

Tandem duplications contribute to not one but two distinct phenotypes

Johnathan Watkins^{a,b,c}, Andrew Tutt^{b,c,d}, and Anita Grigoriadis^{a,b,c,1}

Menghi et al. (1) report a metric to classify tumors into those with and without a tandem duplicator phenotype (TDP), using the frequency of tandem duplications (TDs) in 277 whole-genome sequenced samples. Building on a previous method (2), Menghi et al. (1) identified TDs from SNP array data, and found that the TDP was strongly associated with response to the DNA damaging chemotherapeutic, cisplatin. These findings supplement the growing recognition that genome-wide signatures of mutator phenotypes may prove to be important additions to the companion diagnostic repertoire (3, 4). Although the findings of this report are highly stimulating, accumulating evidence suggests that an elevated abundance of TDs features in not just one but two distinct phenotypes.

Two of the original studies on the TDP reported a mutual exclusion with breast cancer 1/2 (*BRCA1/2*) inactivation (2, 5), which conflicts with the enrichment of *BRCA1* loss among TDP cancers observed by Menghi et al. (1). Using The Cancer Genome Atlas (TCGA) breast cancer data (6), we established allele-specific copy number profiles using ASCAT (7) before calling TDP status, as described previously (2), using two different size ranges for the TD-like features: (i) between 1 Kbp and 2 Mbp in accordance with the study by Menghi et al. (1); and (ii) between 2 Mbp and 10 Mbp. Five samples with *BRCA1* inactivation exhibited the TDP when considering only shorter TDs (Fig. 1A); however, we found no instances of tumors with *BRCA1* inactivation among 2- to 10-Mbp TDP cancers (Fig. 1B). Furthermore, although 56% of the Menghi et al. (1) study's TDP calls were shared with our 1-Kbp to 2-Mbp TDP calls, only 10% of the Menghi et al.

study's TDP calls agreed with our 2- to 10-Mbp TDP calls (Fig. 1C). In addition, we found that although the 1-Kbp to 2-Mbp TDP calls and the Menghi et al. study's TDP calls were enriched for triple-negative breast cancers ($P < 0.001$, Fisher's exact test), the 2- to 10-Mbp TDP calls were not ($P = 0.81$, Fisher's exact test). These findings support the notion that the study by Menghi et al. captures one particular TDP distinguishable from a second TDP by length and contrasting relationships with loss of *BRCA1* function.

Our results are reinforced by two recent analyses. The first study extracted two TD-enriched rearrangement signatures from 560 whole breast cancer genomes (8). "Signature-1" mostly comprised TDs between 1 and 10 Mbp, whereas "signature-3" mostly comprised TDs ≤ 100 Kbp. Signature-3 was associated with *BRCA1* disruption, signatures of homologous recombination deficiency, and was observed in $\sim 15\%$ of the cohort. In contrast, signature-1 was independent of *BRCA1/2* disruptions, exhibited links with mutational signatures of both homologous recombination deficiency and mismatch repair deficiency, and presented in $\sim 8.5\%$ of the cohort. The second study identified an ovarian and prostate cancer-linked TDP featuring TDs up to 10 Mbp, mutual exclusion with *BRCA1/2* inactivation, and enrichment for inactivation of the *CDK12* kinase (9).

In conclusion, we propose that there are actually two TDPs, with the study by Menghi et al. (1) providing a comprehensive characterization of the *BRCA1* inactivation-linked TDP. The existence of two TDPs has important implications for the robust development of genomic instability-based biomarkers of drug response.

^aCancer Bioinformatics, King's College London, London SE1 9RT, United Kingdom; ^bBreast Cancer Now Research Unit, King's College London, London SE1 9RT, United Kingdom; ^cDivision of Cancer Studies, King's Health Partners Academic Health Science Centre, Faculty of Life Sciences and Medicine, King's College London, London SE1 9RT, United Kingdom; and ^dThe Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, United Kingdom

Author contributions: J.W., A.T., and A.G. designed research; J.W., A.T., and A.G. performed research; J.W. and A.G. analyzed data; and J.W., A.T., and A.G. wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. Email: anita.grigoriadis@kcl.ac.uk.

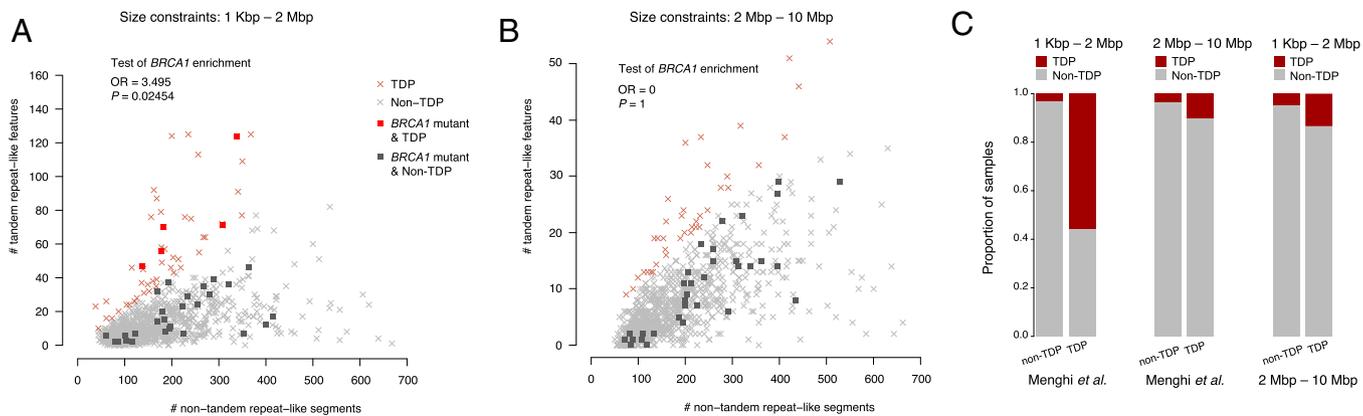


Fig. 1. Tandem duplication phenotypes in 940 TCGA breast cancers. (A and B) TDP status was determined using genomic segments between 1 Kbp and 2 Mbp (A), or 2 Mbp and 10 Mbp (B), followed by Gaussian mixture modeling of the ratio of TDs to non-TD segments (total number of segments minus double the number of TD segments as per ref. 2). Odds ratio and P value represent Fisher's exact test of BRCA1 mutation enrichment in the TDP subset of tumors. BRCA1 loss was defined as germ-line or somatic point mutation or deletion. TDP tumors are colored in red and non-TDP tumors in gray. All samples are denoted by an "x," with the exception of tumors with BRCA1 loss, which are denoted by a square. (C) Bar plots illustrate the overlaps between the different TDP calling methods.

- 1 Menghi F, et al. (2016) The tandem duplicator phenotype as a distinct genomic configuration in cancer. *Proc Natl Acad Sci USA* 113(17):E2373–E2382.
- 2 Ng CKY, et al. (2012) The role of tandem duplicator phenotype in tumour evolution in high-grade serous ovarian cancer. *J Pathol* 226(5):703–712.
- 3 Schouten PC, et al. (2015) Robust BRCA1-like classification of copy number profiles of samples repeated across different datasets and platforms. *Mol Oncol* 9(7):1274–1286.
- 4 Watkins JA, Irshad S, Grigoriadis A, Tutt AN (2014) Genomic scars as biomarkers of homologous recombination deficiency and drug response in breast and ovarian cancers. *Breast Cancer Res* 16(3):211.
- 5 McBride DJ, et al. (2012) Tandem duplication of chromosomal segments is common in ovarian and breast cancer genomes. *J Pathol* 227(4):446–455.
- 6 Cancer Genome Atlas Network (2012) Comprehensive molecular portraits of human breast tumours. *Nature* 490(7418):61–70.
- 7 Van Loo P, et al. (2010) Allele-specific copy number analysis of tumors. *Proc Natl Acad Sci USA* 107(39):16910–16915.
- 8 Nik-Zainal S, et al. (2016) Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 534(7605):47–54.
- 9 Popova T, et al. (2016) Ovarian cancers harboring inactivating mutations in CDK12 display a distinct genomic instability pattern characterized by large tandem duplications. *Cancer Res* 76(7):1882–1891.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

- Q: 1_Please contact PNAS_Specialist.djs@sheridan.com if you have questions about the editorial changes, this list of queries, or the figures in your article. Please include your manuscript number in the subject line of all email correspondence; your manuscript number is 201610228.
- Q: 2_Please (i) review the author affiliation and footnote symbols carefully, (ii) check the order of the author names, and (iii) check the spelling of all author names, initials, and affiliations. Please check with your coauthors about how they want their names and affiliations to appear. To confirm that the author and affiliation lines are correct, add the comment “OK” next to the author line. This is your final opportunity to correct any errors prior to publication. Misspelled names or missing initials will affect an author’s searchability. Once a manuscript publishes online, any corrections (if approved) will require publishing an erratum; there is a processing fee for approved erratum.
- Q: 3_Please review and confirm your approval of the short title: Tandem duplications: Two distinct phenotypes. If you wish to make further changes, please adhere to the 50-character limit. (NOTE: The short title is used only for the mobile app and the RSS feed.)
- Q: 4_Please review the information in the author contribution footnote. PNAS requires all articles to have the author contributions listed. Typically, for a Letter, “wrote the paper” is the only contribution. However, you may add others as you think necessary. The other standard contributions are “performed research,” “designed research,” “contributed new reagents/analytic tools,” and “analyzed data.”
- Q: 5_Per PNAS style, certain compound terms are hyphenated when used as adjectives and unhyphenated when used as nouns. This style has been applied consistently throughout where (and if) applicable.
- Q: 6_PNAS italicizes genes and alleles. All genes, alleles, and proteins will appear as indicated in your proofs. Please carefully check use of italics and capital letters throughout the proof and correct as necessary. If, by “■■■■ gene,” you mean “the gene that encodes protein ■■■■,” then italic type is not necessary. (NOTE: If all instances of a gene/allele should be changed, please make only one correction and indicate that it is a global change.)
- Q: 7_Please provide a departmental affiliation and postal code for affiliation "d."
-
-