Combining High Resolution Magic Angle Spinning H1 NMR and Molecular Genomics Predicts Survival in Brain Tumor Patients Better than Either Methodology Alone

L. G. Astrakas[1,2], K. D. Blekas[3], O. C. Andronesi[1,4], M. N. Mindrinos[5], P. M. Black[6], L. G. Rahme[7], and A. A. Tzika[1,4]

[1]NMR Surgical Laboratory, Department of Surgery, Massachusetts General Hospital and Shriners Burns Institute, Harvard Medical School, Boston, MA, United States, [2]Department of Medical Physics, University of Ioannina, Ioannina, Greece, [3]Department of Computer Science, University of Ioannina, Ioannina, Greece, [4]Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Athinoula A. Martinos Center of Biomedical Imaging, Boston, MA, United States, [5]Stanford Genome Technology Center, Department of Biochemistry, Stanford University School of Medicine, Palo Alto, CA, United States, [6]Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States, [7]Molecular Surgery Laboratory, Department of Surgery, Massachusetts General Hospital and Shriners Burn Institute, Harvard Medical School, Boston, MA, United States

Introduction: Our objective was to develop and optimize a novel approach that combines biomarkers detected with high-resolution magic angle spinning (HRMAS) ^{1}H NMR and molecular genomics to improve prognostication of brain tumors. Our hypothesis is that current tissue typing can be enhanced by developing and applying a classification strategy analysis algorithm that produces unique tumor fingerprints by combining biomarker profiles from MRS and whole-genome expression profiling performed on microscale brain tumor biopsies. Fusion of different sources of information can be used to improve system performance and facilitate detection, recognition, identification, tracking, change detection, and decision making in defense, robotics, and medicine [1, 2]. Few previously described classifiers have attempted to combine data from different sources [3-6]. An efficient fusion scheme using complementary information can improve confidence and accuracy. Here, we focus on fusing HRMAS ^{1}H NMR data of brain tumor biopsies and gene expression data received from the same brain tumor biopsies.

Methods: We used a previously designed 2D adiabatic Total through Bond correlation SpectroscopY (TOBSY) HRMAS ^{1}H NMR pulse sequence, based on novel concepts rooted in solid-state NMR spectroscopy (7). All HRMAS experiments were performed on a Bruker Bio-Spin Avance NMR spectrometer (600.13 MHz) using a 4mm triple resonance (^{1}H, ^{13}C, ^{2}H) HRMAS probe. After HRMAS, RNA was extracted, purified, and quantified and genomic analysis was performed following standard Affymetrix protocols (Affymetrix, CA, USA). We then employed a Support Vector Machine (SVM) classifier [8] with linear kernels, implemented in the LIBSVM environment for multi-class SMV [9]. The original feature space comprised 54,675 genes. A **f**eature **s**election (FS) scheme using minimum redundancy – maximum relevance (MRMR) [10] was employed to reduce the dimensionality of feature space (Figure 1). MRMR is a powerful and robust FS framework that captures class characteristics in a broad spectrum by reducing mutual redundancy within the feature set. We tested our algorithm with the standard leave-one-out training/testing scheme. Our dataset had N=55 gene expression profiles derived from normal (9 cases) and tumor (46 cases) classes. The tumor class (N=46 samples) comprised three categories: high grade (H) [20 cases: 12 glioblastoma multiforme (GBM); 8 anaplastic astrocytoma (AA)], low grade (L) (17 cases: 7 meningioma; 7 schwanoma; 7 pylocitic astrocytoma) and metastasized (M) (11 cases: 5 adenocarcinoma; 3 breast cancer metastasis; 3 other metastasis). The following 15 NMR features were used: choline (Cho), phosphocholine (PC), glycerophosphocholine (GPC), phosphoethanolamine (PE), ethanolamine (Etn), γ-aminobutyric acid (GABA), n-acetyl aspartate (NAA), aspartate (Asp), alanine (Ala), polyunsaturated fatty acids (PUFA), glutamine (Gnl), glutamate (Glu), lactate (Lac), taurine (Tau) and lipids (Lip).

Results: The architecture of our classification system is shown in Fig. 1. A typical *ex vivo* HRMAS ^{1}H NMR spectra using TOBSY in anaplastic astrocytoma is shown in Fig.2. We performed multiple stepwise logistic regression analysis to evaluate how gene expression values, HRMAS NMR data, and their combination predict survival. We chose the 15 best genes according to their MRMR algorithm rank and 15 metabolite values (Ala, Asp, Cho, Etn, GABA, Gln, Glu, GPC, Lac, Lip, Myo, NAA, PC, PE, PUFA), (see above) corresponding to 49 available binary clinical outcomes (33 survived vs. 16 deceased). Our results proved that the combination of HRMAS NMR and genomic data improves our ability to predict clinical outcome. More specifically, gene data alone achieved high sensitivity, predicting 15 out of the 16 deceased cases (sensitivity 94%), high specificity, predicting 32 out of the 33 cases (specificity 97%), and high accuracy (96%). HRMAS NMR data had inferior sensitivity (11/16, 69%), specificity (28/32, 85%), and accuracy (80%). However, logistic regression on the combined HRMAS NMR and genomics data achieved a perfect classification (100% for all indices) of survived and diseased cases.

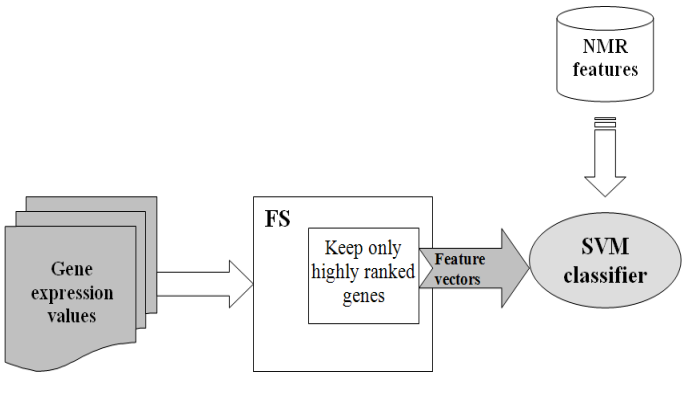

Figure 1. The architecture of the classification system

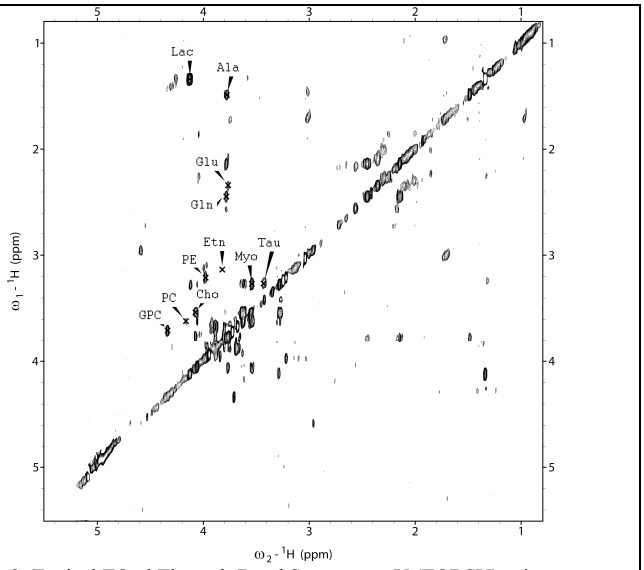

Figure 2. Typical TOtal Through-Bond SpectroscopY (TOBSY) using *ex vivo* HRMAS MRS on anaplastic astrocytoma

Discussion: Although we believe that these promising results are affected by sample size, they clearly demonstrate that the combination of NMR gene and expression data predict a clinically meaningful parameter such as survival better than either technique alone. Our results, using adult brain tumor biopsies, demonstrate that with appropriate quality control, we are able to produce meaningful data and introduce a novel classification scheme that complements and substantiates the current hypothesis of cancer stem cells [11] as a means of determining brain tumor classification and treatment. The data suggest that clinical MRI, MRS and MR imaging of gene expression *in vivo* can be combined to produce improved combined images, which could then be used to readily discriminate between different tumor types as well as metastasis and high-grade gliomas, a distinction not made adequately at present [3].

References

1. Hall DL, Linas J. *Handbook of Multisensor Data Fusion*, CRC Press, 2001
2. Roemer MJ et al, *Aerospace Conference 2001*, Big Sky, MT, IEEE, 6: 2979–2988
3. De Edelenyi FS et al, Nat Med 2000;6:1287-1289
4. Simonetti AW et al, NMR Biomed 2005;18:34-43
5. Galanaud D et al, Magn Reson Med 2006;55:1236-1245
6. Luts J et al, Artif Intell Med 2007;40:87-102
7. Andronesi OC et al., J Magn Reson., 193(2):251, 2008.
8. Vapnik V. Statistical Learning Theory, Wiley-Interscience, 1998
9. Hsu CW, Lin CJ, IEEE Trans Neural Netw 2002;13:415-425
10. Ding C, Peng H, J Bioinform Comput Biol 2005;3:185-205
11. Sakariassen PO, Immervoll H, Chekenya M, Neoplasia 2007;9:882-892