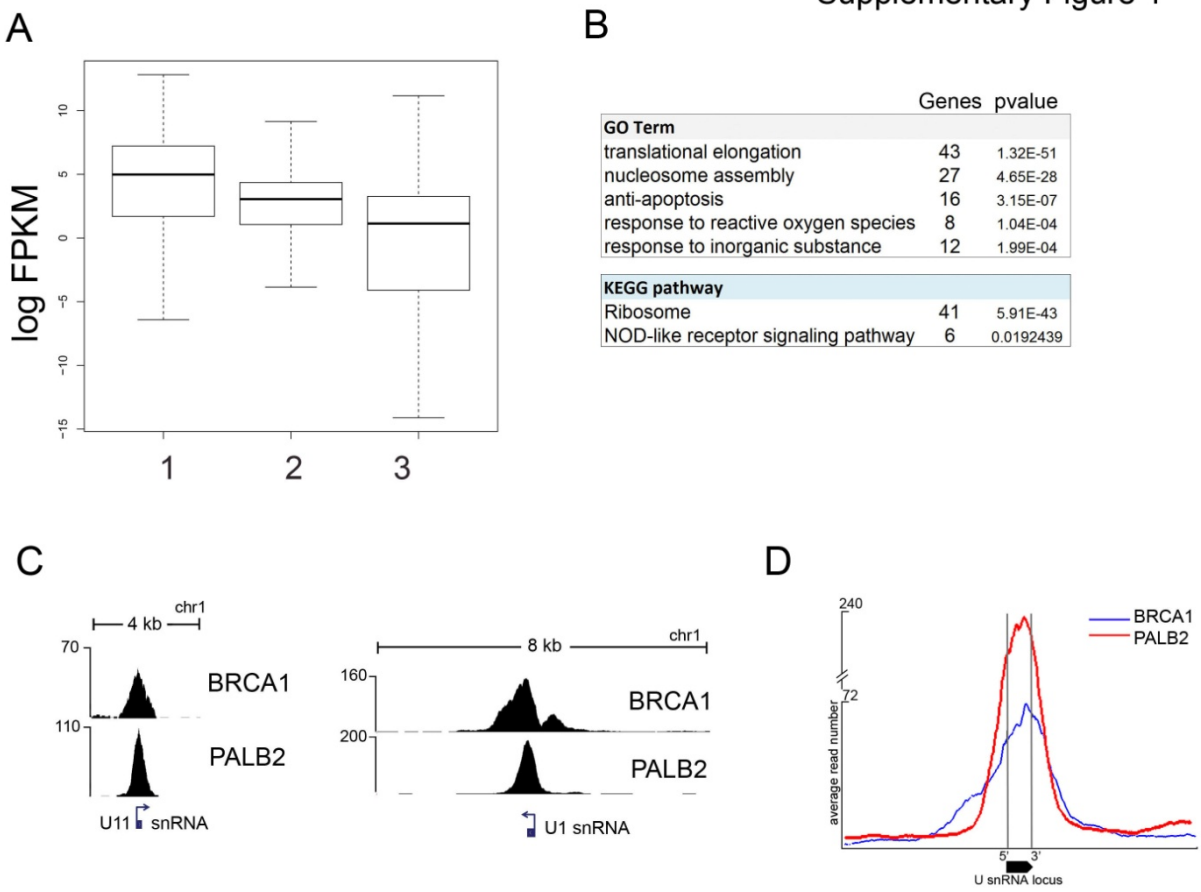
Supplementary Figure 1



Supplementary Figure 1

(A) Binding of BRCA1 and PALB2 correlates with highly transcribed genes. The box plot represents the average expression level of transcripts from the three classes of genes (calculated as logFPKM) identified by unbiased clustering (Fig. 1A). FPKM values were calculated from the sequencing of the PolyA+ fraction of total RNA extracted from asynchronously growing MCF10A cells. The differences of median expression between the classes are statistically significant ($p<0.001$).

(B) Functional annotation of BRCA1/PALB2 common target genes. 373 target genes were functionally annotated using DAVID (http://david.abcc.ncifcrf.gov), and clustered according to either the GO FAT term or the KEGG PATHWAY collection (http://www.genome.ad.jp/kegg/pathway.html). The table reports a selection of

categories that are significantly enriched, number of genes belonging to each category and the associated p-value are also indicated.

(C) BRCA1 and PALB2 bind to the highly transcribed U snRNA genes (see Supplementary Table 1 for the full list). A UCSC Genome Browser hg18 snapshot is shown for two representative loci. The loci were confirmed to be actively transcribed based on the presence of RNAPII by ChIP analysis (data not shown).

(D) Average binding profiles of BRCA1 and PALB2 across all U snRNA target loci. Profiles were generated with seqMINER using Kmeans linear as the method of clustering, with the following parameters: left and right extension=1 kb, internal bins=10, flanking region bins=80 number of cluster=5.