SUPPLEMENTARY INFORMATION

Supplementary Methods

Supplementary Table I

Supplementary Figure Legends

Supplementary Fig. S1

TOP2A transcript levels in ovarian and lung cancers stratified by *DLX4* expression.

Supplementary Fig. S2

Overexpression and knockdown of TOP2α.

Supplementary Fig. S3

DLX4 does not stimulate HR-mediated repair of DSBs.

Supplementary Fig. S4

DLX4 increases the magnitude of erroneous end-joining repair of DSBs.

SUPPLEMENTARY METHODS

Bioinformatic analyses. Analysis of *DLX4* and *TOP2A* transcript levels in tumors was performed using publicly available gene expression data from the following sources. Expression data of breast cancer (n=537 cases), lung cancer (n=155 cases) and ovarian cancer (n=594 cases) from the Cancer Genome Atlas (TCGA) Project were downloaded from the TCGA data portal site (http://tcga-data.nci.nih.gov/tcga/). Expression data of breast cancer (n=295 cases) from the study of Van de Vijver *et al* (32) were downloaded from the Division of Molecular Biology, Netherlands Cancer Institute website (http://bioinformatics.nki.nl/data.php). Expression data of breast cancer linked to patient outcomes following anthracycline-based chemotherapy were downloaded from the National Center for Biotechnology Information Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo). Datasets were as follows: [1] GSE25055, referred to as the Discovery Cohort in Hatzis *et al* (34) (n=310 cases of which 4 without outcomes data were excluded from analysis, where patients received anthracycline-taxane chemotherapy and then endocrine therapy if ER+); [2] GSE22093 (35) (n=103 cases of which 6 without outcomes data were excluded from analysis, where patients were treated with doxorubicin or epirubicin in combination with 5-fluorouracil and cyclophosphamide); [3] GSE20194 (36) (n=278 cases, where patients were treated with doxorubicin in combination with 5-fluorouracil, cyclophosphamide and paclitaxel); [4] GSE4779 (37) (n=102 cases, where patients were treated with epirubucin in combination with 5-fluorouracil and cyclophosphamide). Normalized values for genes as provided by contributors of each dataset were used. Where there were multiple probe sets for an individual gene, the mean value for the given gene for each case was used. Cases in each cohort were stratified according to transcript levels of *TOP2A* and *DLX4* in tumors, where transcript levels for each gene were defined as High (> upper quartile) and Low (< lower quartile).

Supplementary Table I:

IC_{50} values for chemotherapeutic agents in vector-control and +DLX4 stable tumor cell lines [1]

	MDA-MB-468			U2OS		
	Vector-control	+DLX4	*P*-value [2]	Vector-control	+DLX4	*P*-value [2]
Doxorubicin	100 ± 20 nM	215 ± 21 nM	*P* = 0.005	360 ± 68 nM	741 ± 98 nM	*P* = 0.004
Paclitaxel	1.4 ± 0.3 nM	1.1 ± 0.3 nM	*P* = 0.24	100 ± 10 nM	127 ± 40 nM	*P* = 0.33
5-Fluorouracil	98 ± 30 μM	110 ± 28 μM	*P* = 0.54	473 ± 40 μM	307 ± 61 μM	*P* = 0.017
Cyclophosphamide	n.d. [3]	n.d. [3]		7.5 ± 0.3 mM	7.7 ± 0.6 mM	*P* = 0.75

1. Cell viability was measured by MTT assay at 2 d after incubation with each drug. Shown are mean ± s.d. of IC_{50} values (concentrations that induce 50% loss of cell viability) determined from 3 independent assays.
2. Significance of difference in IC_{50} values between vector-control and +DLX4 cell lines determined by Student *t*-test
3. IC_{50} value not determined (n.d.). Cells resistant at maximum concentration used (10 mM).

SUPPLEMENTARY FIGURE LEGENDS

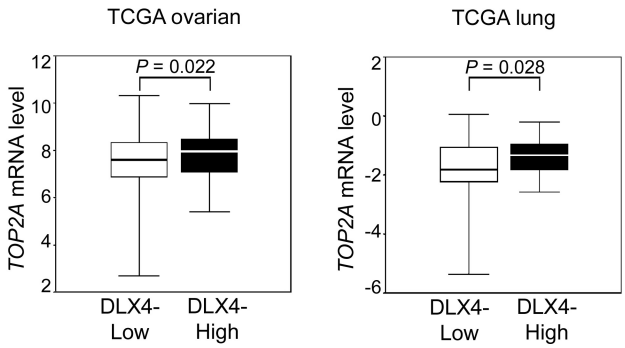
Supplementary Fig. S1. *TOP2A* transcript levels in ovarian and lung cancers stratified by *DLX4* expression. Cases of ovarian cancer (n=594) and lung cancer (n=155) in TCGA datasets were stratified according to *DLX4* expression in tumors, where *DLX4* transcript levels were defined as High (> upper quartile) and Low (< lower quartile) in each dataset. Significance of differences in *TOP2A* transcript levels (log2 scale) between upper and lower quartile sub-groups was evaluated by Mann-Whitney *U*-test.

Supplementary Fig. S2. Overexpression and knockdown of TOP2α. Western blot of TOP2α levels in **[A]** MDA-MD-468 cells transiently expressing TOP2α fused to GFP (indicated by arrow) and **[B]** TOV112D cells transiently expressing non-targeting shRNA and shRNAs targeting two different sites of *TOP2A* (shA-TOP2A, shB-TOP2A).

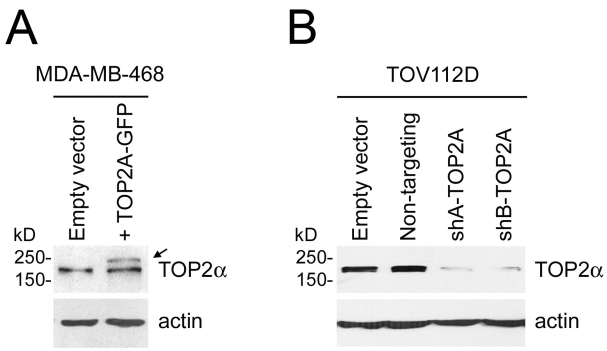
Supplementary Fig. S3. DLX4 does not stimulate HR-mediated repair of DSBs. **[A]** Schematic diagram of the HR repair assay using the U2OS DR-GFP reporter cell line. Repair of the I-*Sce* I-induced DSB in the *SceGFP* gene is facilitated by using the adjacent *iGFP* gene as a template. Repair results in loss of the I-*Sce* I restriction enzyme site and restoration of GFP expression by the *SceGFP* gene. **[B]** Western blot of DLX4 and BRCA1 levels in transfected cells of the U2OS reporter line. **[C]** Restoration of GFP expression in transfected cells was detected by flow cytometry and is expressed as % of GFP+ cells. **[D]** A 650 bp fragment around the DSB was amplified and digested with I-*Sce* I. Repair of the *SceGFP* gene was confirmed by detection of the 650 bp PCR product that is resistant to I-*Sce* I digestion. Shown are representative results of 3 independent assays.

Supplementary Fig. S4. DLX4 increases the magnitude of erroneous end-joining repair of DSBs. Representative examples of sequences of pUC19 plasmid DNA isolated from individual white colonies of end-joining assays using nuclear extracts of **[A]** vector-control and +DLX4 U2OS cells and **[B]** TOV112D cells expressing non-targeting and *DLX4* shRNAs. Indicated are deleted nucleotides (dotted lines) surrounding the *Eco* RI-cut DSB and tracts of nucleotide microhomology around breakpoint junctions (underlined).

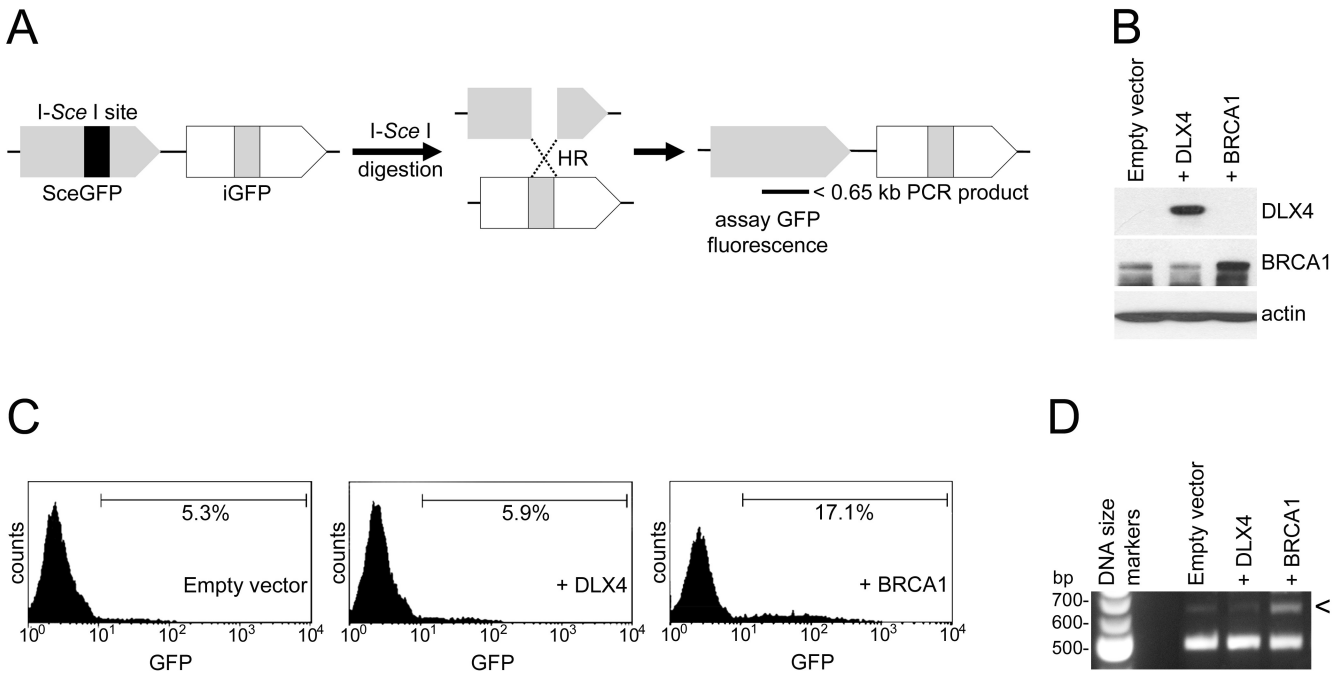
Supplementary Fig. S1



Supplementary Fig. S2



Supplementary Fig. S3



A

U2OS extracts

intact pUC19

Eco RI

```
ATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAG
```

Empty vector

Deletion (bp)

```
Clone 1  ATGTGCTGGAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGG----------TACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAG  19
Clone 2  ATGTGCTGGAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCC----------CCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAG  19
Clone 3  ATGTGCTGGAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTG-------GCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAG   7
```

\+ DLX4

```
Clone 1  ATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCA----------TTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAG  42
Clone 2  ATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACG----------AGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAG  26
Clone 3  ATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGAC----------CTGCAGGCATGCAAG  59
Clone 4  ATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGT----------GCAGGCATGCAAG  59
```

B

TOV112D extracts

intact pUC19

```
GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAA
AATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGT
AACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTT
TCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCT
TTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC
TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG
```

Non-targeting shRNA

Deletion (bp)

```
GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGC----------
----------
----------
Clone 1  ----------  722
----------
----------
------GTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG

GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAA
AATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGT
AACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACG----------
Clone 2  ----------  257
----------GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC
TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG
```

shA-DLX4

```
GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAA
AATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGT
AACGCCAGGGTTTTCCCAGTCACGACG----------AGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTT
Clone 1  TCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCT  26
TTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC
TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG

GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAA
AATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGT
AACGCCAGGGTTTTCCCAG----------GAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTT
Clone 2  TCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCT  33
TTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC
TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG
```