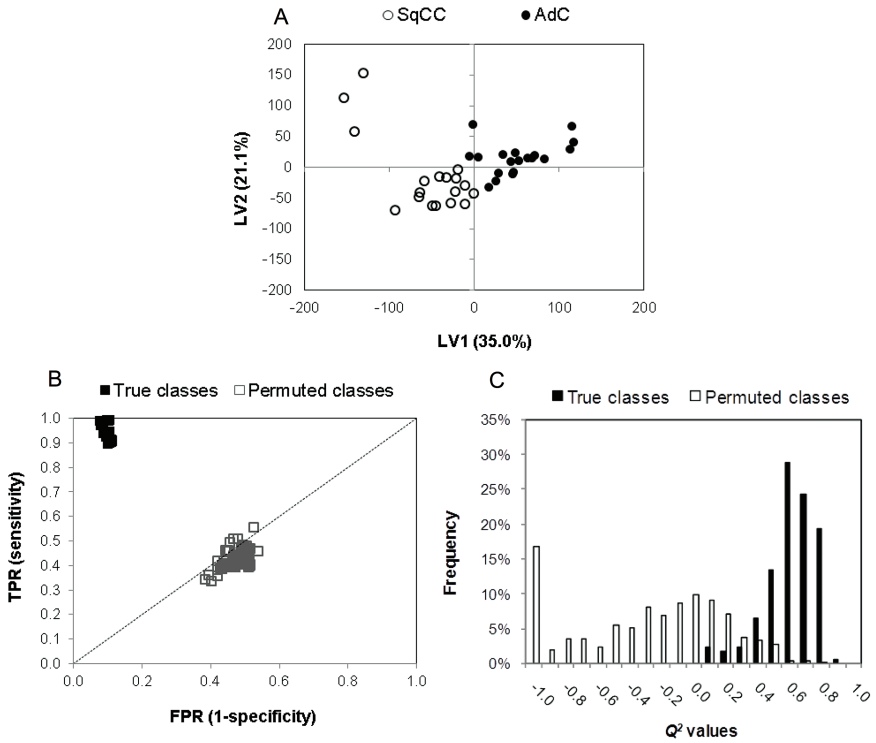

Fig. 4. Multivariate analysis of ^{1}H NMR spectra from AdC (n 18) and SqCC (n 18) tissues, after exclusion of T13 and T37 (characterized by very high lipid levels) and application of variable selection: (A) PLS-DA scores scatter plot, (**B**) ROC space and (**C**) Q^2 histogram obtained from MCCV of the PLS-DA model (500 iterations).

Table II. Metabolites showing statistically significant differences ($P < 0.002$, Bonferroni-corrected) between AdC and SqCC, as assessed by spectral integration of selected signals

Metabolite, δ (multiplicity)[a]	% variation	Effect size	*P* value
Creatine (δ 3.93, s)	55.8 ± 11.7	1.3 ± 0.71	6.1×10^{-4}
GSH (δ 2.16, m)	36.8 ± 9.8	1.1 ± 0.69	7.0×10^{-4}
myo-Inositol (δ 3.62, t)	−54.5 ± 13.5	−1.8 ± 0.76	8.5×10^{-6}
PC (δ 3.22, s)	−71.6 ± 21.2	−1.7 ± 0.75	2.0×10^{-6}
PE (δ 3.98, m)	−22.4 ± 4.6	−1.8 ± 0.76	3.1×10^{-6}
Taurine (δ 3.43, t)	−29.0 ± 10.1	−1.1 ± 0.69	1.9×10^{-3}

Positive/negative variations refer, respectively, to metabolites increased/decreased in SqCC compared to AdC. The magnitude and statistical significance of the differences between average values are expressed by the effect size and *P* value, respectively.
[a]s, singlet; t, triplet; q, quartet; m, multiplet.

metabolism in cancer and suggested to promote malignant transformation via mitogenic signal transduction (9). As GPC is solely a product of PtdCho degradation, it is likely that the observed increases of PC and GPC, together with increased PE levels, reflect both anabolic and catabolic processes, which our results suggest to be more relevant in AdC. Additionally, significantly decreased acetate levels in AdC (not observed in SqCC) may also relate to lipid metabolism as this metabolite can be converted to acetyl-CoA to feed *de novo* biosynthesis of lipids. Indeed, the shift from citrate oxidation to lipogenesis is another typical feature of cancer metabolism (42). However, although total lipids did show a (non-significant) trend to be higher in tumours, their levels seemed to be influenced by other factors like necrosis and did not prove determinant in the discrimination between malignant and non-malignant tissue. AdC tumour spectra were also characterized by increased contributions from peptide moieties, while no significant difference existed between SqCC and their respective controls. Concordantly, Wu *et al.* (43) have found elevated levels of 90 dipeptides in lung AdC tissues (from nine patients) relatively to normal tissues. In another study (44), 17 dipeptides were found to be increased in cancer-associated fibroblasts (CAFs) isolated from lung tumours, compared to matched normal fibroblasts obtained from non-neoplastic tissues. The authors postulated that the observed increase could relate to enhanced autophagy, a lysosome-related protein degradation mechanism, which has recently emerged as a key regulator of multiple aspects of cancer biology (45,46). Finally, uracil nucleotides (UDP and/or UTP) were found to be significantly increased in tumours, especially in AdC. Given the important role of these molecules as extracellular signalling molecules (47) and in cancer-related glycosylation of proteins and lipids (48), this is an interesting clue to further investigate their biological relevance in the context of lung cancer.

Regarding the differential classification of AdC and SqCC, the main metabolites highlighted as being significantly different between the two tumour types were creatine and GSH (higher in SqCC), together with PC, PE, *myo*-inositol and taurine (higher in AdC). Interestingly, the latter two metabolites, known to play a role in cellular osmoprotection, were important in AdC versus SqCC discrimination in spite of not being significantly different between each tumour type and their respective control tissues.

In summary, this work provided new, clear evidence that lung AdC and SqCC show distinct metabolic behaviour. Major alterations in AdC tumours regarded phospholipid metabolism and protein catabolism, whereas SqCC had a much stronger glycolytic and glutaminolytic profile. This difference could have important implications in the selective definition of the most appropriate targets for developing new anticancer therapies or imaging tracers in lung cancer research. Furthermore, NMR metabolomics has shown great potential for discriminating between AdC and SqCC tissues, based on a set of selected variables, thus highlighting the valuable role of this approach in the subtyping of tumours. Future work will focus on the metabolic profiling of more challenging samples, with uncertain histomorphology, with a view to compare the classification accuracy of metabolic markers with that of established immunohistochemical markers.

Supplementary material

Supplementary Figures 1–8 and supplementary Tables 1–3 can be found at http://carcin.oxfordjournals.org/

Funding

European Regional Development Fund (FEDER) through the Competitive Factors Thematic Operational Programme (COMPETE) and national funds through the Foundation for Science and Technology (FCT), Portugal, under grant to C.R. (SFRH/BD/63916/2009) and projects CICECO–FCOMP-01-0124-FEDER-037271 (Ref. FCT PEst-C/CTM/LA0011/2013), PTDC/QUI/68017/2006 and QOPNA (Ref. FCT PEst-C/QUI/UI0062/2013).

Acknowledgements

The authors are grateful to Dr Manfred Spraul, Bruker BioSpin, Germany, for providing access to spectral databases. CIMAGO (Faculty of Medicine, University of Coimbra) and the Portuguese National NMR (PTNMR) Network, supported with FCT funds, are also acknowledged.

Conflict of Interest Statement: None declared.

References

1. Ferlay,J. *et al*. (2010) Estimates of cancer incidence and mortality in Europe in 2008. *Eur. J. Cancer*, **46**, 765–781.
2. Kerr,K.M. (2012) Personalized medicine for lung cancer: new challenges for pathology. *Histopathology*, **60**, 531–546.
3. Travis,W.D. *et al*. (2004). *World Health Organization classification of tumours; tumours of lung, pleura, thymus and heart*. IARC Press, Lyon, France.
4. Cooper,W.A. *et al*. (2011) What's new in non-small cell lung cancer for pathologists: the importance of accurate subtyping, EGFR mutations and ALK rearrangements. *Pathology*, **43**, 103–115.
5. Warth,A. *et al*. (2012) Large-scale comparative analyses of immunomarkers for diagnostic subtyping of non-small-cell lung cancer biopsies. *Histopathology*, **61**, 1017–1025.
6. Hanahan,D. *et al*. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646–674.
7. Mazurek,S. (2007) Tumour cell energetic metabolome. In: Saks,V. (ed) *Molecular System Bioenergetic: Energy for Life*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, pp. 521–540.
8. Cantor,J.R. *et al*. (2012) Cancer cell metabolism: one hallmark, many faces. *Cancer Discov*., **2**, 881–898.
9. Moestue,S. *et al*. (2011) HR MAS MR spectroscopy in metabolic characterization of cancer. *Curr. Top. Med. Chem*., **11**, 2–26.
10. Rocha,C.M. *et al*. (2010) Metabolic profiling of human lung cancer tissue by ^{1}H high resolution magic angle spinning (HRMAS) NMR spectroscopy. *J. Proteome Res*., **9**, 319–332.
11. Duarte,I.F. *et al*. (2010) Can nuclear magnetic resonance (NMR) spectroscopy reveal different metabolic signatures for lung tumours? *Virchows Arch*., **457**, 715–725.
12. Chen,W. *et al*. (2011) Study on metabonomic characteristics of human lung cancer using high resolution magic-angle spinning ^{1}H NMR spectroscopy and multivariate data analysis. *Magn. Reson. Med*., **66**, 1531–1540.
13. Guo,J. *et al*. (2004) *In vitro* proton magnetic resonance spectroscopic lactate and choline measurements, ^{18}F-FDG uptake, and prognosis in patients with lung adenocarcinoma. *J. Nucl. Med*., **45**, 1334–1339.
14. Yokota,H. *et al*. (2007) Lactate, choline, and creatine levels measured by vitro ^{1}H-MRS as prognostic parameters in patients with non-small-cell lung cancer. *J. Magn. Reson. Imaging*, **25**, 992–999.
15. Hori,S. *et al*. (2011) A metabolomic approach to lung cancer. *Lung Cancer*, **74**, 284–292.
16. Kami,K. *et al*. (2013) Metabolomic profiling of lung and prostate tumor tissues by capillary electrophoresis time-of-flight mass spectrometry. *Metabolomics*, **9**, 444–453.
17. Wishart,D.S. *et al*. (2009) HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res*., **37**(Database issue), D603–D610.

18. Cloarec,O. *et al*. (2005) Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic ^{1}H NMR data sets. *Anal. Chem*., **77**, 1282–1289.
19. Veselkov,K.A. *et al*. (2009) Recursive segment-wise peak alignment of biological ^{1}H NMR spectra for improved metabolic biomarker recovery. *Anal. Chem*., **81**, 56–66.
20. Dieterle,F. *et al*. (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics. *Anal. Chem*., **78**, 4281–4290.
21. Diaz,S.O. *et al*. (2013) Second trimester maternal urine for the diagnosis of trisomy 21 and prediction of poor pregnancy outcomes. *J. Proteome Res*., **12**, 2946–2957.
22. Xu,Q.S. *et al*. (2001) Monte Carlo cross validation. *Chemom. Intell. Lab. Syst*., **56**, 1–11.
23. Berben,L. *et al*. (2012) Effect size estimation: methods and examples. *Int. J. Nurs. Stud*., **49**, 1039–1047.
24. Ambrosini,V. *et al*. (2012) PET/CT imaging in different types of lung cancer: an overview. *Eur. J. Radiol*., **81**, 988–1001.
25. Brown,R.S. *et al*. (1999) Glucose transporters and FDG uptake in untreated primary human non-small cell lung cancer. *J. Nucl. Med*., **40**, 556–565.
26. de Geus-Oei,L.F. *et al*. (2007) Biological correlates of FDG uptake in non-small cell lung cancer. *Lung Cancer*, **55**, 79–87.
27. Meijer,T.W. *et al*. (2012) Differences in metabolism between adeno- and squamous cell non-small cell lung carcinomas: spatial distribution and prognostic value of GLUT1 and MCT4. *Lung Cancer*, **76**, 316–323.
28. Wise,D.R. *et al*. (2008) Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc. Natl Acad. Sci. USA*, **105**, 18782–18787.
29. Mohamed,A. *et al*. (2014) Altered glutamine metabolism and therapeutic opportunities for lung cancer. *Clin. Lung Cancer*, **15**, 7–15.
30. Gamcsik,M.P. *et al*. (2012) Glutathione levels in human tumors. *Biomarkers*, **17**, 671–691.
31. Blair,S.L. *et al*. (1997) Glutathione metabolism in patients with non-small cell lung cancers. *Cancer Res*., **57**, 152–155.
32. Ferruzzi,E. *et al*. (2003) Blood glutathione as a surrogate marker of cancer tissue glutathione S-transferase activity in non-small cell lung cancer and squamous cell carcinoma of the head and neck. *Eur. J. Cancer*, **39**, 1019–1029.
33. Ilonen,I.K. *et al*. (2009) Oxidative stress in non-small cell lung cancer: role of nicotinamide adenine dinucleotide phosphate oxidase and glutathione. *Acta Oncol*., **48**, 1054–1061.
34. Traverso,N. *et al*. (2013) Role of glutathione in cancer progression and chemoresistance. *Oxid. Med. Cell. Longev*., **2013**, 972913.
35. Cook,J.A. *et al*. (1991) Cellular glutathione and thiol measurements from surgically resected human lung tumor and normal lung tissue. *Cancer Res*., **51**, 4287–4294.
36. Somashekar,B.S. *et al*. (2011) Magic angle spinning NMR-based metabolic profiling of head and neck squamous cell carcinoma tissues. *J. Proteome Res*., **10**, 5232–5241.
37. Yang,Y. *et al*. (2013) Study of metabonomic profiles of human esophageal carcinoma by use of high-resolution magic-angle spinning ^{1}H NMR spectroscopy and multivariate data analysis. *Anal. Bioanal. Chem*., **405**, 3381–3389.
38. Joseph,J. *et al*. (1997) Creatine kinase activity and isoenzymes in lung, colon and liver carcinomas. *Br. J. Cancer*, **76**, 600–605.
39. Molina,A.R. *et al*. (2002) Overexpression of choline kinase is a frequent feature in human tumour-derived cell lines and in lung, prostate, and colorectal human cancers. *Biochem. Biophys. Res. Commun*., **296**, 580–583.
40. Molina,A.R. *et al*. (2007) Expression of choline kinase alpha to predict outcome in patients with early-stage non-small-cell lung cancer: a retrospective study. *Lancet Oncol*., **8**, 889–897.
41. Li,M. *et al*. (2013) Value of 11C-choline PET/CT for lung cancer diagnosis and the relation between choline metabolism and proliferation of cancer cells. *Oncol. Rep*., **29**, 205–211.
42. Costello,L.C. *et al*. (2005) 'Why do tumour cells glycolyse?': from glycolysis through citrate to lipogenesis. *Mol. Cell. Biochem*., **280**, 1–8.
43. Wu,M. *et al*. (2013) Liquid chromatography/mass spectrometry methods for measuring dipeptide abundance in non-small-cell lung cancer. *Rapid Commun. Mass Spectrom*., **27**, 2091–2098.
44. Chaudhri,V.K. *et al*. (2013) Metabolic alterations in lung cancer-associated fibroblasts correlated with increased glycolytic metabolism of the tumor. *Mol. Cancer Res*., **11**, 579–592.
45. Kimmelman,A.C. (2011) The dynamic nature of autophagy in cancer. *Genes Dev*., **25**, 1999–2010.
46. Singh,R. *et al*. (2011) Autophagy in the cellular energetic balance. *Cell Metab*., **13**, 495–504.
47. Morrone,F.B. *et al*. (2003) Extracellular nucleotides and nucleosides induce proliferation and increase nucleoside transport in human glioma cell lines. *J. Neurooncol*., **64**, 211–218.
48. Christiansen,M.N. *et al*. (2014) Cell surface protein glycosylation in cancer. *Proteomics*, **14**, 525–546.

Received August 21, 2014; revised October 8, 2014;
accepted October 20, 2014