Title

Murine Microenvironment Metaprofiles Associate with Human Cancer Etiology and Intrinsic Subtypes

Journal

Clinical Cancer Research, 19(6)

ISSN

1078-0432

Authors

Nguyen, David H
Fredlund, Erik
Zhao, Wei
et al.

Publication Date

2013-03-15

DOI

10.1158/1078-0432.ccr-12-3554

Published in final edited form as:
Clin Cancer Res. 2013 March 15; 19(6): 1353–1362. doi:10.1158/1078-0432.CCR-12-3554.

Murine microenvironment metaprofiles associate with human cancer etiology and intrinsic subtypes

David H. Nguyen[1], **Erik Fredlund**[2], **Wei Zhao**[3], **Charles M. Perou**[3], **Allan Balmain**[2], **Jian-Hua Mao**[4], and **Mary Helen Barcellos-Hoff**[1,5]

[1]Department of Radiation Oncology, New York University School of Medicine, New York, NY 10016, USA

[2]Cancer Institute, University of California, San Francisco, San Francisco, CA 94143, USA

[3]Department of Pathology and Laboratory Medicine, University of North Carolina, Chapel Hill, NC 27599, USA

[4]Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley CA 94720, USA

Abstract

Purpose—Ionizing radiation is a well established carcinogen in rodent models and a risk factor associated with human cancer. We developed a mouse model that captures radiation effects on host biology by transplanting unirradiated *Trp53* null mammary tissue to sham or irradiated hosts. Gene expression profiles of tumors that arose in irradiated mice are distinct from those that arose in naïve hosts. We asked whether expression metaprofiles could discern radiation-preceded human cancer or be informative in sporadic breast cancers.

Experimental Design—Affymetrix microarray gene expression data from 56 *Trp53* null mammary tumors were used to define gene profiles and a centroid that discriminate tumors arising in irradiated hosts. These were applied to publicly available human cancer data sets.

Results—Host irradiation induces a metaprofile consisting of gene modules representing stem cells, cell motility, macrophages and autophagy. Human orthologs of the host irradiation metaprofile discriminated between radiation-preceded and sporadic human thyroid cancers. An irradiated host centroid was strongly associated with estrogen receptor negative breast cancer. When applied to sporadic human breast cancers, the irradiated host metaprofile strongly associated with basal-like and claudin-low breast cancer intrinsic subtypes. Comparing host irradiation in the context of TGFβ levels showed that inflammation was robustly associated with claudin-low tumors.

Conclusions—Detection of radiation-preceded human cancer by the irradiated host metaprofile raises possibilities of assessing human cancer etiology. Moreover, the association of the irradiated host metaprofiles with estrogen receptor negative status and claudin-low subtype suggests that host processes similar to those induced by radiation underlie sporadic cancers.

Keywords

ionizing radiation; breast cancer; intrinsic subtypes; etiology

---

[5]Corresponding Author: Mary Helen Barcellos-Hoff, Department of Radiation Oncology, New York University, School of Medicine, 566 First Avenue, New York, NY10016, Phone: 212-263-3021, Fax: 212-263-6274, mhbarcellos-hoff@nyumc.org.

Introduction

Ionizing radiation is one of very few environmental exposures unequivocally associated with increased cancer risk in humans (1), particularly in thyroid and breast cancer following exposure at a young age (2–3). Breast cancer increases in women who survived the atomic bombs (4), received diagnostic radiation for tuberculosis (5), or were treated with radiation for benign breast disease (6). Twenty percent of women treated with radiation for Hodgkin's lymphoma develop breast cancer before the age of 40 (7). Breast cancer is a complex disease that consists of at least six intrinsic subtypes identified by gene expression profiling (8–10) that can be prognostic (11–12). Recent studies suggest that prior exposure to radiation promotes aggressive, estrogen receptor (ER) negative tumors(13–15).

Radiation is a complete carcinogen able to both initiate and promote cancer. Initiation, thought to be due to oncogenic mutations from mis-repaired double DNA breaks, is widely believed to be the critical event for radiation carcinogenesis (16), however host systemic and stromal responses can also contribute to radiation's carcinogenic potential (17–20). To test whether host biology contribute to radiation carcinogenesis, we established a radiation chimera model that separates radiation effects on the host from those on the target epithelium (21). In this model, the mammary gland is cleared of endogenous epithelium of host mice, which are subsequently irradiated and then transplanted orthotopically with non-malignant *Trp53* null mammary tissue (22), which has many similarities to human breast cancer, including progression from pre-neoplastic lesions to ductal carcinoma *in situ* to tumors with diverse histopathologies (23–24). Even though host irradiation occurred many months before tumor development and the mammary epithelium was never irradiated, the course of *Trp53* null carcinogenesis is significantly altered by host irradiation as evidenced by decreased tumor latency and more rapid tumor growth rate. Unexpectedly, host irradiation also increased the proportion of ER negative tumors. Expression profiles of *Trp53* null tumors arising in an irradiated host compared to those arising in non-irradiated hosts were also distinct, suggesting that the biology elicited by radiation has long lasting effects on tumor development (22). Network analysis of the irradiated host signature implicated two critical factors, TGFβ and mammary stem cells that were validated with additional experiments, demonstrated at least two distinct mechanisms by which the irradiated microenvironment promotes breast cancer (22).

Detailed understanding of the basis for cancer characteristics and clinical behaviors in irradiated populations could improve risk predictions, and may uncover means to reduce risk. We speculated that expression metaprofiles might discern radiation-preceded human cancer and be informative in sporadic breast cancers. We used bioinformatics to evaluate the value of murine irradiated host signatures for classifying radiation-preceded human cancers and its associations with sporadic breast cancer.

Methods

Data from Affymetrix mouse Genechip MG-430 2.0 arrays from our prior study (22), GSE18216 NCBI Gene Expression Omnibus database accession number, were used, in addition to 8 addition samples (merged under GSE42742). Background was normalized using robust multichip average algorithm (25), R software v2.10.1, with widgets specific to the Affymetrix platform. Unsupervised hierarchical clustering using Gene Cluster v3.0 software was visualized using Java TreeView v1.1.4r3 software. Data was mean centered; gene clustering was done by an uncentered-correlation and array clustering was done by Spearman Rank correlation; under complete linkage. Pathways were identified with Ingenuity Pathway Analysis, or ConceptGen (http://conceptgen.ncibi.org/core/conceptGen/index.jsp).

Irradiated Murine Host Signature

Significance of analysis of microarray (SAM) used a two-class analysis with 100 permutations per comparison of the reference class to the target class, followed by a fold-change cut-off of 1.5 (26). To increase stringency, a secondary, or "tandem," bootstrapping was done by running the above SAM analysis iteratively, removing one sample from the reference class each time, including an iteration that removed no samples, to generate a list of genes regulated ≥1.5-fold present in 80% of the secondary SAM analyses.

Microarray data from 1608 human breast tumors from Ringner et al. and 337 untreated human breast cancer from Prat et al. (27–28) classified into molecular subtypes were used for cross-species comparison (28). Human orthologs of murine genes present on the human array platforms were used to cluster human microarray data using Gene Cluster as above. Genes in human microarray data were filtered by a criteria of Log 2-value >5 in 80% of the samples, before isolating the genes for clustering. In some analyses, the validity of clustering was tested against the performance of 10,000 randomly selected gene sets of the same size. Spearman's correlation and complete linkage was used to assess the distribution of irradiated and sham samples in the first divisions of the clustering dendrogram, excluding dendrograms with less than five samples in these branches.

Irradiated Host Centroid

To construct the irradiated host centroid, gene expression data (n=32) was filtered to include only genes with an expression above background in more than 75% of the samples. A t-test based p-value of the median centered probes present in the irradiated host signature was calculated for each gene based on separation of samples on irradiation status. Gene expression centroids were calculated for probes with a <0.01 p-value, leaving 133 of the original 323 probes. For classification of human tumors, mouse probes were translated to human genes using cross-referenced Entrez Gene IDs, leaving 72 genes in the centroid. Breast cancers were assayed for nearest centroid classification using Spearman correlation for the irradiated(IR-host)and the sham (S-host)gene expression centroids.

Results

Host irradiation induces a distinct metaprofile in *Trp53 null* murine mammary tumors

Expression profiles from 56 tumors arising from *Trp53* null fragments transplanted to cleared mammary glands of control or previously irradiated wildtype or *Tgfb1* heterozygote 3 month old mice were previously reported (22). Subsequently, Herschkowitz et al. showed that the transcriptional profiles of *Trp53* null tumors can be classified into molecular subtypes, including two basal-like classes, luminal, claudin-low similar to human breast cancer, and a subtype unique to this model (29). To determine whether the host irradiation affected the spectrum of intrinsic molecular subtypes, a intrinsic gene list, previously defined for mouse tumors (30), was used for hierarchical clustering analysis of expression profiles from these 56 tumors with 187 murine mammary tumors, including 50 other *Trp53* null tumors, and 10 mouse mammary glands. SigClust (31) was used to assess the significance of tumor clustering and objectively determine significant groups/subtypes (Figure 1, Table S1). SigClust assigned these newly analyzed 56 *Trp53* null tumors to basal-like(5/56), claudin-low (14/56), luminal (19/56) and p53 null (6/56) intrinsic subtypes. Two clusters of tumors, mostly from irradiated hosts (11/12), were unclassified by this method. The distribution of tumor subtypes as a function of either host irradiation and/or host genotype was not significantly different as determined by Chi-square (data not shown).

Yet host irradiation confers a distinct expression signature on tumor transcriptomes (22). Since tumors arising in irradiated hosts were not enriched in a particular tumor subtype, we concluded that the gene lists that define tumors arising in irradiated hosts are metaprofiles that overlay intrinsic subtype. To further explore this biology, we generated an irradiated host profile list of 323 genes (Table S2) significantly regulated by at least 1.5-fold in at least 80% of the secondary SAM bootstraps. As expected, tumors were clustered according to prior host irradiation (Figure 2A). Ingenuity Pathway Analysis (IPA) implicated inflammation as a key process imposed by the irradiated host environment. The top IPA interaction network included inflammatory response, cell-to-cell signaling and interaction, and organismal survival (Figure 2B; IPA Score 54). Specific inflammatory programs included proliferation of T-lymphocytes (p=8.5E-5, 21 genes), chemotaxis (p=3.1E-4, 20 genes), and cell movement of phagocytes (p=0.002, 16 genes). Four gene networks representing two cell types, stem cells and macrophages, and two processes, motility and autophagy were evident(Figure 2C).

The irradiated host signature segregates radiation preceded human cancer

To test whether this host biology is applicable to other experimental models, we applied this gene list to published data from sporadic or radiation-preceded rat sarcomas (32) and murine *Ptch1* mutant medulloblastoma (33). Most radiation-induced sarcoma and medulloblastoma were clustered by this profile (Figure S1A&B).

Encouraged by this evidence that the biology captured by the radiation chimera is useful across tumor types and species, we searched for expression profile microarray of human sporadic cancers compared to those preceded by radiation. Very few radiation-induced tumor microarray data sets are amenable to analysis due to sample size or platform differences (15, 32, 34–36). We applied the irradiated host signature to radiation-preceded papillary thyroid carcinomas (36) and radiotherapy associated sarcomas (32) (Figure 2D,E). Clustering using this subset of human gene orthologs present from the murine signature resulted in segregation of sporadic from radiation-preceded cancers. Permutation analysis showed that segregation by the genes from the irradiated host was significantly better than randomly selected genes in radiation-preceded thyroid cancers (p<0.02). This analysis suggests that host response to radiation, as defined by the radiation chimera model, also significantly affects the transcriptome of cancers arising in irradiated humans.

Association of irradiated host metaprofile and human breast cancer intrinsic subtypes

We devised a metric by which to classify breast cancers as similar to tumors from irradiated hosts by making a centroid classifier from the 323 irradiated host signature, defined herein. One centroid represents tumors from non-irradiated wildtype hosts, and second one represents tumors from irradiated WT hosts, both based on 72 of the most differentially expressed genes of the 323 gene list. As expected, the 72-gene centroid (Table S3) discriminates *Trp53* null tumors from sham-irradiated hosts and irradiated hosts (Figure 3A). ER status, based on immunostaining, was distributed independently (Figure 3B), even though more ER-negative breast tumors arose in irradiated hosts (22), which is consistent our previous report of a distinct signature that discriminated ER-negative tumors arising in an irradiated host (22). We used the centroid to assign human breast cancers compiled by Ringnér et al. (28) according to similarity to sham or irradiated host centroid. Most ER-positive breast cancers associated with the sham host signature while ER-negative breast cancers were strongly associated with irradiated host signature (Figure 3C). This suggests that the transcriptome in sporadic ER-negative, basal-like human breast cancer is influenced by tissue processes similar to those that promote ER-negative tumorigenesis in the radiation-chimera murine model. Several intrinsic subtypes are represented in ER-negative breast cancer, as evident in the data set from Prat et al. (10). Principle component analysis of the 72 genes in the centroid within the Prat UNC-337 data set demonstrated that claudin-low tumors were most strongly associated with the irradiated host centroid (Figure 3D).

Metaprofiles are gene expression modules consisting of co-expressed genes that represent key biological processes that have prognostic or treatment predictive power for cancer (37–38). Equipped with a gene list that had many orthologs present on human microarray platforms, we next determined how the biology represented in the irradiated host metaprofile applied to two data sets of sporadic human breast cancers. The first consists of 1608 breast cancers (28). Human orthologs of 182 of the 323 irradiated host gene list clustered tumors into two major groups, each of which had two subgroups (Figure 4A). Each of the four major subgroups was enriched with a particular intrinsic subtype, which were associated with different up-regulated genes (Table S4).

Subgroup 1A consisted of predominantly luminal A and B breast cancers that have an intermediate prognosis and disease free survival compared to the other groups (Figure 4 B,C). Group 1B contained a block of basal-like breast cancers and ERBB2 tumors and fared poorly, as is consistent with prior reports for these intrinsic subtypes (10, 39). Group 2C contained many normal-like tumors and luminal subtypes, and had a similar overall survival and relapse free survival as Group 1A. Notably Group 2D consisted of a mix of tumors that exhibited longer overall and disease free survival.

It is thought that the poor prognosis of basal-like and luminal-B subtypes is due to increased proliferation, as indicated by expression of many proliferation-related genes. To test the extent to which proliferation was driving the ability of irradiated host gene list to cluster the compiled breast cancer profiles, we removed 64 genes that were identified as being involved in “proliferation of eukaryotic cells” as annotated in IPA (Table S5). The remaining 118 genes of the irradiated host gene list still segregated breast cancers into four main subgroups enriched for molecular subtypes (Figure S2A). Without proliferation genes, the luminal-B cancer no longer shared the main bifurcation with the basal-like subcluster. Evenso, 1B was both strongly enriched in basal-like breast cancer and had a much worse relapse-free survival compared to the other subgroups, suggesting that the biology elicited by irradiated host is an important factor in prognosis of these tumors.

We next tested the utility of the irradiated host gene list in the data set from Prat et al. (10), which classified 337 human breast cancers (UNC337) into 6 intrinsic subtypes with the addition of the 6th type characterized as claudin-low. Using 203 human orthologs of the irradiated host genes clustered basal-like tumors, normal-like tumors and two distinct groups of claudin-low tumors (Figure 5A). Notably, claudin-low and basal-like tumors were on different arms.

Three gene clusters appear to define the clustering of the subtypes (Table S6). Cluster *a* contains 16 genes involved in tumorigenesis (p=8.9E-4); cluster *b* contains 19 genes involved in immune response (p=5.3E-8); and cluster *c* contains 29 genes involved in genetic disorders (p=2.3E-3). Expression of each of these gene clusters was significantly different among the six subtypes (Figure 5B–D). Principle component analysis of these data indicates that basal-like breast cancers are enriched for genes in cluster *a*, claudin-low are enriched in cluster *b*, and both tumors are depleted of genes in cluster *c*.

Claudin-low tumors exhibit epithelial-to-mesenchymal transition (EMT) features and are enriched in genes associated with stem cell biology (10). TGFβ is a key driver of EMT and mediates various aspects of stem cell biology, and is highly induced by ionizing radiation (reviewed in (40)). The radiation chimera experiment conducted in *Tgfb1* heterozygote hosts showed that the effect of host irradiation on tumor latency and growth rates was TGFβ dependent, while the effect on ER-status was not (22). We speculated that the profiles of tumors arising in irradiated *Tgfb1* heterozygote hosts compared to those arising in control mice would be informative. Significance of analysis of microarray (SAM) was done using a two-class analysis with 100 permutations per comparison of the reference class to the target class, followed by a fold-change cut-off of 1.5, followed by a secondary "tandem," bootstrapping(26). Interestingly, tumors arising in unirradiated hosts of either genotype were indistinguishable using this method. A list of 199 genes that were present in 100% of the secondary SAM analyses were able to segregate tumors of irradiated *Tgfb1* heterozygote hosts from those of non-irradiated heterozygote hosts under unsupervised hierarchical clustering, independent of ER status or histopathology subtype (Figure 6A). It did not do so when applied to tumors from wild type hosts (Figure 6B). Thus, host-irradiation, in conjunction with host TGFβ levels, elicits distinct transcriptional biologies of *Trp53* null tumors.

IPA revealed enrichment for cancer related genes, along with inflammatory processes such as recruitment and activation of lymphocytes and phagocytes (all p<0.005). Gene enrichment analysis using the ConceptGen database identified extracellular matrix programs and activation of monocytes, macrophages, and dendritic cells (all p<0.005) (data not shown). Eight genes, seven of which are "stem-related", are present in gene lists from both irradiated wildtype and *Tgfb1* heterozygote mice. However, five of seven genes are oppositely regulated between the two profiles. *Tp63, Igfbp2* and *Id4* are up-regulated in the irradiated host signature from wildtype irradiated mice, but down-regulated in that from *Tgfb1* heterozygote mice. This inverse pattern indicates that TGFβ is a critical component of the radiation response. For example, *Cd133*, a marker of progenitor cells and cancer initiating cells in several cancer types including breast cancer (41–43), is present only in profile from *Tgfb1* heterozygote mice.

We then applied this gene list to the UNC337 breast cancers. This list clustered breast cancer into two arms that represented roughly basal-like, ER-negative cancers and luminal, ER-positive cancers (Figure 6C). In contrast to clustering using the wildtype irradiated host signature, claudin-low and basal-like were no longer in distinct arms, indicating that the biology resulting from TGFβ provides an important distinction between ER-negative basal-like and claudin-low tumors (Table S6). Together these analyses suggest that processes promoting cancer in the irradiated mouse are strongly associated with spontaneous basal-like and claudin-low human breast cancer, the latter of which is particularly influenced by TGFβ associated inflammatory processes.

Conclusions

Here we show that a gene signature derived from a murine mammary radiation chimera model is informative in both radiation-preceded and sporadic human cancer, underscoring the contribution of host biology during cancer evolution. The *Trp53* null tumor subtype distribution was not particularly affected by host irradiation; rather, a distinct tumor microenvironmental transcriptome signature could be discerned, suggesting that tumors were "imprinted" by prior host radiation exposure. Together our analyses support the idea that the radiation response of the microenvironment is a significant component of the carcinogenic process. Moreover, subtype segregation using this signature suggests that similar processes may underlie the development of specific subtypes of apparently sporadic breast cancers.

The 323 irradiated host signature identified herein was enriched for genes indicative of inflammation, including a macrophage module, suggesting that either the recruitment or activation of inflammatory cells may underlie the effect of radiation on cancer. The human gene orthologs of a centroid classified breast cancers into distributions that suggest that a subgroup of ER-negative, basal-like intrinsic subtypes were like *Trp53* null tumors that arose in the irradiated hosts. The relevance of the *Trp53* null mammary model is supported by recent report of the Cancer Genome Atlas network on the molecular portraits of human breast cancer (44). The study group found that 80% of basal-like breast cancers harbored mutations in TP53, most of which were nonsense and frame shift mutations. The ER-negative, TP53 mutant, basal-like subtype was distinct from the other subtypes across all mRNA, miRNA, sequencing, and DNA copy number array platforms, suggesting that perhaps similar mechanisms may be detected in radiation-preceded cancer. The human orthologs of the irradiated murine host signature clustered two breast cancers datasets into groups with distinct outcomes and discriminated between closely related basal-like and claudin-low breast cancers. Of particular interest is that TGFβ mediated inflammatory processes strongly define the claudin-low breast cancers (45).

The signature derived from the radiation chimera model also provided important insights into features of sporadic human breast cancer. Several recent studies have turned attention to the stroma to derive prognostic value by using expression profiling of stromal and extratumoral tissues (46). Using microdissected stroma from breast cancer, Finak et al. showed that a stroma-derived prognostic predictor stratifies disease outcome based on a signature of immune mediators, hypoxia and angiogenesis (47). Analysis of the expression profiles from invasive breast cancer and ductal carcinoma *in situ* provides evidence that stromal biology is a key determinant of progression (48). Consistent with this, the presence of distinct subtypes of microenvironment, an active versus inactive cancer-adjacent microenvironment, influences the aggressiveness and outcome of ER positive human breast cancers (46). Our data suggests that host biology induced in the radiation chimera model has strong parallels to the biology that underlies aggressive, ER-negative sporadic breast cancers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by funding from NASA Specialized Center for Research in Radiation Health Effects, NNX09AM52G, and the Department of Energy, Office of Biological and Environmental Research program on Low Dose Radiation. CMP and ZW received support from the NCI Breast SPORE program P50-CA58223, by RO1-CA138255 and RO1-CA148761, and by the Breast Cancer Research Foundation.

Abbreviations

ER	estrogen receptor
SAM	significance of analysis of microarray
TGFβ	transforming growth factor β1

References

1. Ronckers CM, Erdmann CA, Land CE. Radiation and breast cancer: a review of current evidence. Breast Cancer Res. 2005; 7:21–32. [PubMed: 15642178]
2. Shore RE, Hildreth N, Dvoretsky P, Andresen E, Moseson M, Pasternack B. Thyroid cancer among persons given X-ray treatment in infancy for an enlarged thymus gland. Am J Epidemiol. 1993; 137:1068–80. [PubMed: 8317436]
3. Shore RE, Woodard E, Hildreth N, Dvoretsky P, Hempelmann L, Pasternack B. Thyroid tumors following thymus irradiation. J Natl Cancer Inst. 1985; 74:1177–84. [PubMed: 3858590]
4. Tokunaga M, Land CE, Tokuoka S. Follow-up studies of breast cancer incidence among atomic bomb survivors. J Radiat Res (Tokyo). 1991; 32 (Suppl):201–11. [PubMed: 1762108]

5. Boice JD Jr, Preston D, Davis FG, Monson RR. Frequent chest x-ray fluoroscopy and breast cancer incidence among tuberculosis patients in Massachusetts. Radiat Res. 1991; 125:214–22. [PubMed: 1996380]
6. Shore RE, Hildreth N, Woodard E, Dvoretsky P, Hempelmann L, Pasternack B. Breast cancer among women given X-ray therapy for acute postpartum mastitis. J Natl Cancer Inst. 1986; 77:689–96. [PubMed: 3462410]
7. De Bruin ML, Sparidans J, van't Veer MB, Noordijk EM, Louwman MWJ, Zijlstra JM, et al. Breast cancer risk in female survivors of Hodgkin's lymphoma: lower risk after smaller radiation volumes. J Clin Oncol. 2009; 27:4239–46. [PubMed: 19667275]
8. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000; 406:747–52. [PubMed: 10963602]
9. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001; 98:10869–74. [PubMed: 11553815]
10. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. 2010; 12:R68. [PubMed: 20813035]
11. Haibe-Kains B, Desmedt C, Piette F, Buyse M, Cardoso F, Van't Veer L, et al. Comparison of prognostic gene expression signatures for breast cancer. BMC Genomics. 2008; 9:394. [PubMed: 18717985]
12. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, et al. Concordance among gene-expression-based predictors for breast cancer. N Engl J Med. 2006; 355:560–9. [PubMed: 16899776]
13. Castiglioni F, Terenziani M, Carcangiu ML, Miliano R, Aiello P, Bertola L, et al. Radiation effects on development of HER2-positive breast carcinomas. Clin Cancer Res. 2007; 13:46–51. [PubMed: 17200337]
14. Wang SE, Narasanna A, Whitell CW, Wu FY, Friedman DB, Arteaga CL. Convergence of p53 and Transforming Growth Factor beta (TGFbeta) Signaling on Activating Expression of the Tumor Suppressor Gene maspin in Mammary Epithelial Cells 10.1074/jbc.M608499200. J Biol Chem. 2007; 282:5661–9. [PubMed: 17204482]
15. Broeks A, Braaf LM, Wessels LF, van de Vijver M, De Bruin ML, Stovall M, et al. Radiation-associated breast tumors display a distinct gene expression profile. Int J Radiat Oncol Biol Phys. 2010; 76:540–7. [PubMed: 20117289]
16. UNSCEAR. Sources and effects of ionizing radiation. New York: United Nations; 2006.
17. Tubiana, M.; Aurengo, A.; Averbeck, D.; Bonnin, A.; Le Guen, B.; Masse, R., et al. Dose-effect relationships and estimation of the carcinogenic effect of low doses of ionizing radiation. Paris: Académie des Sciences -Académie Nationale de Médecine; 2005 Mar.
18. Barcellos-Hoff MH, Park C, Wright EG. Radiation and the microenvironment - tumorigenesis and therapy. Nat Rev Cancer. 2005; 5:867–75. [PubMed: 16327765]
19. Little MP, Filipe JAN, Prise KM, Folkard M, Belyakov OV. A model for radiation-induced bystander effects, with allowance for spatial position and the effects of cell turnover. Journal of Theoretical Biology. 2005; 232:329–38. [PubMed: 15572058]
20. Durante M, Cucinotta FA. Heavy ion carcinogenesis and human space exploration. Nat Rev Cancer. 2008; 8:465–72. [PubMed: 18451812]
21. Barcellos-Hoff MH, Ravani SA. Irradiated mammary gland stroma promotes the expression of tumorigenic potential by unirradiated epithelial cells. Cancer Res. 2000; 60:1254–60. [PubMed: 10728684]
22. Nguyen DH, Oketch-Rabah HA, Illa-Bochaca I, Geyer FC, Reis-Filho JS, Mao JH, et al. Radiation Acts on the Microenvironment to Affect Breast Carcinogenesis by Distinct Mechanisms that Decrease Cancer Latency and Affect Tumor Type. Cancer Cell. 2011; 19:640–51. [PubMed: 21575864]
23. Jerry DJ, Kittrell FS, Kuperwasser C, Laucirica R, Dickinson ES, Bonilla PJ, et al. A mammary-specific model demonstrates the role of the p53 tumor suppressor gene in tumor development. Oncogene. 2000; 19:1052–8. [PubMed: 10713689]

24. Medina D, Kittrell FS, Shepard A, Stephens LC, Jiang C, Lu J, et al. Biological and genetic properties of the p53 null preneoplastic mammary epithelium. Faseb J. 2002; 16:881–3. [PubMed: 11967232]
25. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on bias and variance. Bioinformatics. 2003; 19:185–93. [PubMed: 12538238]
26. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001; 98:5116–21. [PubMed: 11309499]
27. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. Journal of Clinical Oncology. 2009; 27:1160–7. [PubMed: 19204204]
28. Ringnér M, Fredlund E, Häkkinen J, Borg Å, Staaf J. GOBO: Gene Expression-Based Outcome for Breast Cancer Online. PLoS ONE. 2011; 6:e17911. [PubMed: 21445301]
29. Herschkowitz JI, Zhao W, Zhang M, Usary J, Murrow G, Edwards D, et al. Comparative oncogenomics identifies breast tumors enriched in functional tumor-initiating cells. Proceedings of the National Academy of Sciences. 2011; 109:2778–83.
30. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. 2007; 8:R76. [PubMed: 17493263]
31. Liu Y, Hayes DN, Nobel A, Marron JS. Statistical significance of clustering for high-dimension, low-sample size data. J Am Stat Assoc. 2008; 103:1281–93.
32. Hadj-Hamou NS, Ugolin N, Ory C, Britzen-Laurent N, Sastre-Garau X, Chevillard S, et al. A transcriptome signature distinguished sporadic from post-radiotherapy radiation-induced sarcomas. Carcinogenesis. 2011
33. Ishida Y, Takabatake T, Kakinuma S, Doi K, Yamauchi K, Kaminishi M, et al. Genomic and gene expression signatures of radiation in medulloblastomas after low-dose irradiation in Ptch1 heterozygous mice. Carcinogenesis. 2010; 31:1694–701. [PubMed: 20616149]
34. Detours V, Wattel S, Venet D, Hutsebaut N, Bogdanova T, Tronko MD, et al. Absence of a specific radiation signature in post-Chernobyl thyroid cancers. Br J Cancer. 2005; 92:1545–52. [PubMed: 15812549]
35. Detours V, Delys L, Libert F, Weiss Solis D, Bogdanova T, Dumont JE, et al. Genome-wide gene expression profiling suggests distinct radiation susceptibilities in sporadic and post-Chernobyl papillary thyroid cancers. Br J Cancer. 2007; 97:818–25. [PubMed: 17712314]
36. Delys L, Detours V, Franc B, Thomas G, Bogdanova T, Tronko M, et al. Gene expression and the biological phenotype of papillary thyroid carcinomas. Oncogene. 2007; 26:7894–903. [PubMed: 17621275]
37. Huang E, Cheng SH, Dressman H, Pittman J, Tsou MH, Horng CF, et al. Gene expression predictors of breast cancer outcomes. The Lancet. 2003; 361:1590–6.
38. Brunet J-P, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. Proceedings of the National Academy of Sciences. 2004; 101:4164–9.
39. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A. 2003; 100:8418–23. [PubMed: 12829800]
40. Moses H, Barcellos-Hoff MH. TGF-b Biology in Mammary Development and Breast Cancer. Cold Spring Harbor Perspectives in Biology. 2011; 3:a003277. [PubMed: 20810549]
41. Wright MH, Calcagno AM, Salcido CD, Carlson MD, Ambudkar SV, Varticovski L. Brca1 breast tumors contain distinct CD44+/CD24− and CD133+ cells with cancer stem cell characteristics. Breast Cancer Res. 2008; 10:R10. [PubMed: 18241344]
42. Wang J, Wakeman TP, Lathia JD, Hjelmeland AB, Wang X-F, White RR, et al. Notch Promotes Radioresistance of Glioma Stem Cells. STEM CELLS. 2010; 28:17–28. [PubMed: 19921751]
43. Hwang-Verslues WW, Kuo W-H, Chang P-H, Pan C-C, Wang H-H, Tsai S-T, et al. Multiple Lineages of Human Breast Cancer Stem/Progenitor Cells Identified by Profiling with Stem Cell Markers. PLoS ONE. 2009; 4:e8377. [PubMed: 20027313]

44. TCGA. Comprehensive molecular portraits of human breast tumours. Nature. 2012 advance online publication.

45. Bruna A, Greenwood W, Le Quesne J, Teschendorff A, Miranda-Saavedra D, Rueda OM, et al. TGFβ induces the formation of tumour-initiating cells in claudin low breast cancer. Nat Commun. 2012; 3:1055. [PubMed: 22968701]

46. Roman-Perez E, Casbas-Hernandez P, Pirone J, Rein J, Carey L, Lubet R, et al. Gene expression in extratumoral microenvironment predicts clinical outcome in breast cancer patients. Breast Cancer Research. 2012; 14:R51. [PubMed: 22429463]

47. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, et al. Stromal gene expression predicts clinical outcome in breast cancer. Nat Med. 2008; 14:518–27. [PubMed: 18438415]

48. Lee S, Stewart S, Nagtegaal I, Luo J, Wu Y, Colditz G, et al. Differentially Expressed Genes Regulating the Progression of Ductal Carcinoma in Situ to Invasive Breast Cancer. Cancer Research. 2012

Translational Relevance

Radiation has long been established as a risk factor for breast cancer with recent evidence suggesting that it promotes estrogen receptor negative cancers in mice and humans. Human breast cancer comprises at least six transcriptional subtypes, each with distinct molecular programs. We report that the human orthologs of a mouse mammary tumor transcriptional program elicited by prior host irradiation can be used to segregate radiation preceded cancers. Moreover, this program is associated with ER-negative human breast cancer subtypes, suggesting that sporadic cancer is promoted by similar biological processes.

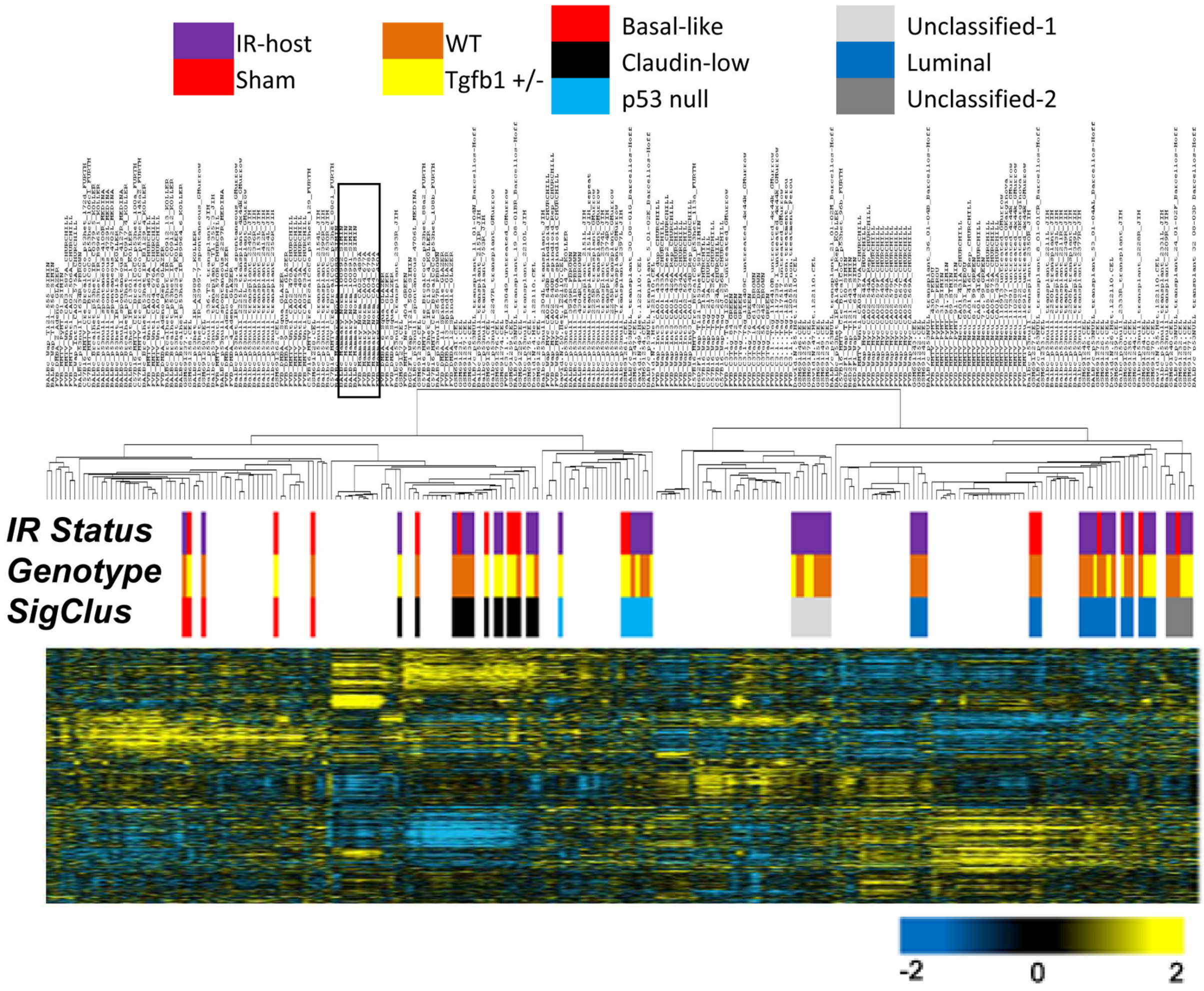

Figure 1. *Trp53* null murine mammary tumors classified into intrinsic molecular subtypes

Gene expression profiling of 56 *Trp53* null murine mammary tumors arising in the radiation chimera model consisting of either wild type or *Tgfb1* heterozygote hosts were classified by the SigClust method into intrinsic mouse subtypes (basal-like, red; claudin-low, black; *Trp53* null, light blue; luminal, dark blue; unclassified type 1, light gray; unclassified type 2, dark gray). They were also clustered along with 187 other tumors from various mouse mammary tumor models, including 10 normal mammary glands (box). (Sham-irradiated host, red; irradiated-host, purple; wild type, brown; *Tgfb1* +/−, yellow.)

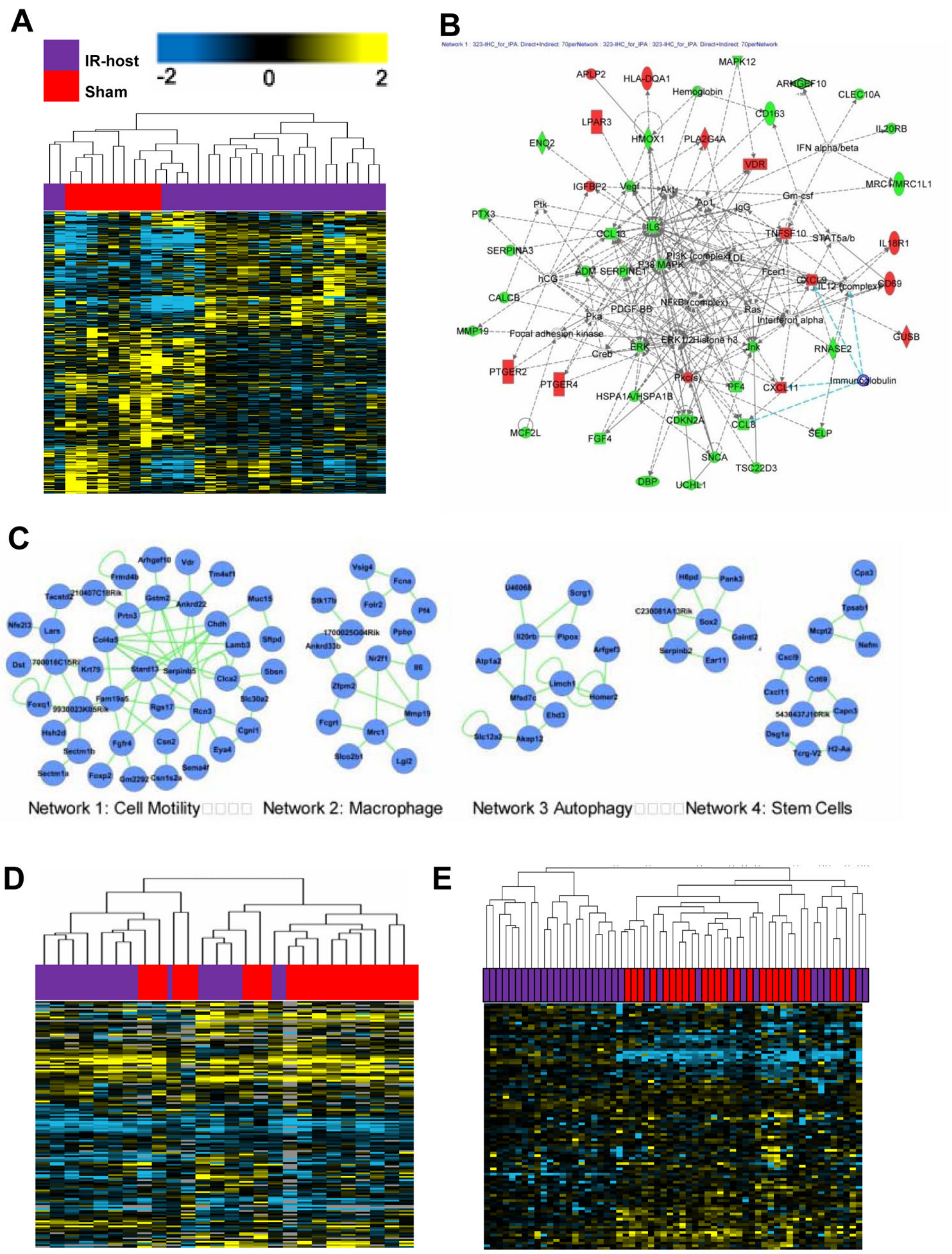

Figure 2. Tumors arising in irradiated hosts exhibit distinct gene expression program

(A) SAM identified 323 genes regulated by at least 1.5-fold in tumors arising in wild type irradiated hosts, which cluster host-irradiation status apart from sham-host-irradiation. (B) IPA interactome of this gene list implicated inflammation, proliferation and development as major biological activities. (C) The four identified gene networks represent macrophages, stem cells, autophagy, and cell motility. (D) Radiation-preceded human thyroid cancers were clustered by 139 of the murine genes present in that data set. Sporadic thyroid cancers were segregated from radiation-preceded cancers in children from Chernobyl. Chi-square test of association between irradiation status and the main dendrogram bifurcation, p=0.02. (E) Radiotherapy associated sarcomas were clustered by 92 human orthologs present from the murine data set. Red: Sporadic; Purple: radiation-preceded

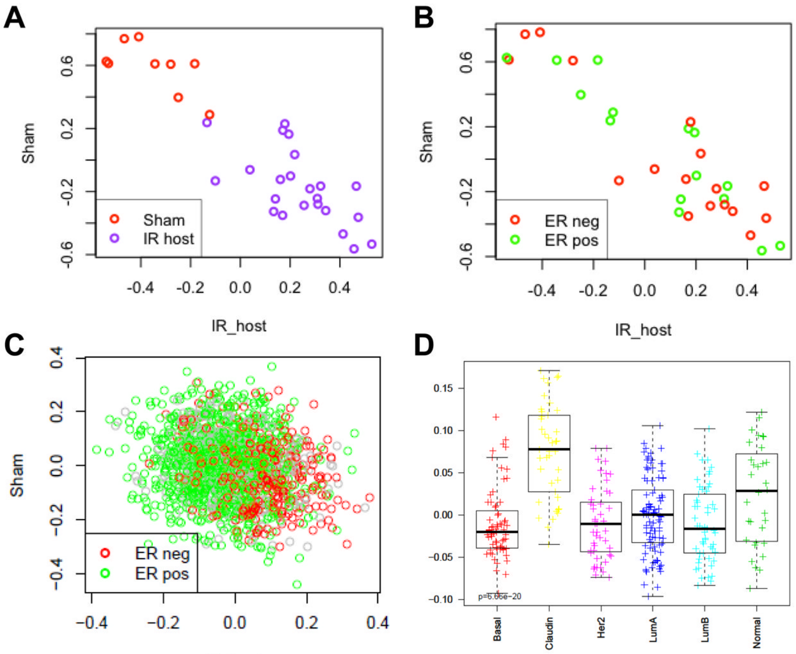

Figure 3. The irradiated host centroid associates with distinct human breast cancer intrinsic subtypes

Two centroids, one representing non-irradiated hosts (S-host) and the other irradiated hosts (IR-host), were derived based on 72 genes that were most differentially expressed within the 323 host irradiated profile. (A) The correlation of murine *Trp53* null tumors with the IR-host and S-host centroids demonstrates robust discrimination, as expected. (B) The distribution of murine tumors as a function of ER status is homogenous (ER-positive, green; ER-negative, red). (C) Contour plot of ER-positive and negative status of 1608 human breast cancers from by Ringnér et al. (28), after calculating their correlations to the two centroids. The distributions of ER-positive breast cancer (green) and ER-negative breast cancer (red) are significantly different (KS test p-value = 6.5E-36). (D) Analysis of variance plots of the expression of the 72 genes across each breast cancer subtype in the UNC337 data set indicate distinct behaviors of the claudin-low breast cancers compared to other subtypes.

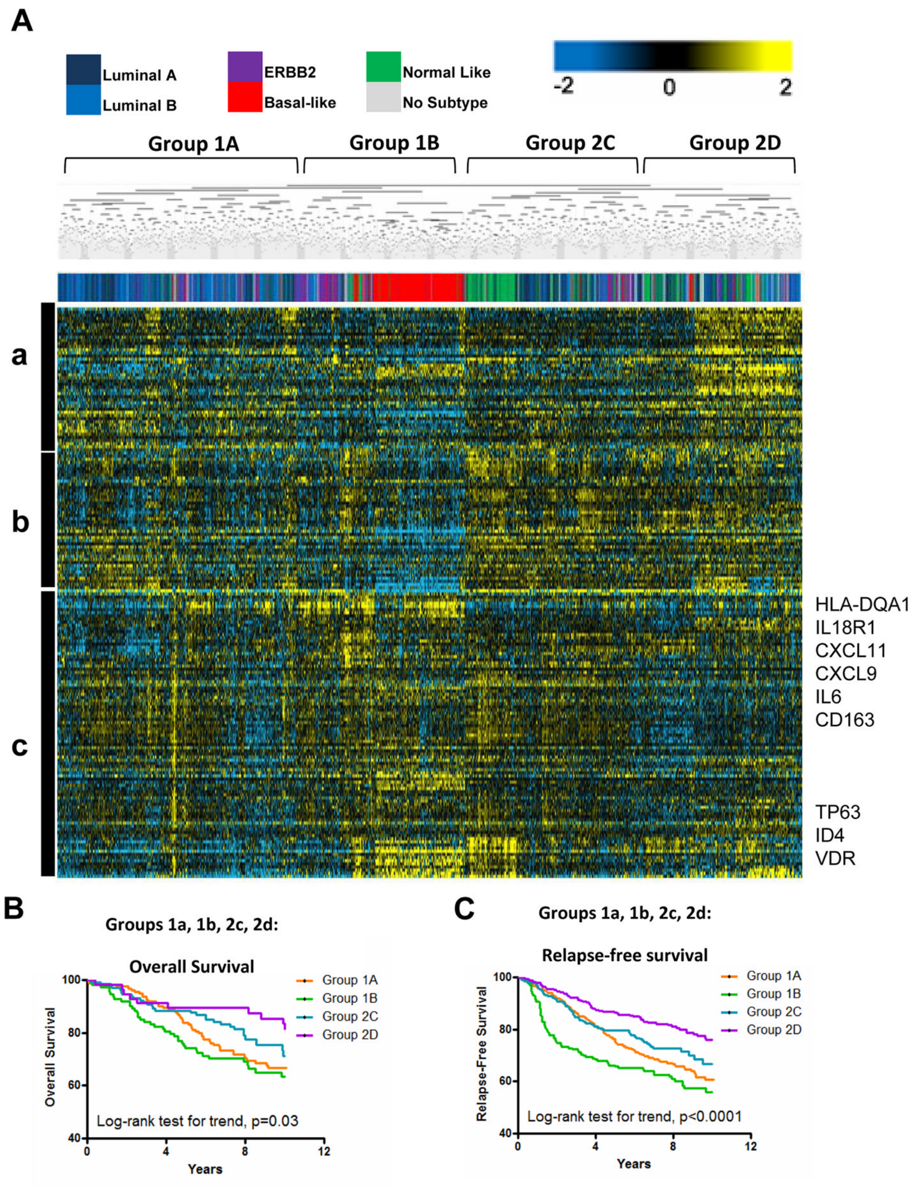

Figure 4. The irradiated host gene profile stratifies human breast cancers into prognostic groups

(A) Unsupervised clustering of 1,608 human breast cancers compiled from 10 independent studies were clustered using 182 orthologs of the murine genes. Tumors were segregated into four subgroups, some of which contained a predominant molecular subtype (luminal-A, dark blue; luminal-B, light blue; ERBB2/HER2, purple; basal-like, red; normal-like, green; and unclassified, gray). Black bars represent the three gene clusters (a, b, c) that represent genes induced within the four subgroups. (B, C) The four major subgroups exhibited significantly different overall survival (B) and relapse-free survival (C); (Group 1A, orange; Group 1B, green; Group 2C, blue; Group 2D, purple).

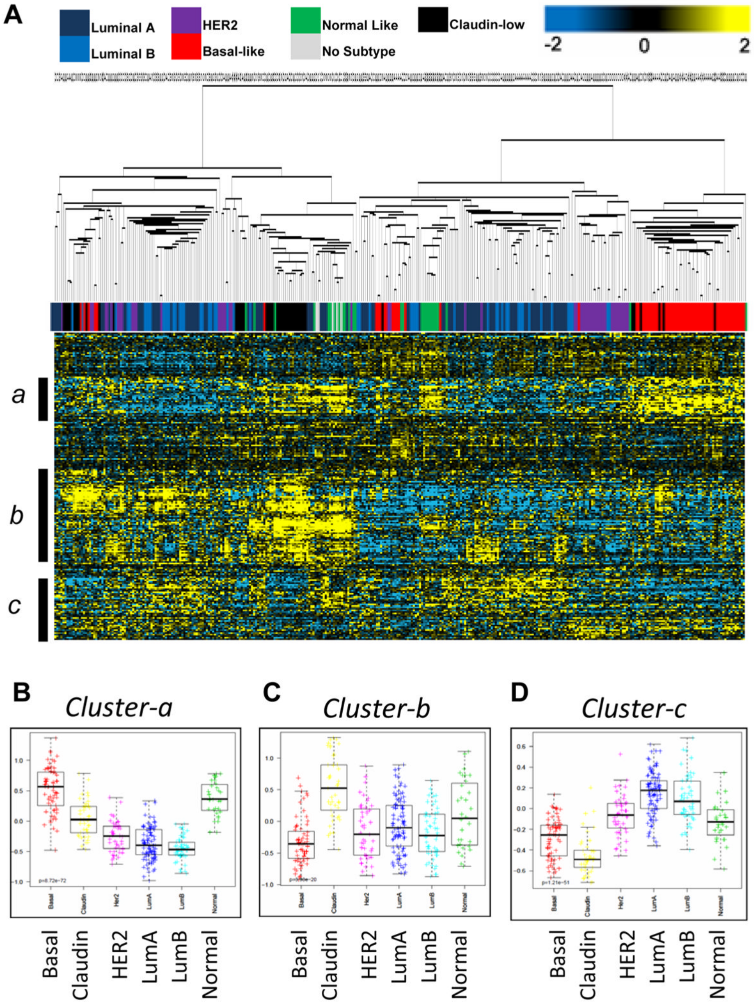

Figure 5. The irradiated host gene profile stratifies claudin-low breast cancers apart from other subtypes

(A) 337 human breast cancers from Prat et al. (2010) were clustered by 203 of the murine genes present and resulted in four subgroups that were enriched for particular subtypes of breast cancer (luminal-A, dark blue; luminal-B, light blue; HER2, purple; basal-like, red; claudin-low, black; normal-like, green; unclassified, gray). Gene clusters, a, b, and c, are indicated by black bars. (B-D) Analysis of variance of the median expression level for each gene module highlighted in (A) across each of the six intrinsic subtypes. (B) Basal-like tumors are strongly associated with gene cluster a (ANOVA p=8.7E-72), representing tumorigenesis-related genes. (C) Claudin-low tumors are strongly associated with gene cluster b (ANOVA p=3.9E-20) representing immune response genes. (D) Both basal-like and claudin-low tumors are negatively associated with gene cluster c (ANOVA p=1.2E-51), representing genes involved in genetic disorders.

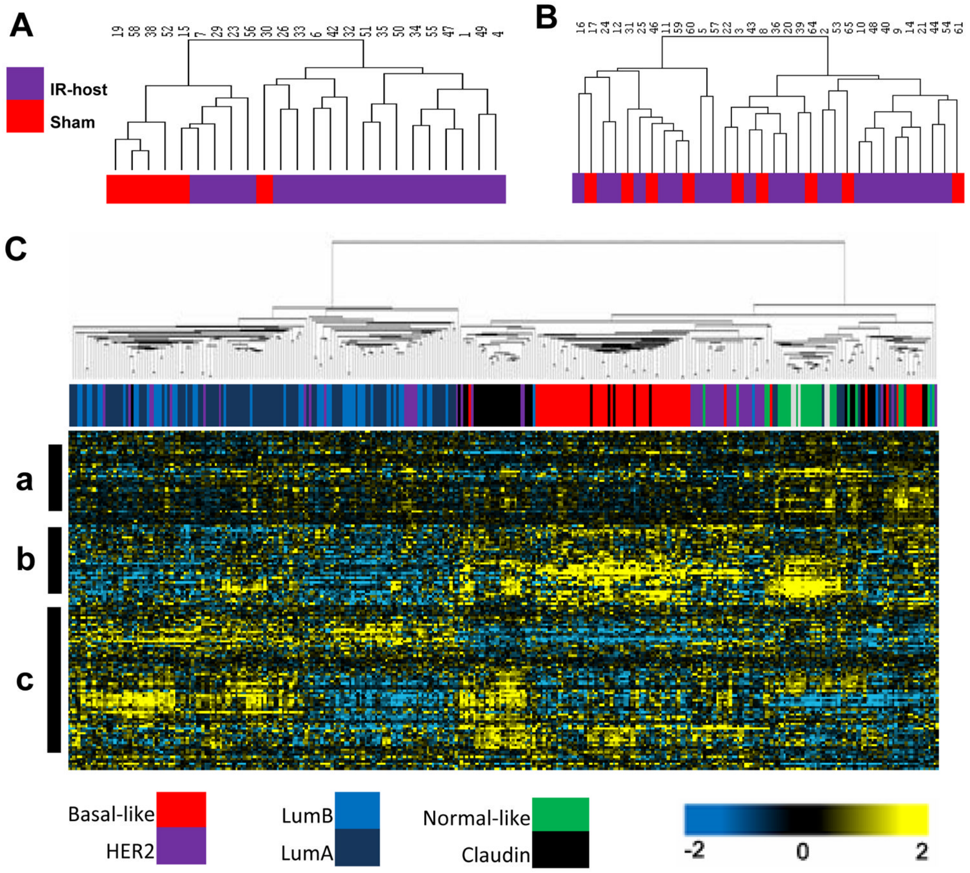

Figure 6. Host irradiation interacts with *Tgfb1* genotype to yield distinct gene profile that detects ER-status of human breast cancers

SAM analysis of 24 *Trp53* null tumors arising in *Tgfb1* heterozygote hosts resulted in 199 genes regulated by at least 1.5-fold in tumors arising after host-irradiation. (A) The 199 genes segregate tumors of irradiated hosts (purple) apart from those in non-irradiated hosts (red) in the heterozygote background, but did not do so for tumors from the wild type background (B). (C) 337 human breast cancers from Prat et al. (2010) were clustered by the human orthologs of the 199 genes present in that platform (luminal-A, dark blue; luminal-B, light blue; HER2, purple; basal-like, red; claudin-low, black; normal-like, green; unclassified, gray).