Gene essentiality in cancer cell lines is modified by the sex chromosomes

Shahar Shohat, Ethel Vol and Sagiv Shifman

Supplemental Information:

Figure S1. Characterization of cell lines with different sex chromosome dosages.

Figure S2. Expression analysis of genes on the pseudoautosomal regions.

Figure S3. Enrichment of genes expressed predominantly in the testis.

Figure S4. Fine mapping of Y chromosome genes that modify the essentiality of their paralogs.

Figure S5. Genes with a sex-bias rate of somatic mutations in human cancers are enriched with testis-specific genes.

Figure S1

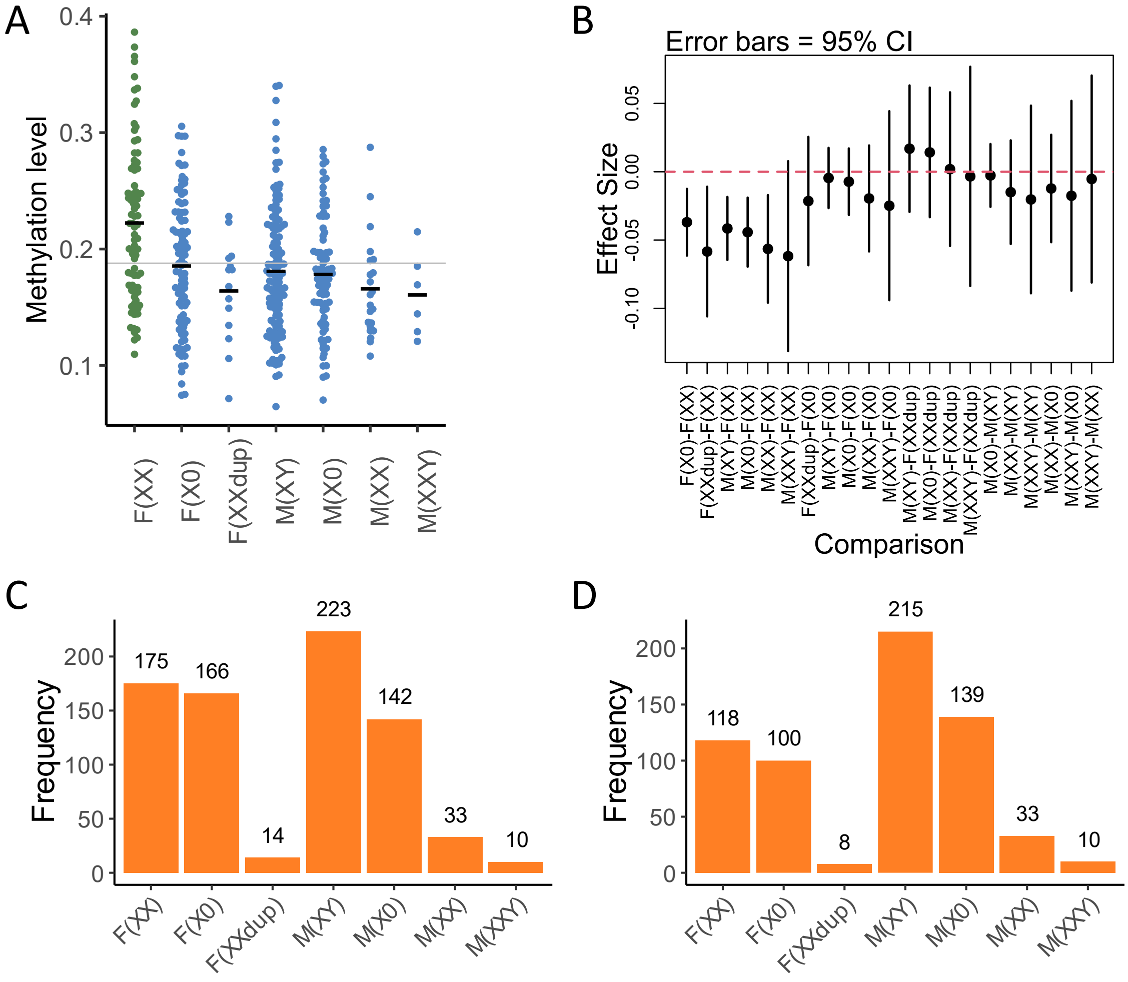

Figure S1. Characterization of cell lines with different sex chromosome dosages.

(A) Higher level of methylation in female XX cells. Values are the percentage of methylated CpGs in promotors of X chromosome genes (not including XCI escape genes). The difference between the groups is significant (P = 1.3x10^{-7}). The grey line is the average across the cell lines. The black lines are the mean value for each group. (B) Multiple comparisons of methylation level using the Tukey honest significant difference test. Values are differences in the observed means ± 95% confidence interval. (C) The number of cell lines in each group. (D) The number of cell lines in each group after excluding cell lines originated from male- and female-specific tissues.

Figure S2

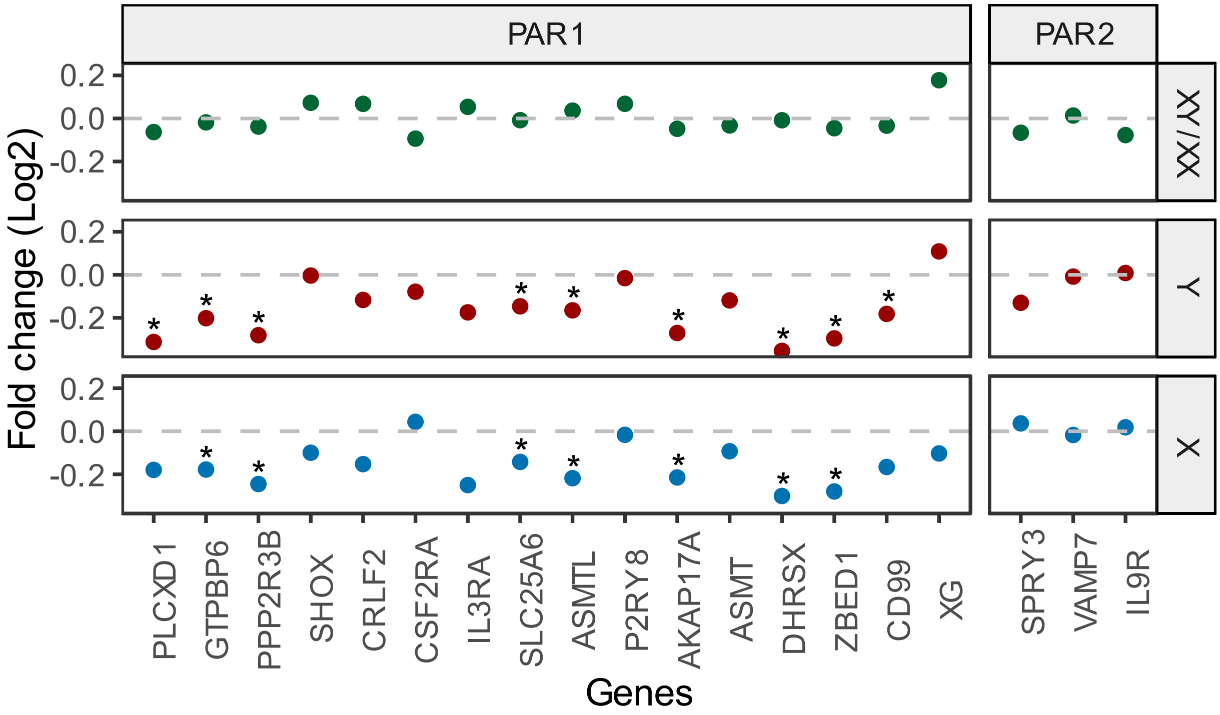

Figure S2. Expression analysis of genes on the pseudoautosomal regions.

The values are the fold change in expression (log2) for genes in PAR1 and PAR2. The data is shown for the difference between female XX and male XY cells (top plot) and the effect of Y (middle) and X chromosomes (bottom) on gene expression. Negative fold change means that the gene is expressed less in cells with one copy of the X or Y chromosome. The dashed line shows a log2 fold change = 0. The asterisks indicate an FDR < 0.05.

Figure S3

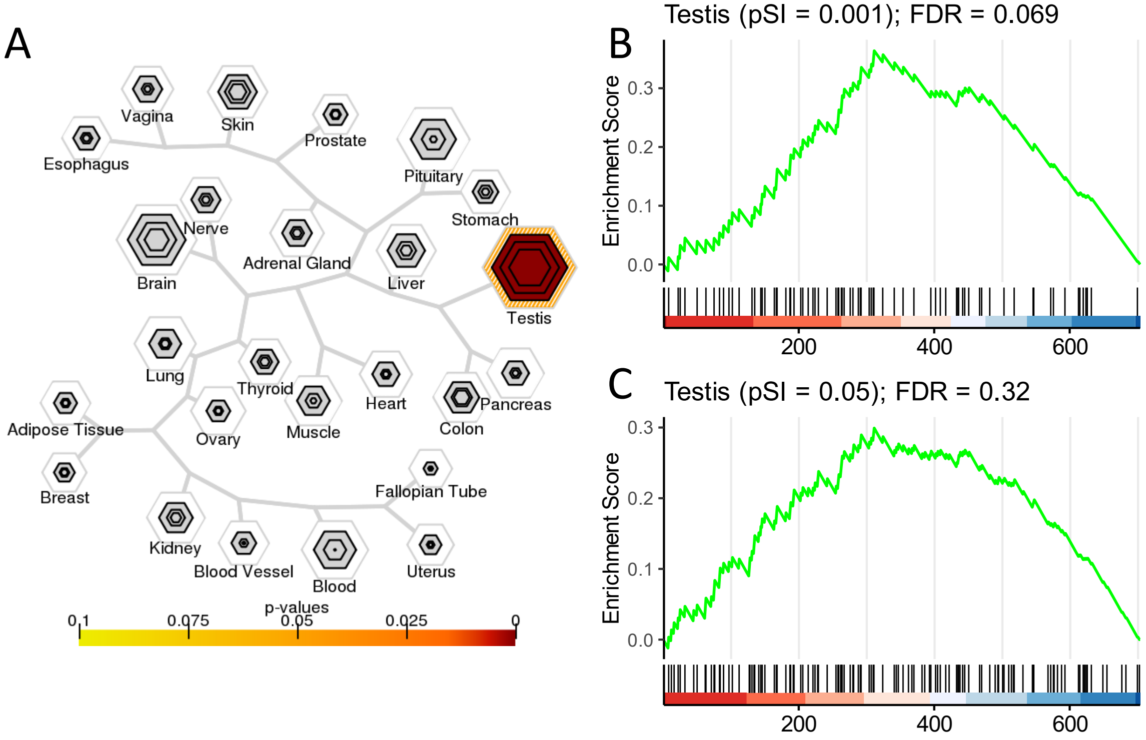

Figure S3. Enrichment of genes expressed predominantly in the testis.

(A) Tissue-Specific Expression Analysis (TSEA) of genes associated with the X chromosome. Each hexagon shows the enrichment of genes in a different tissue. The largest hexagon is for pSI < 0.05, and smaller embedded hexagons are for more stringent thresholds (pSI < 0.01, 0.001, and 0.0001). The size of the hexagons is scaled to the number of genes. The color represented the P-values. (B-C) Results of gene set enrichment analysis (GSEA) of genes associated with the X chromosome with two sets of genes preferentially expressed in the testis (pSI < 0.001 and 0.05). The position of genes in the gene set is marked with vertical bars. The genes are sorted based on the association significance with the X chromosome. The green curve is the enrichment score based on a weighted running sum. pSI, P-value for the specificity Index.

Figure S4

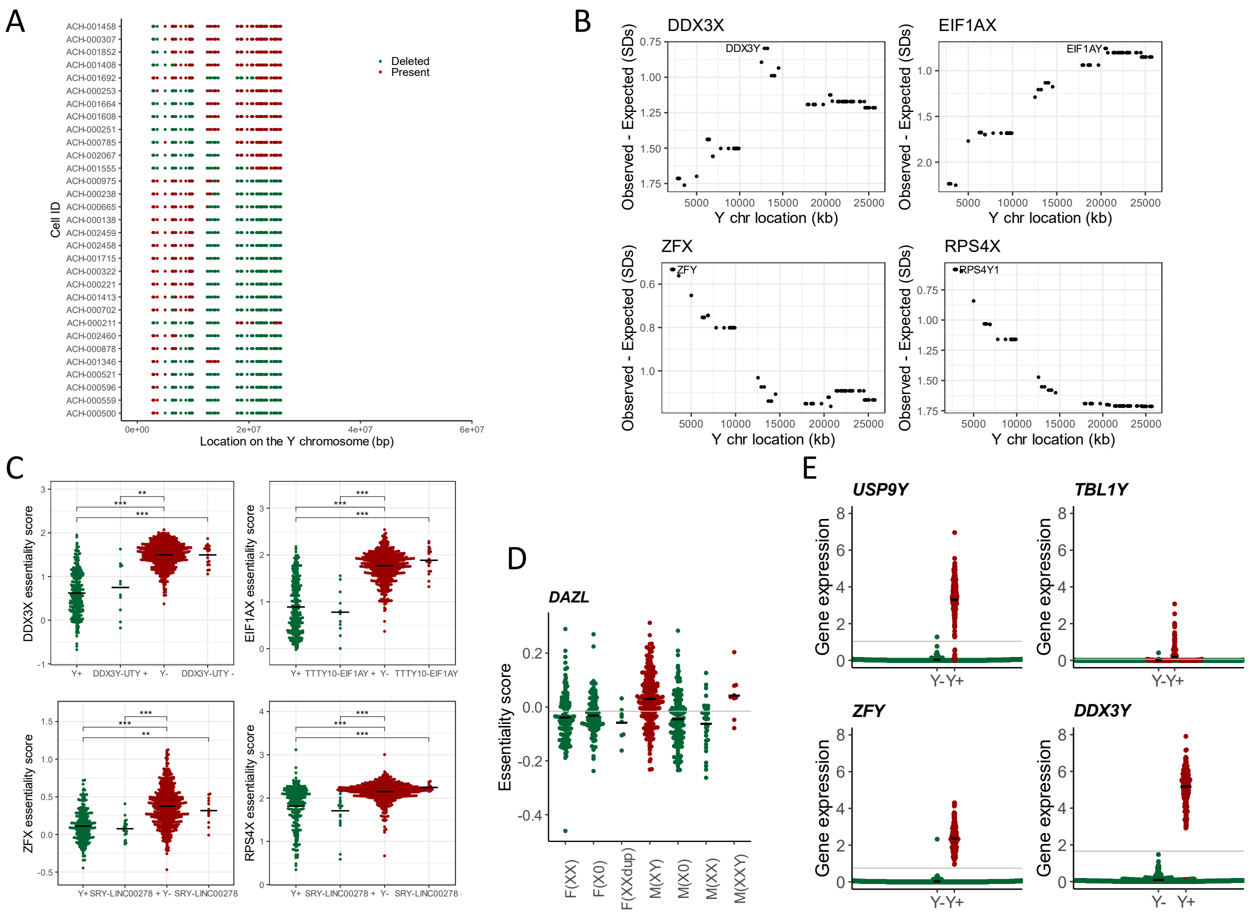

Figure S4. Fine mapping of Y chromosome genes that modify the essentiality of their paralogs.

(A) Partial deletion of the Y chromosome in 31 cell lines was used for fine mapping. Each point is a gene with available copy number data, where red means that the gene exists and green means that it is deleted. (B) Fine mapping results for four X-linked genes (DDX3X, EIF1AX, ZFX, and RPS4X). The values are the number of standard deviations for the difference between the expected and observed dependency scores, assuming the causal gene is in the indicated area. The labels are the names of the proposed causal genes. (C) Dependency score differences between Y^+ and Y^- cells and between cells with and without the region most likely to include the causal gene. (D) The distribution of essentiality scores for the DAZL gene, which is more essential for cells with a Y chromosome. (E) The expression of four Y-linked genes with high sequence identity to a paralog on the X chromosome. The grey line is the average across the cell lines. The black lines are the mean value for each group.

Figure S5

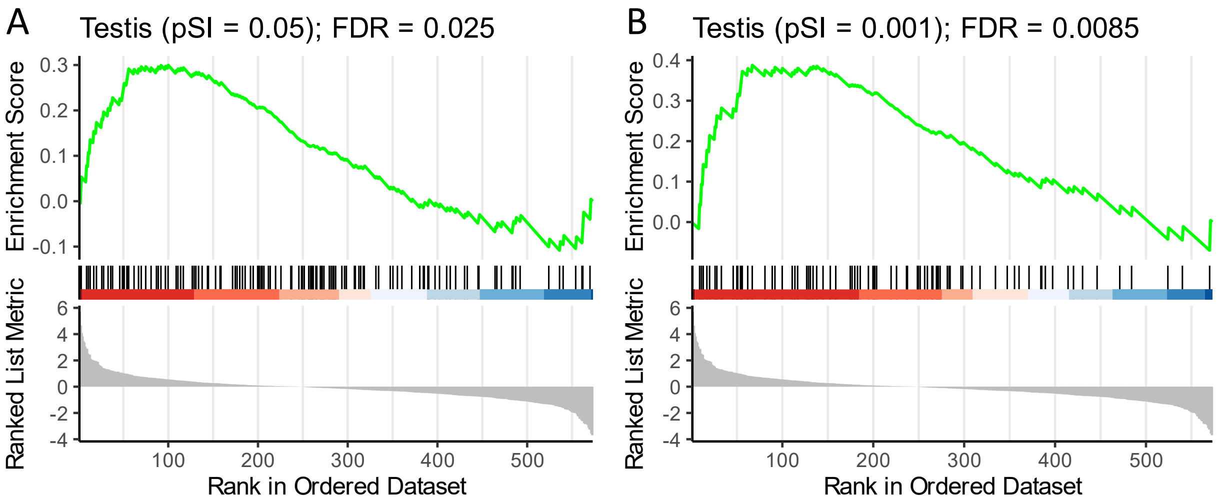

Figure S5. Genes with a sex-bias rate of somatic mutations in human cancers are enriched with testis-specific genes.

(A-B) Gene-set enrichment analysis (GSEA) plot with two sets of genes preferentially expressed in the testis (pSI < 0.05 and 0.001). The vertical black lines indicate the position of genes in the sets relative to all X-linked genes that are ranked from the most significant male-biased genes to the most significant female-biased gene. The green curve is the enrichment score based on a weighted running sum.