Germline microRNA-based variants predicting late radiation toxicity in breast cancer

J. Weidhaas[1], K. McGreevy[2], S. McCloskey[3], J. Hegde[1], M. Steinberg[1], J. Baker[4], M. DiNome[5], C. Thompson[6], A. Kusske[6], J. Glaspy[7], R. Callahan[7], K. McCann[7], M. Teshome[7], N. Mcandrew[7], A. Bardia[7], D. Telesca[8];
[1]Radiation Oncology, UCLA, Los Angeles, CA, [2]Biostatistics, MiraDx, Los Angeles, CA, [3]Radiation Oncology, UCLA, Santa Monica, CA, [4]Surgical Oncology, UCLA, Los Angeles, CA, [5]Surgery, Duke, Durham, NC,
[6]Breast Surgery, UCLA, Santa Monica, CA, [7]Medical Oncology, UCLA, Los Angeles, CA, [8]Department of Biostatistics, UCLA, Los Angeles, CA.

Background

Germline microRNA-disrupting variants (mirSNPs) have been proven to predict late radiation toxicity, which generally presents as fibrosis, in several cancer types. mirSNP signatures of toxicity are radiation fractionation (size of delivered dose) dependent, meaning that identifying patients at risk of toxicity from one form of radiation treatment affords the opportunity to choose safer alternatives for their management. As radiation is an integral part of breast cancer management, with several fractionation approaches gaining acceptance, we tested the hypothesis that mirSNP signatures could predict late radiation toxicity for breast cancer patients. We also investigated the applicability of existing late toxicity signatures used in prostate cancer (PROSTOX) and evaluated if toxicity signatures could be improved with additional mirSNPs.

Cohort Demographics

The study cohort consisted of 161 breast cancer patients treated with conventionally fractionated radiation (CFRT, 39.8%) or moderately hypo-fractionated radiation (MHFRT, 16-20 fractions, 60.2%).

Most of the training cohort was White (93.4%) while the test cohort was more diverse (42.5%). Most of the training cohort was female (98.1%), and most had radiation to an intact breast after breast conservation (88.2%) versus post-mastectomy (12.8%).

Table 1. Cohort Information

	Train (N=121)	Test (N=40)
Age at Treatment		
Mean (SD)	61.9 (11.5)	59.3 (12.1)
Race		
Asian	3 (2.5%)	7 (17.5%)
Black	1 (0.8%)	4 (10.0%)
Other	4 (3.3%)	6 (15.0%)
White	113 (93.4%)	23 (57.5%)
Ethnicity		
Hispanic	14 (11.6%)	2 (5.0%)
Not Hispanic	107 (88.4%)	38 (95.0%)
Fractionation		
CFRT	47 (38.8%)	17 (42.5%)
MHFRT	74 (61.2%)	23 (57.5%)

Methods

DNA was isolated from buccal swabs and tested for over 120 mirSNPs. We split the cohort into 75% train (n=121) and 25% test sets (n=40). We applied a previously developed mirSNP signature of late grade $\geq$ 2 CFRT radiation toxicity defined in prostate cancer patients (PROSTOX standard)[1]. We then explored the addition of clinical factors to this signature to see if any improved predictive accuracy including post-mastectomy radiation or implant, chemotherapy treatment, alcohol and smoking history, autoimmunity, and DNA repair mutant status.

We used transfer learning techniques [3] to incorporate germline mirSNPs marginally associated with late toxicity (Fisher or Jonckheere-Terpstra $p < 0.1$) in our breast cohort alongside the PROSTOX standard signature in a boosted tree model framework to make our new CFRT and MHFRT toxicity signature for breast cancer, referred to as BREASTOX. We compared performance of models using standard performance metrics stratified by training and test sets.

BREASTOX Performance

The PROSTOX standard late radiation toxicity signature achieved robust predictive performance for late grade $\geq$ 2 toxicity in our training and test breast patient cohorts, with high specificity (**Table 2**). Notably, all patients with grade 0 late toxicity were accurately identified as low risk, and the model predicted grade 3 late toxicity more accurately than grade 2 (60% vs. 30%, p=0.317 for difference in proportions).

We found that none of the clinical factors improved the performance of the PROSTOX *standard* radiation toxicity signature in our cohort. In contrast, adding just 6 additional mirSNPs, including MAT1A and *KRAS*, boosted the AUC to 0.890 and increased the sensitivity from 0.50 (PROSTOX *standard* alone) to 0.828 (BREASTOX) (**Figure 1A**).

Table 2. Performance metrics of PROSTOX *standard* and BREASTOX in predicting late toxicity

Breast Data	mirSNP Model	Sensitivity	Specificity	PPV	NPV	AUC
Training (n=121)	BREASTOX, LOOCV	0.828	0.848	0.632	0.940	0.890
Test (n=40)	BREASTOX	0.857	0.879	0.600	0.967	0.918
Training (n=121)	PROSTOX standard	0.500	0.956	0.790	0.853	0.728
Test (n=40)	PROSTOX standard	0.429	0.970	0.750	0.889	0.706
Training (n=121)	PROSTOX standard + Clin Vars	0.517	0.902	0.625	0.856	0.710

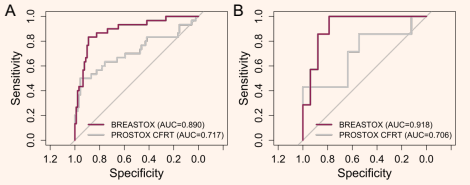


Figure 1. Performance of PROSTOX *standard* compared to newly constructed BREASTOX. A) Training data (n=121) with LOOCV metrics for BREASTOX. B) Held out test data (n=40) applied to BREASTOX and PROSTOX *standard*

Important Genetic Factors for Nodal Burden

Both MAT1A and *KRAS* variants appeared protective (**Table 3**): the rate of late toxicity dropped from 38% to 7% (wild type to mutant MAT1A, Fisher p=0.0005) and from 27% to 5% (wild type to heterozygous *KRAS*, Fisher p=0.041).

In women, testosterone is primarily metabolized into estradiol (E2), whereas in men it is largely converted into dihydrotestosterone (DHT). Notably, treatment of ER-positive breast cancer cells with estradiol has been shown to increase *MAT1A* mRNA expression[4], whereas treatment with DHT has not been shown to have this effect. This observation suggests a potential sex-specific regulatory pathway in which *MAT1A* plays a more prominent role in females than in males. Thus, the identification of a mirSNP within *MAT1A* as the strongest predictor in our CFRT/MHFRT model for breast cancer toxicity is biologically compelling and consistent with sex hormones.

Table 3. Proportion of Patients with Late Toxicity by Genotype

mirSNP	Homo-zygous wildtype	Hetero-zygous	Homo-zygous mutant	Fisher p-value
MAT1A, A>G, intronic	37.5%	40.4%	7.1%	0.0005
KRAS, A>C, 3'UTR	27.5%	5.3%	-	0.0414

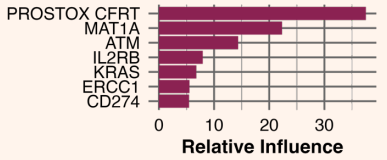


Figure 2. Importance of Variables in BREASTOX. Relative influence is the relative improvement in the loss function via boosted tree modeling. PROSTOX *standard* improves late toxicity prediction by 38% and MAT1A by 22%.

Conclusions

Our findings demonstrate that germline mirSNP signatures not only predict late radiation toxicity within a specific cancer type but also generalize across tissues and fractionation schemes. Remarkably, the PROSTOX *standard* model, originally developed for prostate cancer, retained strong predictive performance in breast cancer patients treated with both CFRT and MHFRT, highlighting the conserved biology underlying late radiation toxicity.

Building upon this foundation, we show that adding just six additional mirSNPs tailored to breast cancer further improves predictive accuracy in both training and independent test sets—culminating in the new BREASTOX model.

The identification of sex-specific regulators such as *MAT1A* and the *KRAS*-variant, whose variants are protective and mechanistically linked to estrogen-driven pathways, reinforces the biological plausibility of our approach. Together, these findings underscore the potential of mirSNP signatures to guide truly personalized, biology-driven radiation treatment decisions for breast cancer patients.

Contact

Joanne B. Weidhaas, MD, PhD, MSM
Organizations: University of California, Los Angeles
Email: jweidhaas@mednet.ucla.edu

References

1. Kishan AU, Marco N, et al. Germline variants disrupting microRNAs predict long-term genitourinary toxicity after prostate cancer radiation. Radiother Oncol. 2022 Feb;167:226-232.
2. Kishan AU, et al. Validation and Derivation of MicroRNA-based Germline Signatures Predicting Radiation Toxicity in Prostate Cancer. Clin Cancer Res. April 2025.
3. Hosna, A, et al. Transfer learning: a friendly introduction. *J Big Data* **9**, 102 (2022).
4. Wang CY, et al. Gene signatures and potential therapeutic targets of amino acid metabolism in estrogen receptor-positive breast cancer. Am J Cancer Res. 2020 Jan 1;10(1):95-113.