Molecular variation between spontaneous primary breast cancers and breast cancers arising in the setting of previous thoracic irradiation in survivors of childhood malignancies

Working Groups:
Primary: Genetics
Secondary: Second Malignancy

Investigators:

Christopher Pelloski, MD
Ohio State Medical Center/Nationwide Children's Hospital
Department of Radiation Oncology
300 W. 10th Avenue
Room 094F CHRI
Columbus, OH 43210
Phone: 614-366-2729
Fax: 614-293-7602
Christopher.pelloski@osumc.edu

Allison Quick, MD
Department of Radiation Oncology
Ohio State Medical Center
300 W. 10th Avenue
Room 083A CHRI
Columbus, OH 43210
Phone: 614-293-8595
Fax: 614-293-4044
Allison.quick@osumc.edu

Denise Fabian
Department of Radiation Oncology
Ohio State Medical Center
300 W. 10th Avenue
Columbus, OH 43210
Denise.fabian@osumc.edu

Sue Hammond, MD
Nationwide Children's Hospital
Department of Pathology and Laboratory Medicine
700 Children’s Drive, Columbus OH 43205
Phone: 614-722-5310
Fax: 614-722-5308

Sue.hammond@nationwidechildrens.org

Kun Huang, Ph.D.
Department of Biomedical Informatics
The Ohio State University Medical Center
Phone: 614-292-5607
E-mail: kun.huang@osumc.edu

Les L Robison, PhD
St. Jude Children's Research Hospital
Memphis, TN
Email: Les.Robison@stjude.org

Joseph P Neglia, MD, MPH
Department of Pediatrics
University of Minnesota
Email: jneglia@umn.edu

Smita Bhatia, MD, MPH
Department of Population Sciences
City of Hope
Duarte, CA
Phone: 626-471-7321
Email: sbhatia@coh.org

Marilyn Slovak

Wendy Leisenring, ScD (CCSS statistician)
Cancer Prevention and Clinical Statistics
Fred Hutchinson Cancer Research Center
Seattle, WA
wleisenr@fhcrc.org

Background and Rationale:

Female survivors of childhood cancers, who receive thoracic radiation as part of their therapy, are at a significantly increased risk of developing breast cancer. The most well-studied cohorts consist of childhood Hodgkin's lymphoma (HL) survivors. In a previous Childhood Cancer Survivor Study Cohort, 33.8% of HL survivors developed a second solid malignancy, the most common being breast cancer (38%). The cumulative incidence of female breast cancer by age 40 was 12.9% for survivors of HL who received radiation. [1] For female childhood cancer survivors, 68% of women who developed a second breast cancer were HL survivors and 97% of these survivors were treated with chest radiation.[2] Secondary breast cancers most commonly develop within the previous HL radiation field and are likely due to the sensitivity of developing breast tissue to

ionizing radiation.[3] Pathological characteristics are thought to be similar to primary breast cancers;[4-5] however, some studies have identified pathologic and molecular differences.[6-7]

Primary breast cancer is a heterogeneous disease, with multiple pathologic and molecular subtypes and varying clinical outcomes. Currently, immunohistochemical (IHC) markers for estrogen receptors (ER) and progesterone receptors (PR) and the HER2 epidermal growth factor receptor are used to classify 4 different clinical subtypes of breast cancer: 1) ER+ and/or PR+, HER2-; 2) ER+ and/or PR+, HER2+; 3) ER-, PR-, HER2-; and 4) ER-, PR-, HER2+. However, global gene expression profiling studies have recently identified 4 main molecular signatures that correspond to the 4 clinical subtypes respectively: luminal A and B (hormone receptor positive), basal-like ("triple-negative"), and HER2 enriched (HER2+) (**Figure 1**).[8-11] A fifth subtype, the normal breast tissue-like group, was also identified. However, the clinical significance of this group is not known and may represent poorly sampled tumor.[8] Clinical outcomes vary between the different subtypes; the luminal A group has the best prognosis and the basal-like and HER2 enriched subtypes are associated with an inferior overall and disease free survival.[8, 10-11]

To our knowledge, no studies have identified the gene expression profile of breast cancers arising in the context of childhood exposure to thoracic radiotherapy for previous malignancies. Since each molecular subtype carries a different clinical outcome and treatment strategy, this knowledge may guide therapeutic decisions and estimation of patient prognosis as well as improve our understanding of radiation-induced oncogenesis in breast tissue.

Specific aims/objectives/research hypotheses/methods:

Our overriding hypothesis is that radiation interacts, enhances, and potentially accelerates the individual patient's genetic predisposition to develop breast cancer, but does not fundamentally alter the molecular phenotype. To test this hypothesis, we plan to obtain the 54 (or more if available) formalin-fixed, paraffin-embedded tissue samples of the secondary breast cancer cases from the CCSS bio-repository from women who received thoracic radiation. We plan to subject these samples to RNA extraction, using established protocols that yield expression profiles that are highly concordant with frozen tissue.[12-13] This mRNA will then be assayed for quality and quantity and amplified accordingly.[14] The resultant cDNA will then be hybridized, at our core facility, using the Affymetrix U133A gene chip. Using the GeneSpring software platform, gene expression data will then be analyzed using supervised methods (which will query the expression levels of specific, hallmark genes that define the molecular subtypes of breast cancer) as well as unsupervised methods (within the set of 54 cases as well as in conjunction with 1,619 general population cases) to determine the prevalence of known, and potentially novel, molecular subclasses of these secondary tumors. Once the subclass designation is established for the 54 secondary breast cancers, the patient demographics and overall survival time after diagnoses will be compared to their spontaneous/general population

counterparts. Generating this working data will enable us to answer the following clinically important questions, which represent our specific aims:

1. Specific Aim 1: What is the prevalence/distribution of Luminal A, Luminal B, Basal, Her2-enriched and breast-like breast cancer subtypes among those women who received thoracic radiation during childhood for their pediatric malignancy versus that observed in the general population of spontaneous breast cancers? Consistent with our overriding hypothesis is that the prevalence of these molecular subtypes will approximate that observed in the general, spontaneous breast cancer population. Therefore, assuming radiation induced breast cancers are representative of the general population, approximately 50% will be Luminal A, 20% will be Luminal B, 20% will be Basal, and 10% will be Her 2+.

2. Specific Aim 2: Is there a clinical outcome difference in terms of overall survival after diagnosis within the molecular subclasses of the 54 cases of secondary breast cancers from the childhood cancer survivors and compared to spontaneous breast cancers from the general population cases with the same molecular subclass? We believe the answer to this question will determine if radiation has an independent role in expediting the development of various breast cancers and/or augmenting the biologic lethality of the disease.

If our overriding hypothesis is incorrect, then 2 other potential phenomena may be identified with this data:

1. Radiation to the chest during childhood induces a specific type of one of the established molecular subclasses of breast cancer, which may or may not have the same virulence as seen in the general breast cancer population.

2. A novel molecular subclass(es), unique to patients with radiation to the chest during childhood, is elucidated. If so, multiple questions can be answered using this data:
    i. What are the hallmark genes that define this group?
    ii. What are the patient demographics at time of onset? And
    iii. What is the expected survival time within this group(s)?

Analysis framework:

Outcome of interest

Primary endpoint:

- The patterns of gene expression and distribution of molecular subtypes of radiation induced breast cancer compared to spontaneous/primary breast cancer

Secondary endpoints:

- Overall survival: Time from diagnosis of breast cancer to death of any cause
- Patient demographics at the time of breast cancer diagnosis.

Subject population:

Inclusion criteria

1. Cases: Survivors from the Childhood Cancer Survivor Study Cohort who subsequently developed breast cancer after thoracic radiation, with available breast tissue
2. Controls: 1,619 patients with spontaneous primary breast cancers, without previous cancer diagnosis or treatment, obtained from previously published research studies for which the gene expression data is publically available, will comprise the general, spontaneous breast cancer population. The previously published gene expression data is analyzed for subtype composition and is presented in **Figure 2**.[15-22]
    a. All ages
    b. All stages
    c. Tumor sample obtained from biopsy or surgical specimen prior to any non-surgical therapy
    d. Biopsy proven invasive ductal, lobular, mucinous, tubular, colloid or papillary (i.e. epithelial) carcinoma of the breast
    e. Primary and secondary breast cancer with any hormone status, including ER+, PR+/-, Her2neu-; ER+,PR+/-, Her2neu +; ER-,PR-, Her2neu-; ER-/PR-, Her2neu +

Exclusion criteria

1. Previously published research using gene expression data sets for primary breast cancers previously treated with chemotherapy, hormone therapy, or radiation prior to analysis will not be used for the control population
2. Previously published research using gene expression data from cell lines and animal studies will not be used for the control population
3. Non-epithelial breast malignancies, including lymphoma or sarcoma of the breast and melanoma and non-melanomatous skin cancers
4. In situ carcinomas of the breast

Materials/Data needed from the CCSS

1. Adequate secondary breast cancer tissue, from the study patients, to yield sufficient, amplifiable mRNA for eventual hybridization with the Affymetrix U133A Gene Chip.
    a. N=54 (or more if possible)
    b. At least five 10-micron scrolls (or equivalent quantity) of tumor tissue, verified by corresponding H&E slide, for each sample
2. Clinical data on study patients:
    a. Age at childhood cancer diagnosis
    b. Childhood primary tumor diagnosis
    c. Childhood radiotherapy details (including total body irradiation and pelvic radiotherapy)
        i. Dose and fractionation to irradiated site

ii. Radiation dose to the chest
iii. Age at start of radiation

d. Age at breast cancer diagnosis
e. Menopausal status at breast cancer diagnosis
f. Age at death or last follow-up with current vital status for survival calculations
g. Use of chemotherapy for treatment of the primary cancer

Statistical and Analytic Plans:

The first layer of analysis will consist of determining the distribution of the molecular subtypes of the secondary breast cancers. The RNA will be isolated from the study samples and in a manner similar to the clustering analysis of the control population, the gene expression status of the 249 genes that define the molecular subclasses will be performed in an unsupervised manner on the 54 study and 1,619 control cases together. To account for batch-effect and paraffin extraction, log-transformation and normalization using the RMA method will be performed, again using GeneSpring software, with all cases included. Genes will undergo an independent QA process (to determine if there is signal susceptibility to the formalin fixation and parafinization process) and will be excluded from the final clustering analysis. The placement of the study cases within the cluster groups, as defined predominantly by the control population, will determine the subclass assignment. A chi-square test will be used to compare the distribution of molecular subtypes between the primary and secondary breast cancers.

The raw data from the .cel files of the study population will also be compared. This will be especially important in the event of the formation of a separate clustering group by the radiation-induced cases and a novel subtype is identified. From this analysis, the genes that define this new subclass will be identified and reported. This approach will also allow the identification of gene expression, unique to previously radiated breast tissue, which spans across all subtypes.

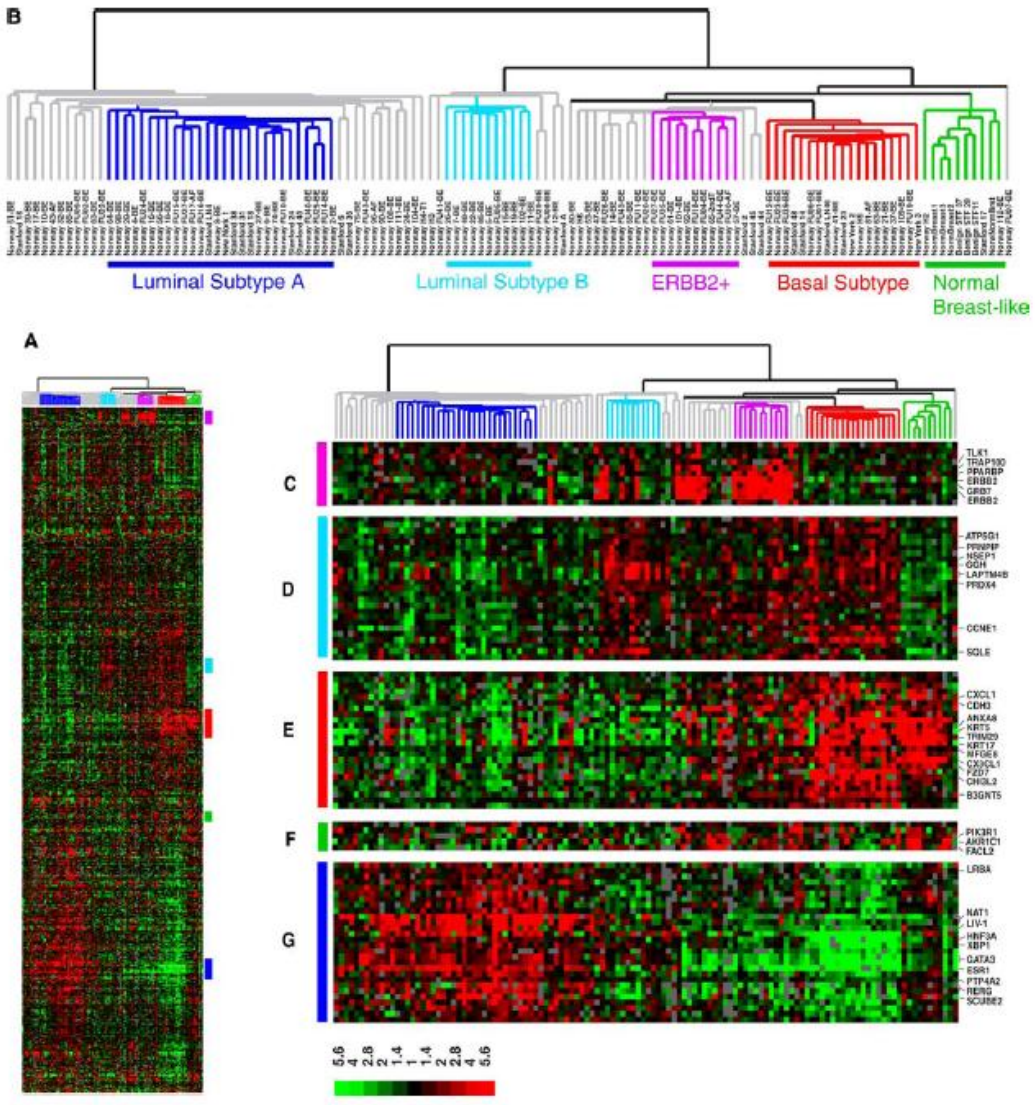
To study our second aim, the molecular subtype status of the study population will be analyzed as a clinicopathologic factor for it's association with survival as well as with other factors such as age at childhood cancer diagnosis, type of childhood primary malignancy, age at time of radiation for childhood malignancy, radiation dose to the chest during childhood, total radiation dose received during childhood, use of chemotherapy for treatment of childhood malignancy, age and menopausal status at breast cancer diagnosis, and stage at breast cancer diagnosis. Correlative analysis (Spearman's rho, t-tests, ect.) will be used to test for univariable associations between the molecular subtype and the cross-sectional demographic variables. For the time-to-event outcomes such as survival and time to onset of breast cancer the log-rank test will be used to determine univariate associations of the molecular subtype (key risk factor of interest) with these outcomes intervals. Additional risk factors such as age at childhood cancer diagnosis, age at time of radiation for childhood malignancy, and radiation dose to the chest during childhood will also be evaluated similarly. Cox proportional hazards models will be used

to evaluate multivariable associations between molecular subtype and each of the time-to-event outcomes, adjusting for other factors found to be significant in univariable analyses, or considered *a priori* to be important to include in the model (e.g. age at diagnosis, breast cancer stage, etc.)

The CCSS database represents the only chance in which the issue of the breast cancer molecular subclass prevalence, in the setting of previous thoracic radiation, could possibly be addressed. If any single institution attempts to do this, there would not be enough cases to draw a statically significant conclusion. Fortunately, there are 54 cases within the CCSS database, and this is sufficient to determine if the 5 molecular subtypes are evenly distributed, or predominately composed of 1 or 2 subtypes. In the most diffuse scenario, if there are at least 10 samples per subtype (5), there will be a >80% power to detect a 2-fold difference between any of the two groups using Chi-squared analysis. A coefficient of variation of 50% and alpha = 0.005 (.05/10, to adjust for 10 comparisons between 5 groups) will be used for the calculations. Since we expect there to be 50% with the Luminal A subtype designation, there will be even greater power to make this conclusion statistically valid. It should be iterated, though, that regardless of statistical power, this is a retrospective study and the large number of subjects from the CCSS provide the only means for attempting to answer this scientific question.

.

Figures:



Hierarchical clustering of 115 tumour tissues and seven non-malignant tissues using the "intrinsic" gene-set. (A) A scaled-down representation of the entire cluster of 540 genes and 122 tissue samples based on similarities in gene expression. (B) Experimental dendrogram showing the clustering of the tumours into five subgroups. Branches corresponding to tumours with low correlation to any subtype are shown in grey. (C) Gene cluster showing the *ERBB2* oncogene and other co-expressed genes. (D) Gene cluster associated with luminal subtype B. (E) Gene cluster associated with the basal subtype. (F) A gene cluster relevant for the normal breast-like group. (G) Cluster of genes including the oestrogen receptor (*ESRI*) highly expressed in luminal subtype A tumours. Scale bar represents fold-change for any given gene relative to the median level of expression across all samples. PNAS, July 8, 2003, vol. 100, no. 14, 8418–8423. Copyright (2003) National Academy of Sciences, USA.

Figure 1: **Molecular subtypes of breast cancer identified.** Gene expression profiles of primary breast cancers demonstrating 5 distinct molecular subtypes: luminal A, luminal B, Basal, ERBB2+ (HER2 enriched [ERBB2+]), and normal breast-like. From Perou CM et al. Molecular portraits of human breast tumours. Nature 2000;406:747-52.

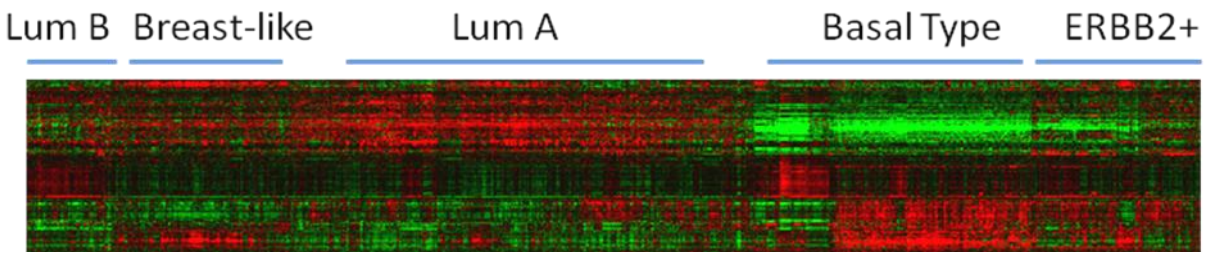


Figure 2: **Our preliminary data.** Supervised clustering analysis using 249 genes identified as hallmark genes in determining the established molecular subtypes applied to 1,619 cases from publicly available U133A-derived .cel files,[15-22] after log-transformation and normalization using the RMA method. The approximate subclass designation is noted (scaling precludes detailed dendrogram). Collectively, these cases reflect an adequate, heterogeneous control population. Data obtained from radiation-induced breast cancers will be analyzed within the context of this general population of spontaneous breast cancers.

Special considerations:
The principle investigator has significant experience in extracting RNA and DNA from paraffin-embedded tissues for global profiling analyses.[23-25]

References:

1. Meadows AT, Friedman DL, Neglia JP, et al. Second Neoplasms in Survivors of Childhood Cancer: Findings From the Childhood Cancer Survivor Study Cohort. Journal of Clinical Oncology 2009;27:2356-62.
2. Kenney LB, Yasui Y, Inskip PD, et al. Breast cancer after childhood cancer: a report from the Childhood Cancer Survivor Study. Annals of Internal Medicine 2004;141:590-7.
3. Travis LB, Hill DA, Dores GM, et al. Breast Cancer Following Radiotherapy and Chemotherapy Among Young Women With Hodgkin Disease. JAMA: The Journal of the American Medical Association 2003;290:465-75.
4. Wolden SL, Hancock SL, Carlson RW, Goffinet DR, Jeffrey SS, Hoppe RT. Management of Breast Cancer After Hodgkin's Disease. Journal of Clinical Oncology 2000;18:765.
5. Yahalom J, Petrek J, Biddinger P, et al. Breast cancer in patients irradiated for Hodgkin's disease: a clinical and pathologic analysis of 45 events in 37 patients. Journal of Clinical Oncology 1992;10:1674-81.
6. Gaffney DK, Hemmersmeier J, Holden J, et al. Breast cancer after mantle irradiation for Hodgkin's disease: correlation of clinical, pathologic, and molecular features including loss of heterozygosity at BRCA1 and BRCA2. International Journal of Radiation Oncology*Biology*Physics 2001;49:539-46.
7. Janov AJ, Tulecke M, O'Neill A, et al. Clinical and Pathologic Features of Breast Cancers in Women Treated for Hodgkin's Disease: A Case-Control Study. The Breast Journal 2001;7:46-52.
8. Sørlie T. Molecular portraits of breast cancer: tumour subtypes as distinct disease entities. European Journal of Cancer 2004;40:2667-75.
9. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406:747-52.
10. Sørlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America 2001;98:10869-74.
11. Sørlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the United States of America 2003;100:8418-23.
12. Roberts L, Bowers J, Sensinger K, Lisowski A, Getts R, Anderson MG. Identification of methods for use of formalin-fixed, paraffin-embedded tissue samples in RNA expression profiling. Genomics 2009;94:341-8.
13. Jacobson TA, Lundahl J, Mellstedt H, Moshfegh A. Gene expression analysis using long-term preserved formalin-fixed and paraffin-embedded tissue of non-small cell lung cancer. International Journal of Oncology 2011;38:1075-81.
14. Turner L, Heath JD, Kurn N. Gene expression profiling of RNA extracted from FFPE tissues: NuGEN technologies' whole-transcriptome amplification system. Methods in molecular biology 2011;724:269-80.
15. Schmidt M, Bohm D, von Torne C, et al. The Humoral Immune System Has a Key Prognostic Impact in Node-Negative Breast Cancer. Cancer Research 2008;68:5405-13.

16. Sotiriou C, Wirapati P, Loi S, et al. Gene Expression Profiling in Breast Cancer: Understanding the Molecular Basis of Histologic Grade To Improve Prognosis. Journal of the National Cancer Institute 2006;98:262-72.
17. Desmedt C, Piette F, Loi S, et al. Strong Time Dependence of the 76-Gene Prognostic Signature for Node-Negative Breast Cancer Patients in the TRANSBIG Multicenter Independent Validation Series. Clinical Cancer Research 2007;13:3207-14.
18. Farmer P, Bonnefoi H, Becette V, et al. Identification of molecular apocrine breast tumours by microarray analysis. Oncogene 2005;24:4660-71.
19. Pawitan Y, Bjohle J, Amler L, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. Breast Cancer Research 2005;7:R953 - R64.
20. Popovici V, Chen W, Gallas B, et al. Effect of training-sample size and classification difficulty on the accuracy of genomic predictors. Breast Cancer Research 2010;12:R5.
21. Tabchy A, Valero V, Vidaurre T, et al. Evaluation of a 30-Gene Paclitaxel, Fluorouracil, Doxorubicin, and Cyclophosphamide Chemotherapy Response Predictor in a Multicenter Randomized Trial in Breast Cancer. Clinical Cancer Research;16:5351-61.
22. Wang Y, Jan GMK, Yi Z, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. The Lancet 2005;365:671.
23. Noushmehr H, Weisenberger DJ, Diefes K, et al. Identification of a CpG Island Methylator Phenotype that Defines a Distinct Subgroup of Glioma. Cancer Cell 2010;17:510-22.
24. Colman H, Zhang L, Sulman EP, et al. A multigene predictor of outcome in glioblastoma. Neuro-Oncology 2010;12:49-57.
25. Rivera AL, Pelloski CE, Gilbert MR, et al. MGMT promoter methylation is predictive of response to radiotherapy and prognostic in the absence of adjuvant alkylating chemotherapy for glioblastoma. Neuro-Oncology 2010;12:116-21.