Brain tumor classification using a novel H1 HRMAS MRS method and robust algorithmic classifiers

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

[1]NMR Surgical Laboratory, MGH & Shriners Hospitals, Harvard Medical School, Boston, MA, United States, [2]Radiology, Athinoula A. Martinos Center for Biomedical Imaging, Boston, MA, United States, [3]Computer Science, University of Ioannina, Ioannina, Greece, [4]Medical Physics, University of Ioannina, Ioannina, Greece, [5]Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

Introduction– We present results combining a novel 2D TOBSY (TOtal Through-Bond SpectroscopY ^{1}H HRMAS MRS [1] with a robust SVM (Support Vector Machine) classifier with several feature-selection schemes [2]. Our results provide the framework for a non-subjective diagnostic approach that relies on highly informative biomarkers. This is justified because tumor classification according to histological features of the assumed cell of origin [3] is frequently controversial, since tumors often do not follow classic histology, and pathological diagnosis can therefore be subjective [4].

Materials and Methods– ♦We analyzed fifty-five biopsies, N=9 control from epileptic surgeries and N=46 from patients with brain tumors. Tumor biopsies belonged to three categories: high-grade [N=20: 12 glioblastoma multiforme (GBM); 8 anaplastic astrocytoma (AA)], low-grade (N=17: 7 meningioma; 7 schwanoma; 3 pylocitic astrocytoma), and brain metastases (N=9: 5 from adenocarcinoma; 4 from breast cancer). Subject age ranged from 17 to 54 years. ♦We performed *ex vivo* MRS employing 2D TOBSY ^{1}H HRMAS [1]. All MRS measurements 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 (Bruker), at -8°C to minimize tissue degradation. An external standard (TSP, M_w=172, d=0.00 ppm) served as reference for resonance chemical shift and for quantification. Typical parameters were, 2 s repetition time, 45 ms mixing time, and 45 min total acquisition time. •We quantified brain metabolites using the ratio of the cross peak volumes of the metabolites *CPV(M)* to the TSP diagonal peak volume *DPV(TSP)*. This ratio normalized to the biopsy weight (*w*) yielded the normalized metabolite intensity $I_c(M)$, $I_c(M)=(1/w)\times CPV(M)/DPV(TSP)$. ♦A linear Support Vector Machine (SVM) [5] was employed and tried with three MRS feature selection schemes: (i) the full feature space comprised of all 16 MRS features, (ii) a reduced feature space comprised of 4 specific features [choline (Cho), lactate (Lac), lipids (Lip), and n-acetyl aspartate (NAA)], and (iii) features selected with the minimum redundancy/maximum relevance (MRMR) feature selection method [6] that captures class characteristics in a broader spectrum by reducing mutual redundancy within the feature set. We employed a standard leave-one-out training/testing scheme. Performance was evaluated using "accuracy" (the percentage of correctly classified cases), "sensitivity" (the ratio of true positives to the sum of true positives and false negatives), and "specificity" (the ratio of true negatives to the sum of false positives and true negatives).

Results– All 16 metabolites were assigned from 2D TOBSY spectra (representative spectra shown in Figure 1) according to Tugnoli V et al [7]. Saturated Lip and polyunsaturated fatty acids (PUFAs) were prominent in high-grade tumors (i.e., GBM) and metastases, but not in low-grade tumors such as the Pilocytic Astrocytomas (Figure 1). The rest of the metabolites are detected at various levels in all tumors. The performance of the linear SVM classifier is summarized in Figure 2.

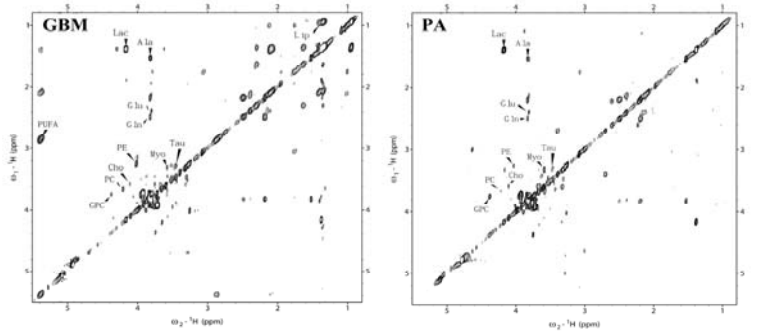

Fig. 1. E*x vivo* TOBSY HRMAS MRS on control and tumor tissue biopsies. 600 MHz HRMAS ^{1}H MR spectra using TOBSY (45 ms mixing time, 3 kHz MAS speed, -8°C). **A.** Glioblastoma multiforme (GBM), **B.** Pilocytic Astrocytoma (PA).

(Ala, alanine; Cho, choline; GABA, γ-aminobutyric acid; Gln, glutamine; Glu, Glutamate; GPC, glycerophosphocholine; Lip, lipids; Myo, myoinositol; PC, phosphocholine; PE, phosphoethanolamine; PUFA, polyunsaturated fatty acids; Tau, taurine). Note that Cho, PC, GPC, PE, Etn, are clearly separable here due to the use of the 2D TOBSY method.

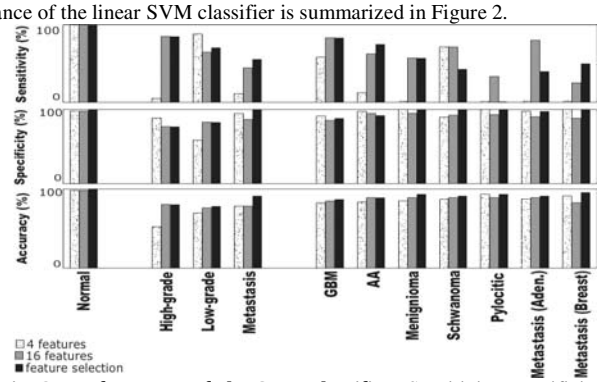

Fig. 2. Performance of the SVM classifier: Sensitivity, specificity, and accuracy for each feature selection scheme (4 features, MRMR, all features), tumor type (normal, high- and low-grade) and subtype [GBM, AA, Meningioma, Schwanoma, Pylocytic, Metastasis (Aden.) and Metastasis (Breast Cancer)]. Sensitivity was higher with SVM+MRMR (black bars) in high-grade biopsies. Specificity was higher in low-grade biopsies. Both sensitivity and specificity were affected in metastases, resulting in higher accuracy.

Discussion– Our novel and sensitive TOBSY ^{1}H HRMAS MRS allowed optimal identification of biomarkers and molecular and/or metabolic assessment of brain tumor biopsies ~2 mg. TOBSY MRS was performed at a low temperature (-8°C), maintaining tissue integrity thus allowing us to subsequently run histopathological, genomic, and/or proteomic analyses to construct molecular cancer signatures. We reliably detected at least 16 metabolites or biologically-relevant molecular species, producing spectra of excellent quality with little overlap. We also employed robust SVM classifiers to characterize brain tumor biopsies with high sensitivity, specificity, and accuracy. One principal finding is that our linear SVM with MRMR feature selection results in robust selection of the reduced feature space, and provides a data-driven criterion for generalizing the feature selection as the training set changes (e.g., increases in size as more data are added) leading to excellent classification performance, comparable to what is obtained when using the full-feature space. These findings enhance our knowledge from prior studies [8-12], since they suggest that the inclusion of more metabolites increases the performance of a given classification system, especially its sensitivity and accuracy. *In vivo* MRS at higher field strengths with 2D methods that allow the detection at least the same 16 metabolites that we reliably quantified, should enable clinicians to characterize and diagnose even inoperable brain tumors with high accuracy. This should allow appropriate therapy selection and the non-invasive monitoring of such therapies *in vivo*, thereby avoiding serial biopsies. Indeed, using molecular information to guide brain tumor therapy has been suggested [13].

References

1. Andronesi OC, Mintzopoulos D et al, J Magn Reson, 2008
2. Andronesi OC, Blekas K, Mintzopoulos D et al, Int J Onc 33: 1017-1025, 2008
3. Kleihues P, WK C, WHO. Lyon, FR: IARC Press, 2000, p. 314
4. Louis DN, Holland EC, and Cairncross JG, Am J Pathol 159: 779-786, 2001
5. Vapnik V, *Statistical Learning Theory*, Wiley-Interscience, 1998
6. Ding C and Peng H, J Bioinform Comput Biol 3: 185-205, 2005
7. Tugnoli V et al, Int J Mol Med 16: 301-307, 2005
8. Poptani H et al, J Cancer Res Clin Oncol 125: 343-349, 1999
9. Huang Y et al, Statistics in Medicine 22: 147–164, 2003
10. Devos A et al, J Magn Reson 170: 164-175, 2004
11. Opstad KS et al, NMR Biomed 20: 763-770, 2007
12. Poullet JB et al, Conf Proc IEEE Eng Med Biol Soc 2007: 5407-5410, 2007
13. Mischel PS and Cloughesy T, Nat Clin Pract Neurol 2: 232-233, 2006