Combining 'omics'; metabolic breast cancer subclass correlation with protein and gene expression subtypes

Tonje H. Haukaas[1,2], Leslie R. Euceda[1], Guro F. Giskeødegård[1], Marit Krohn[2,3], Ellen Schlichting[4], Rolf Kåresen[2,4], Sandra Nyberg[2,3], Kristine Kleivi Sahlberg[2,3], Anne-Lise Børresen-Dale[2,3], and Tone F. Bathen[1,2]

[1]Department of Circulation and Medical Imaging, Faculty of Medicine, NTNU, Trondheim, Norway, [2]K.G. Jebsen Center for Breast Cancer Research, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway, [3]Department of Genetics, Institute for Cancer Research Oslo University Hospital, The Norwegian Radium Hospital, Oslo, Norway, [4]Department of Surgery, Oslo University Hospital, Ullevål, Oslo, Norway

Target audience: Basic scientists investigating metabolomic biomarkers for cancer prognosis and treatment or breast cancer subtypes

Purpose: The heterogeneous biology of breast cancer leads to a high diversity in disease prognosis and treatment response even for patients with the same diagnosis and stage. Identifying underlying mechanisms contributing to this heterogeneity can reveal new cancer targets or clinically relevant subgroups. The molecular subtypes luminal A, luminal B, basal-like, Her2 enriched and normal-like, which have characteristic differences in gene expression [1], show correlation to tumor characteristics and clinical outcome, where basal-like has been associated with shortest survival times, and luminal A are associated with longer relapse-free survival [2]. Breast cancers can be classified into one of these molecular subtypes based on a 50-gene classifier called PAM50 [3]. In addition, reverse phase protein array (RPPA) results of breast cancer tumors have identified subgroups with differences in patient outcomes[4]. Here we combine these two classification systems with metabolic subclasses found from hierarchical clustering of MR metabolic profiles. This novel approach merging transcriptomic, proteomic and metabolic subtypes may improve the understanding of the heterogeneity of breast cancer.

Method: Tumor samples from 229 breast cancer patients were cut into three sections, where an adequate part of the mid piece was used for high resolution magic angle spinning magnetic resonance spectroscopy (HR MAS MRS). When possible, the remnants of the pieces were homogenized and divided into fractions used for DNA, RNA and protein extraction. 186 and 201 were classified by PAM50 and/or RPPA, respectively. Hierarchical clustering (Euclidean distance and Wards linkage) of MR spectral data was used to identify natural metabolic subclasses.

Results and Discussion: Three metabolic subclasses, c1 (n= 112), c2 (n = 59) and c3 (n=58) were identified and chosen for further evaluation (Figure 1). The mean spectra for the subclasses are shown in Figure 2. Relative levels of 18 metabolites were calculated by integration, out of which 16 were found to have significantly different levels between at least two of the metabolic subclasses ($p < 0.05$). c1 exhibited significantly higher levels of lactate and alanine compared to c3; c2 showed significantly higher levels of phosphocholine (PCh) and glycerophosphocholine (GPC) compared to both c1 and c3; glucose levels were significantly higher in c3 than in the other subclasses (Figure 1 and 2). These metabolites have been previously found to be important in breast cancer. Higher lactate levels have been associated with poor prognosis[5] while PCh and GPC correlate with increased proliferation rate and malignancy[6]. β-glucose correlates negatively to proliferation in accordance with the high energy demands of proliferating tumor cells [6]. Distribution of RPPA-subtypes was found to be significantly different between c1, c2 and c3 ($p = 1.94E-04$), while this was not the case for the PAM50-distribution. With regards to the RPPA subtypes, c1 contained most of the Reactive II samples (72%) while most Reactive I samples (56%) cluster in c3. Significance analysis of microarrays (SAM) revealed c1 and c3 to have a significantly higher expression of 752 and 209 mRNA probes, respectively, compared to c2, but no difference between c1 and c3 were observed (Figure 1E). Integrative meta-analysis of expression data (INMEX) [7] uncovered metabolic and transcriptomic differences in glycerophospholipid metabolism between c1 and c2, which corresponds with the significant higher levels of PCh and GPC detected in the latter subclass. Gene set enrichment analysis [8,9] revealed that pathways related to collagens and extracellular matrix were altered in c1 and c3 when compared to c2. 148 of the mRNA probes were found in both comparisons. This might explain why most of the samples with reactive I and II subtypes clustered in c3 and c1, respectively, as characteristic proteins of these RPPA subtypes, such as collagen, are thought to be produced by the tumor microenvironment and cancer-activated fibroblasts[10].

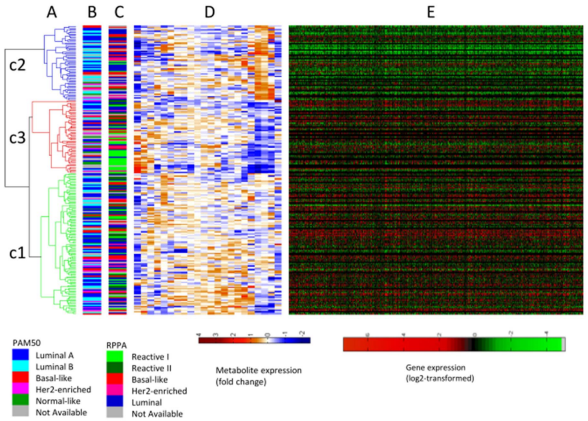

Figure 1: **Main differences between metabolic subtypes**. (A) Hierarchichal clustering of HR MAS MR spectra. (B) PAM50-subtype. (C) RPPA-subtype. (D) Fold change in expression levels of glucose, acetate, myo-inositol, choline, taurine, ascorbate, glutathione, glutathione 2, tyrosine, glutamine&glutamate, glutamine, glutamate, alanine, lactate, succinate, glutamine2, glycine, creatine, phosphocholine, total choline, glycerophosphocholine, and scyllo-inositol. (E) Gene expression levels (quantile normalized, log 2 transformed) for genes significant ($q < 0.01$) between two or more of the groups.

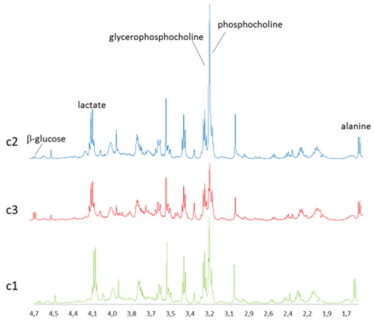

Figure 1: Mean spectra for the three metabolic subclasses.

To conclude, the combination of different molecular levels from the same samples provides insight into the heterogeneity of breast cancers, with differences in metabolite levels and genetic pathways, and the identified subclasses were highly correlated to RPPA previously defined subgroups.

References: **1]** Perou, C.M., et al.,Nature, 2000. 406(6797): p. 747-752. **2]**Sorlie, T., et al., Proc Natl Acad Sci, 2001; 98(19):10869-10874. **3]** Parker, J.S., et al., J Clin Oncol, 2009; 27(8):1160-1167. **4]** Hennessy, B.T., et al., Clin. proteomics, 2010. 6(4): p. 129-151. **5]** Giskeodegard, G.F., et al., NMR Biomed, 2012. 25(11): p. 1271-9. **6]** Moestue, S., et al., Curr Top Med Chem, 2011; 11(1):2-26. **7]** Xia, J., et al., Nucleic Acids Res, 2013. 41(Web Server issue): p. W63-70. **8]** Subramanian, A., et al., Proc Natl Acad Sci, 2005. 102(43): p. 15545-15550. **9]** Mootha, V.K., et al., Nat. Genet., 2003. 34(3): p. 267-273. **10]** TCGA Network, Nature, 2012. 490(7418): p. 61-70