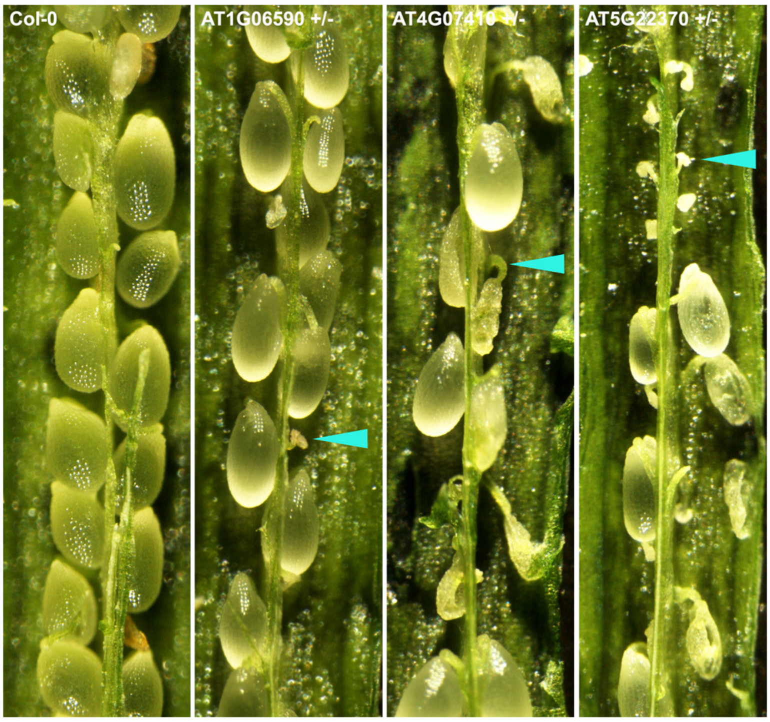

Figure S1. Seedpods of the analyzed hemizygous T-DNA insertion lines (+/-). Aborted seeds are indicated by arrowheads. The number of aborted seeds was correlated with the proportion of expected homozygous seeds, being significant according to the $\chi 2$ statistical test.

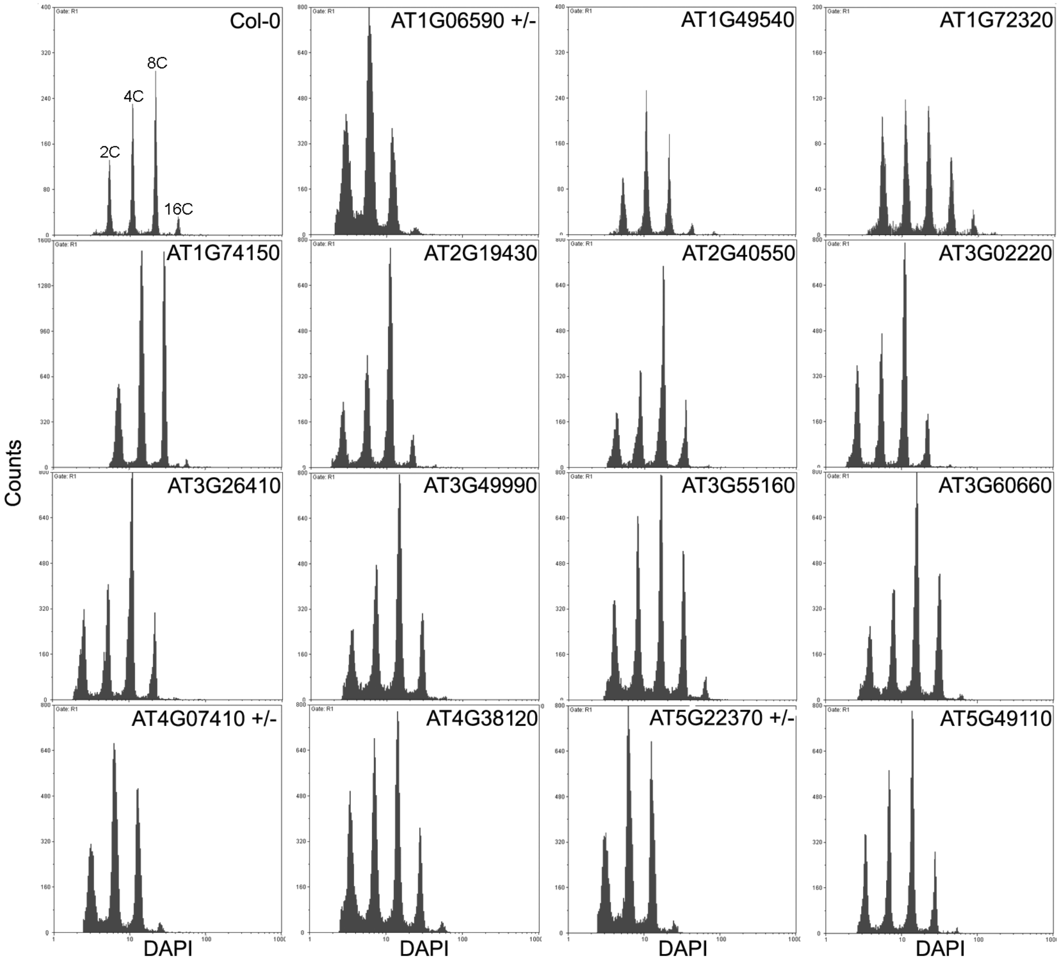

Figure S2. Flow cytometric DNA profiles of Col-0 and the analyzed T-DNA insertion mutants.

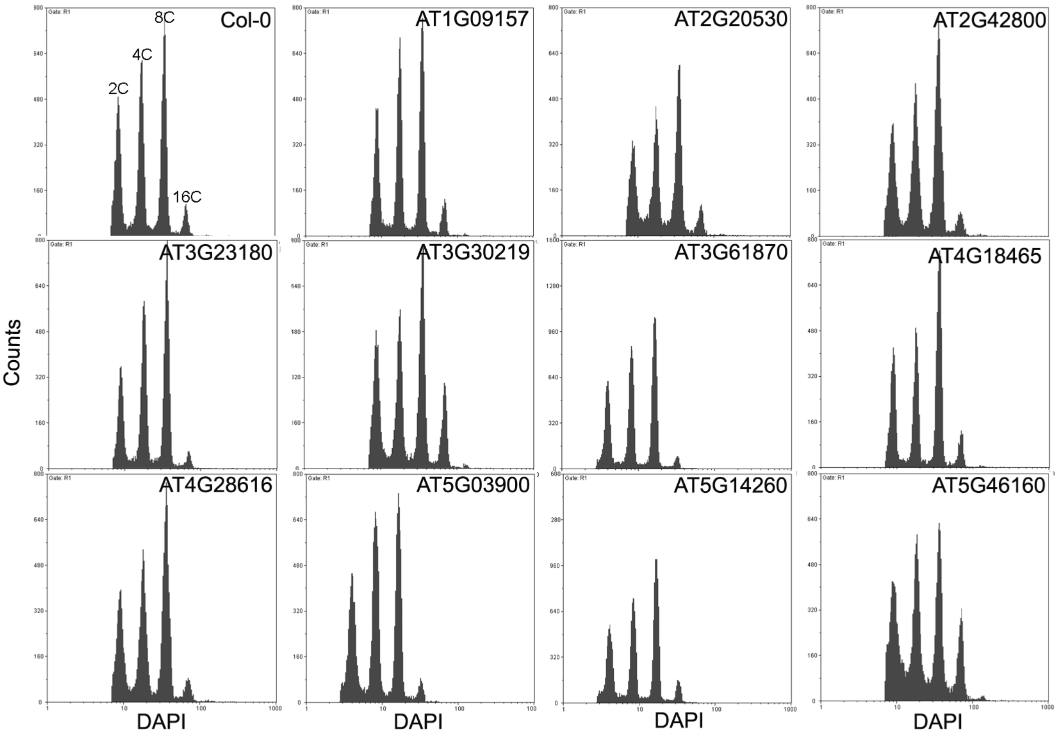

Figure S3. Flow cytometry DNA profiles of Col-0 and 11 randomly selected T-DNA insertion lines. All the lines harbored the T-DNA insertion in-between the translational start and stop codons. No significant changes were observed when compared to control plants (Col-0), except for the T-DNA insertion line of the AT5G46160 gene.

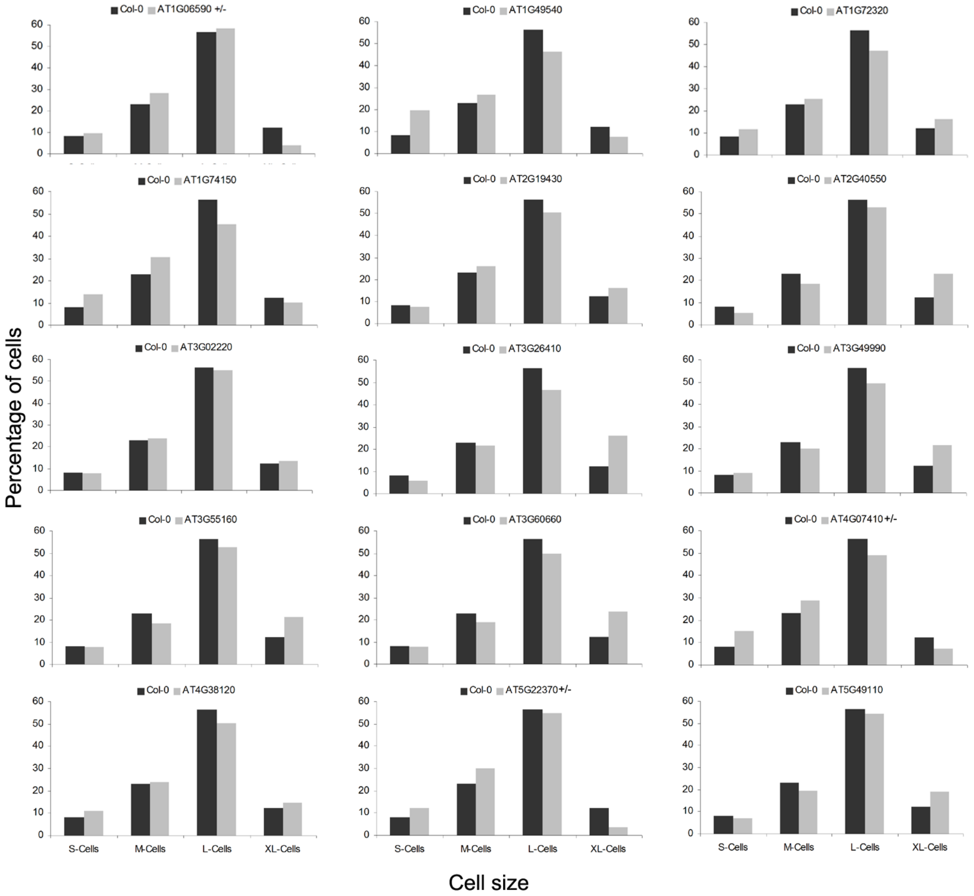

Figure S4. Cell size distributions of pavement cells in Col-0 and the analyzed T-DNA insertion lines. S, M, L and XL cells represent cells with a surface between 1 and 200 µm2, 201 and 700 µm2, 701 to 3000 µm2 and 3001 to 10000 µm2 respectively.

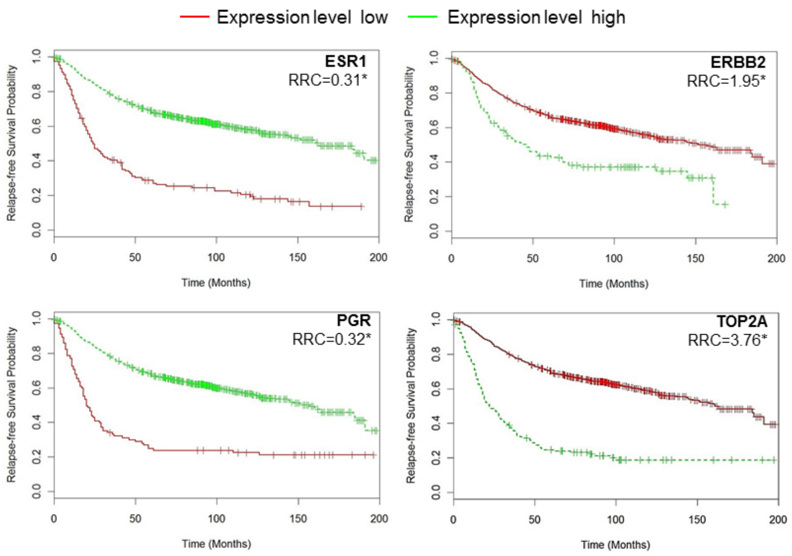

Figure S5. Association of reference genes known to play a role in breast cancer with specific cancer outcomes (Relapse risk). Cox survival plots for estrogen and progesterone receptors (associated with favorable disease outcome), ERBB2 (amplified gene associated with aggressive disease) and TOP2A (gene involved in DNA synthesis) were constructed. A clear association between increased gene expression levels and a diminished probability of relapse-free survival is shown for ERBB2 and TOP2A. The opposite is true for estrogen and progesterone receptors RRC, Relative Risk Coefficient; * represents statistically significant differences in the survival probability *P* <0.01

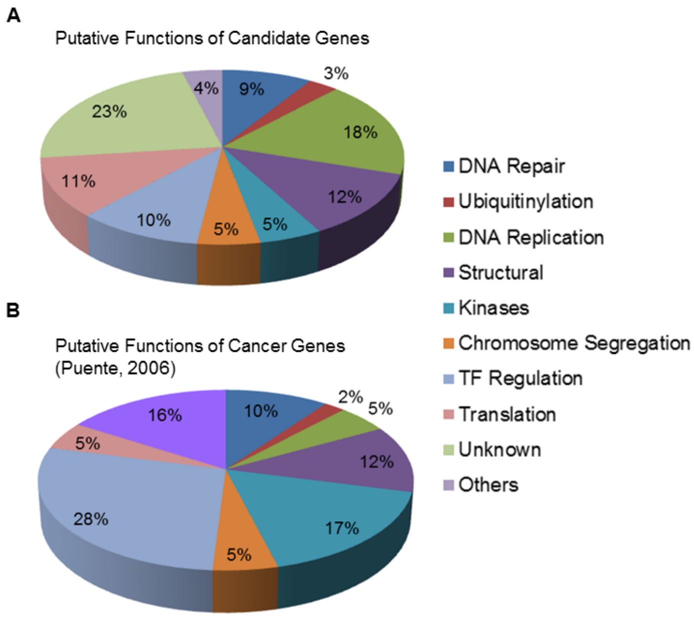

Figure S6. Comparison between **A** the most representative GO categories of the candidate genes final list described in the present research and **B** data of a cancer gene census study, reported by Puente, et al. Strikingly, a highly similar picture can be appreciated, in which all the GO categories described by Puente are represented in our gene list, although in some cases the proportion of genes within the specific category is not the same.

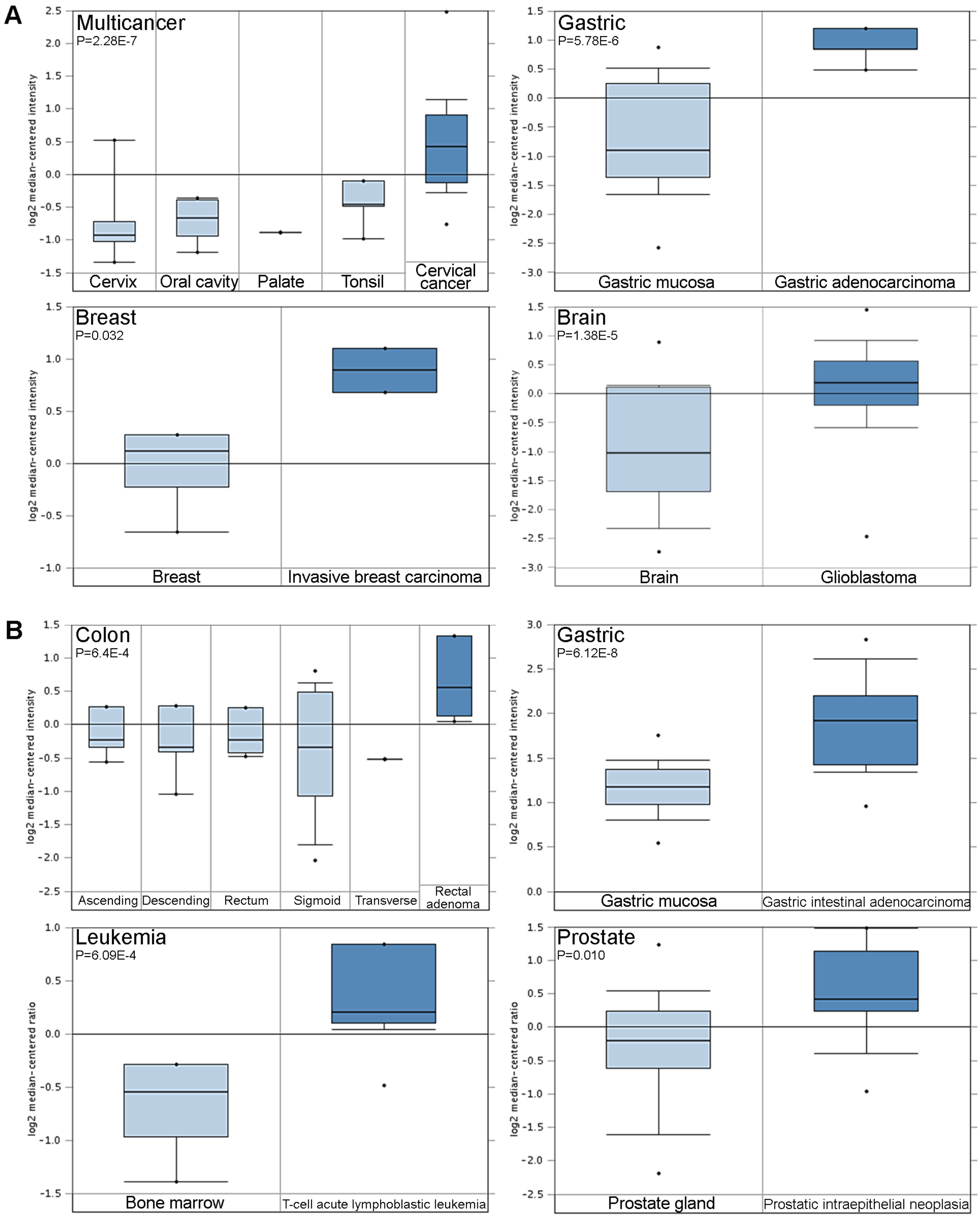

Figure S7 A. Examples taken from Oncomine (www.oncomine.org) of carcinomas in which HEATR6 transcript is significantly up-regulated (P-value<0.05) and B. Malignancies in which C14ORF21 genes is transcriptionally up-regulated (P-value<0.05). For all the given examples, light blue boxes represent normal tissues and dark blue boxes represent specific carcinomas.