

REPLY TO WATKINS ET AL.:

# Whole-genome sequencing-based identification of diverse tandem duplicator phenotypes in human cancers

Francesca Menghi<sup>a</sup> and Edison T. Liu<sup>a,1</sup>

Watkins et al. (1) note that tandem duplications (TDs) are a feature of two distinct cancer phenotypes, distinguished based on TD span size (1 Kb–2 Mb vs. 2–10 Mb) and breast cancer 1 (*BRCA1*) status (inactivation vs. wild-type), present at different frequencies in triple-negative breast cancer (TNBC). The authors suggest that our recent study (2) only captures the first type of TD phenotype (TDP), and that the second form of TDP may have implications in the assessment of genomic instability-based biomarkers of drug response.

We agree that the TDP manifests in at least two distinct genomic configurations. However, we believe that Watkins et al. (1) misidentified the two prevalent TDP configurations by exclusively basing their analysis on SNP array data and focusing on predetermined TD span sizes.

Indeed, when we compared the analytical precision of gold-standard whole-genome sequencing (WGS) vs. SNP array, we found unequivocally that SNP arrays were lacking in sensitivity. Only when we calibrated the threshold of SNP-array TDP calls using matching WGS data were we able to reach adequate sensitivity (0.8) (2). Thus, without training on WGS data, the predictive value of SNP array-based algorithms for the TDP would be problematic. Therefore, we limited all TD span-size analysis only to the WGS dataset.

In our study (2), we analyzed TD span-size distributions across 277 tumor genomes, examined by WGS and classified using a TDP score metric that accounts for both TD number and genomic distribution. We identified two distinct TD span-size peaks, both strongly associated with the TDP: one at 10 Kb and another at 300 Kb (Fig. 1; see also figure 2B in ref. 2). More recently, we

observed that 91% of TDP tumors with 10-Kb TDs (which we now call TDP group 1) are *BRCA1*-deficient, whereas 94% of TDP tumors with 300-Kb TDs (i.e., TDP group 2) are *BRCA1* wild-type. This observation is further supported by the recent study by Nik-Zainal et al. (3) showing *BRCA1*-deficient TNBCs associated with TDs of span size < 100 Kb and *BRCA1* wild-type TNBC associated with TDs of span size 100 Kb–10 Mb, with a mode corresponding to the 100-Kb to 1-Mb interval.

Although we do observe large-span TDs (1–10 Mb), these rearrangements occur sporadically (<25 TDs per genome) and are prevalent only across tumors that do not score as TDP (2). We therefore believe that the 2- to 10-Mb TDP tumors identified by Watkins et al. (1), although rightfully associated with nontriple-negative, *BRCA1* wild-type breast tumors, were misclassified as TDP, likely because of the technical limitations of SNP array in detecting authentic TD rearrangements.

Taking these data together, we find that there are at least two distinct forms of TDP that can be segregated by span size and *BRCA1* status. As more WGS datasets are analyzed with respect to TD number, size, and distribution, we expect to identify additional if less-frequent TDP types and combinations (e.g., the *CDK12*-associated TDP described in ref. 4). However, we do not believe that large-span TDs contribute to the distinctive TDP configuration we showed to be associated with cisplatin sensitivity. Instead, these appear to be sporadic structural rearrangements that are found across many tumor types and are likely involved in the large-amplitude copy-number gain profiles typical of gene amplifications (5).

<sup>a</sup>The Jackson Laboratory for Genomic Medicine, Farmington, CT 06032

Author contributions: F.M. and E.T.L. designed research; F.M. performed research; F.M. and E.T.L. analyzed data; and F.M. and E.T.L. wrote the paper.

The authors declare no conflict of interest.

<sup>1</sup>To whom correspondence should be addressed. Email: edison.liu@jax.org.

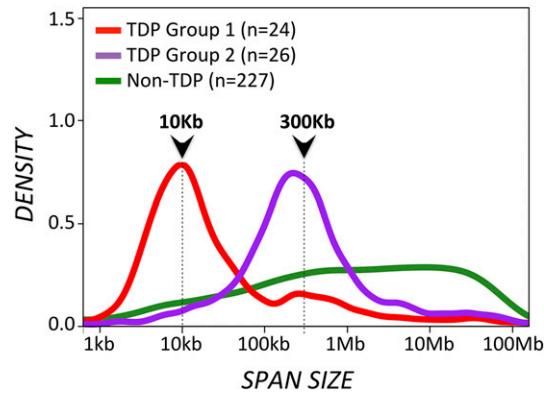


Fig. 1. TD span distributions for the TDP group 1, TDP group 2, and the non-TDP sample groups. The same WGS 277 genomes analyzed in ref. 2 are also analyzed here, based on TDP score (TDP vs. non-TDP) and TD span size distribution (TDP group 1 vs. TDP group 2). Cumulative TD span-size density curves are shown for each separate group. TDP group 1 and group 2 tumors feature narrow TD span-size distributions with discrete modal peaks at  $\sim 10$  Kb and  $\sim 300$  Kb, respectively. In contrast, non-TDP samples feature a plateau-like distribution of much larger TD span sizes, ranging from  $\sim 1$  Mb to  $\sim 10$  Mb.

- 1 Watkins J, Tutt A, Grigoriadis A (2016) Tandem duplications contribute to not one but two distinct phenotypes. *Proc Natl Acad Sci USA* 113:E5257–E5258.
- 2 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.
- 3 Nik-Zainal S, et al. (2016) Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 534(7605):47–54.
- 4 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.
- 5 Inaki K, et al. (2014) Systems consequences of amplicon formation in human breast cancer. *Genome Res* 24(10):1559–1571.