In Breast Cancer, Mutation Signature 30 is Most Prevalent in the Luminal A Subtype and is Associated with Downregulation of snoRNAs

Nicholas A. Grosenbacher[1], Samuel W. Brady[1], Frederick R. Adler[2], Andrea H. Bild[1]
[1]Department of Pharmacology and Toxicology, [2]Departments of Mathematics and Biology, University of Utah

Abstract

Across all cancer types, signatures of mutagenesis have been found and causes for a subset of these signatures have been determined. We demonstrate association of one signature of unknown etiology, signature 30, with luminal A breast cancers. Using pathway analysis, we show that rRNA metabolic pathways are dysregulated in signature 30, possibly via snoRNAs, which are almost entirely downregulated compared to patients that do not have signature 30. In addition, markers for E2F signaling are differentially expressed between patients with high or low signature 30 status.

Background

Well-documented and specific patterns of somatic mutations in cancer have been ascribed to mutagenic processes such as tobacco use and UV exposure. More recently, Alexandrov et al.[1] developed a mathematical model to generate mutation signatures from genomic data. At least 30 SNV signatures have been verified, including 12 present in breast cancers[2]. While some mutation signatures have known etiologies, causes for others, such as signature 30 (Figure 1), are still unknown. By using an existing model to attribute SNVs to signatures, nonnegative weights can be calculated, which estimate the impact of each signature across all of a patient's somatic SNVs. Signature 30, named so because it was the thirtieth signature discovered, has a high proportion of N[**C>T**]A and N[**C>T**]C mutations.

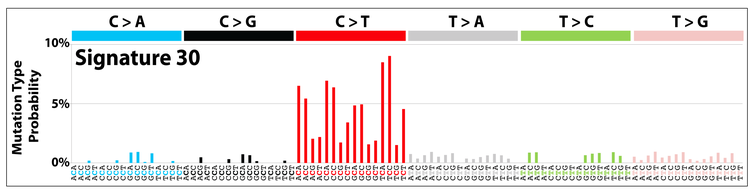

Figure 1: Signature 30 can be visualized as proportions of 96 possible SNVs within a trinucleotide context. Signature 30 is characterized by a high proportion of C>T mutations, especially N[**C>T**]A/C.[3]

Methods

All patient data (SNP information, PAM50 clustering status, RNA-Seq expression data) was obtained from The Cancer Genome Atlas. Patient mutation signature weights were calculated using the deconstructSigs package in R based on reference signatures provided within the package. Signatures 1, 2, 3, 5, 6, 8, 13, 17, 18, 20, 26, and 30 were paneled as these signatures are known to occur in breast cancers. Patients were considered to possess a given signature if that signature had a weight greater than zero. Expression data was log2 transformed and separated by patient mutation signature possession. Welch's t-test was used to compare groups of expression data.

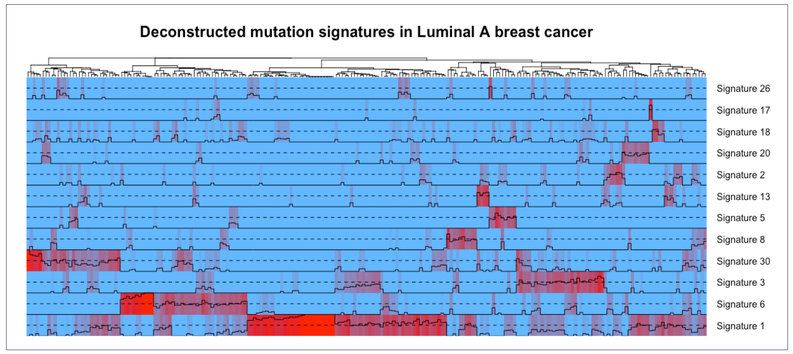

Figure 2: Visual clustering demonstrated low signature overlap in luminal A patients, especially involving signature 30. Red intensity indicates a higher weight for that signature within that patient.

Results

Clustering analysis was performed to identify signatures which might frequently co-occur. Little overlap occurred between the most common signatures (Figure 2). Within luminal A breast cancers, the most frequently occurring signatures were signatures 1 (ageing), 6 (failure of DNA mismatch repair), and 30 (unknown etiology). Signature 30 accounted for 9.31% of signature weights across all breast cancers, including 12.1% of signature weights within the luminal A subtype (Figure 3). As expected, signature 3 (double-strand break repair failure) accounted for a plurality of weights in basal-like breast cancers, reflecting a higher incidence of BRCA1/2 mutations in this subtype.

Pathway analysis of the top 300 differentially expressed genes between signature 30 low and high luminal A breast cancers revealed a phenotype of metabolic dysregulation, particularly of ribosomal RNA. E2F signaling was also higher in samples that lack signature 30 (Table 1, direction not shown). Furthermore, small nucleolar RNAs (snoRNAs) were almost entirely downregulated in samples that possessed signature 30 (Figure 4).

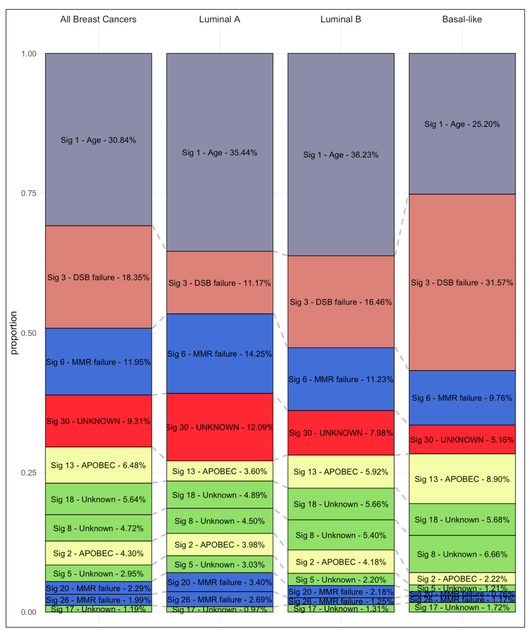

Figure 3: Proportions of paneled signatures found in breast cancer subtypes showed that signature 30 (red) is most prevalent in luminal A breast cancer. As expected, signature 3 (brown) is most prevalent in basal-like as signature 1 (gray) decreases. These are concordant with notions that basal-like breast cancers typically have a higher rate of BRCA1/2 mutations and a younger age of diagnosis.

Table 1: Top Gene Ontology and TRANSFAC terms within top 300 differentially expressed genes in signature 30 positive vs negative luminal A breast cancers. Data generated using GATHER.

Source	Category	p
GO	cytoplasm organization	< 0.0001
GO	ribosome biogenesis	< 0.0001
GO	ribosome biogenesis and assembly	< 0.0001
GO	rRNA processing	0.0001
GO	rRNA metabolism	0.0002
GO	primary metabolism	0.0002
TRANSFAC	E2F_Q6_01	< 0.0001
TRANSFAC	KROX_Q6	0.0001
TRANSFAC	E2F_Q4	0.0002

Discussion

Here, we examined signature 30, a mutation signature of unknown etiology, in luminal A breast cancers by comparing those pathways differentially expressed between luminal A patients positive and negative for signature 30. We found downregulation of almost all snoRNAs in patients with this signature. Interestingly, snoRNAs regulate rRNA biogenesis and mRNA splicing, and have been proposed to impact tumorigenesis and metastasis[4]. Additionally, signaling of transcription factor and cell cycle marker E2F is amplified in tumors lacking signature 30. These findings may be key in determining etiology and clinical impact of signature 30. Notably, previous use of signature-generation based methods indicates that fewer than 5% of breast tumors contain signature 30[5]. However, the attribution-based method we used showed that over 9% of breast cancers were impacted by signature 30. Further investigation of methodology and sample bias is needed to verify findings.

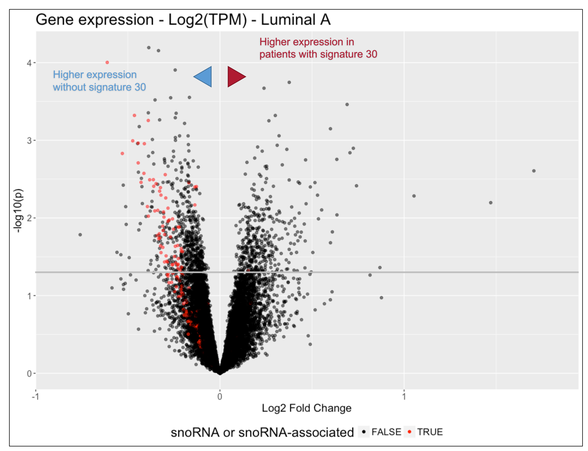

Figure 4: Volcano plot demonstrates that snoRNA expression is lower in luminal A breast cancer patients that have signature 30 compared to those without. Points in red represent snoRNAs and snoRNA-related genes. snoRNAs almost universally have higher mean expression in samples that are negative for signature 30. The gray horizontal line indicates p=0.05.

References

1. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500,** 415–421 (2013).
2. Nik-Zainal, S. *et al.* Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534,** 47–54 (2016).
3. Forbes, S. A. *et al.* COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Research* **45,** D777–D783 (2017).
4. Crea, F., Clermont, P. L., Parolia, A., Wang, Y. & Helgason, C. D. The non-coding transcriptome as a dynamic regulator of cancer metastasis. *Cancer and Metastasis Reviews* **33,** 1–16 (2014).
5. Nik-Zainal, S. & Morganella, S. Mutational Signatures in Breast Cancer: The Problem at the DNA Level. *Clinical Cancer Research* **23,** 2617–2629 (2017).