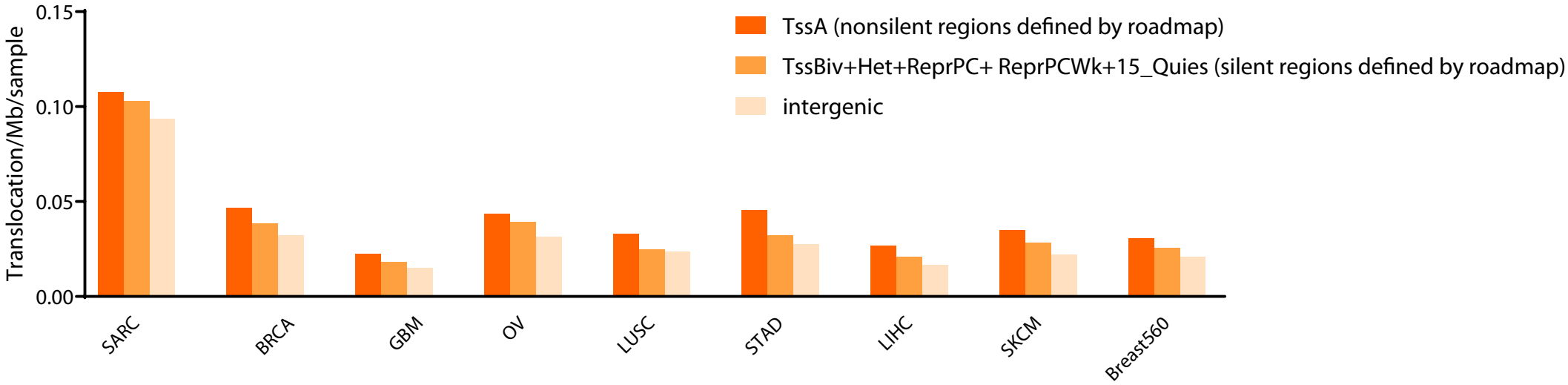

Figure 1. Distribution of chromosomal translocations with SVs from a nature paper and annotation files downloaded in roadmap is quite consisten with the results using SVs from PCACWG and GDC RNA-seq annotations. SVs for SARC, BRCA, GBM, OV, LUSC, STAD, LIHC and SKCM were downloaded in PCAWG (ICGC+TCGA), and Breast560 represents translocations sequenced in a nature paper (PMID: 27135926). The whole genome was classified into three types including genes with promoters overlap with active TSS regions, genes with promoters overlap with heterochromatin, bivalent/poised TSS, repressed polycomb, weak repressed polycomb or quiescent regions, and intergenic regions. Translocation number in each region was calculated and normalized by sample number and region size.

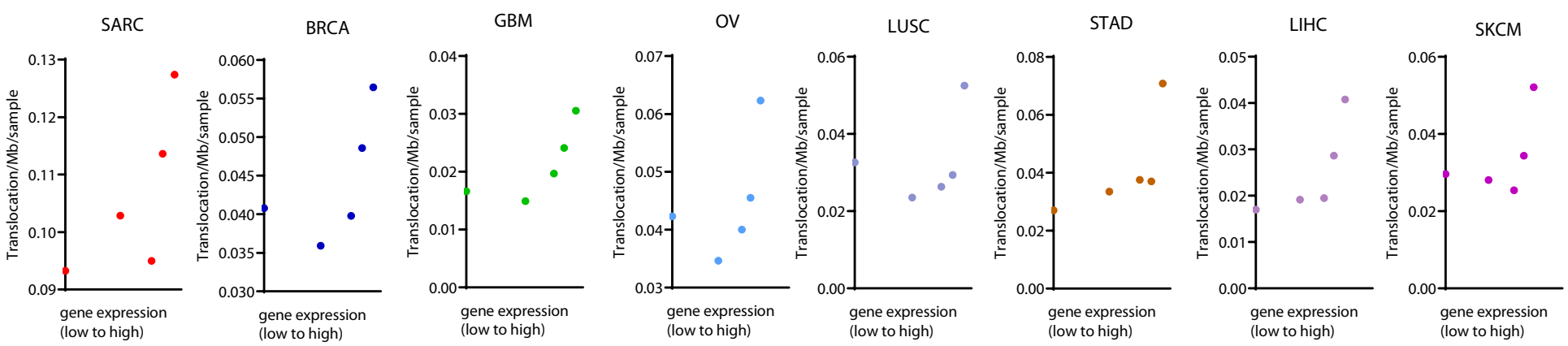

Figure 2. Role of gene expression in chromosomal translocation density with SVs from PCAWG as well as the nature paper and roadmap annotations. All silent genes defined by roadmap annotations were classified as the first bin while all the other genes were sorted by gene expression data from GDC and then divided into 4 bins with the same number of genes. Total translocation number for each bin was analyzed and normalized by bin size and total sample number. Translocation density increases in nonsilent regions with gene expression levels for BRCA, GBM, OV, LUSC, STAD, LIHC and SKCM.

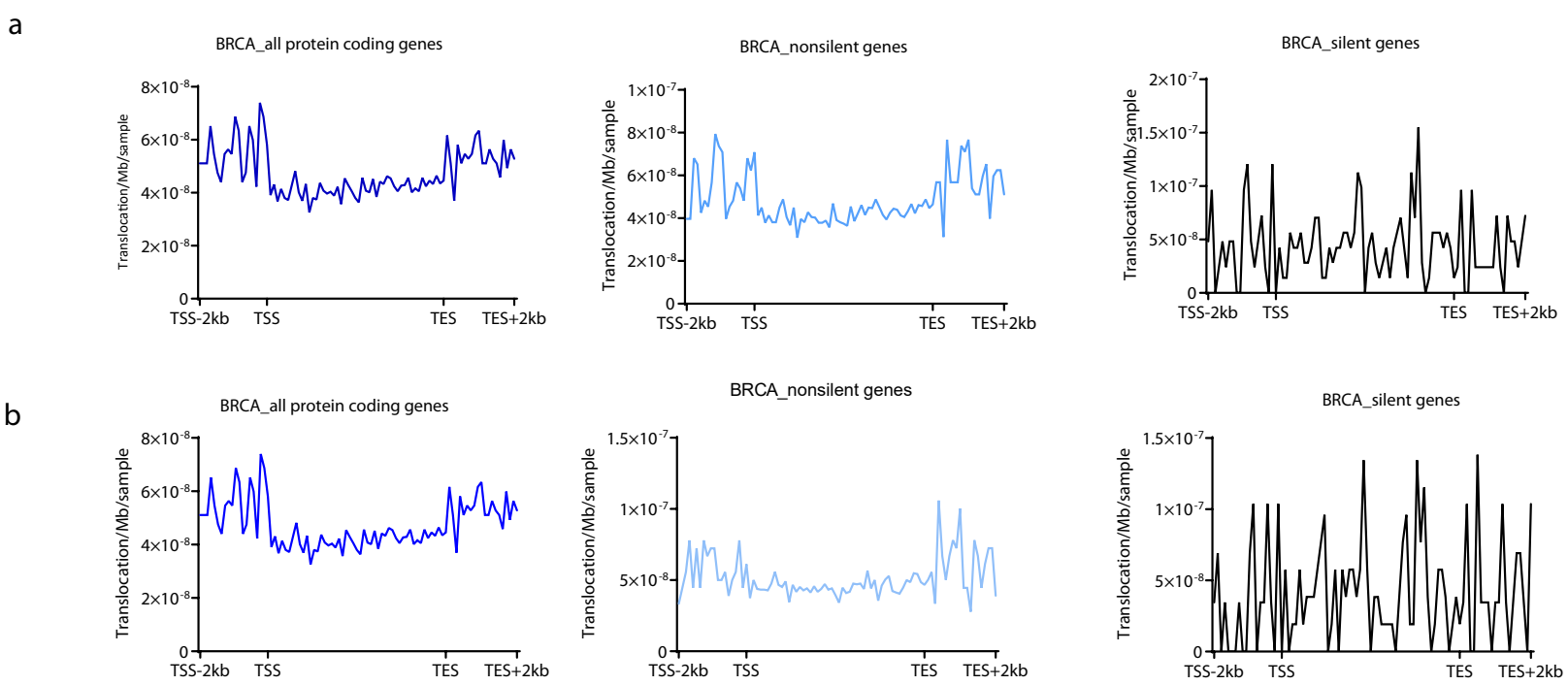

Figure 3. Translocation profles for breast cancer present similar distribution trend when the whole genome is annotated by (a) GDC RNA-seq values or (b) roadmap annottaions