

## Supplementary Tables

Knockout	Pathway	Function during DSB repair
Erc1	HR	Removes non-homologous 3' single-stranded tails in HR and NHEJ. Also active in nucleotide excision repair and interstrand crosslink repair. <sup>1</sup>
Rad52	HR	Promotes HR by displacing RPA from resected ends, allowing Rad51 to bind and start homology search. <sup>2</sup>
Trp53	HR	Multifunctional tumour suppressor gene primarily inducing cell cycle arrest and apoptosis in response to DNA damage. <sup>3</sup>
Wrn	HR	Helicase and exonuclease, promotes NHEJ by suppressing resection. <sup>4</sup>
Lig1	MMEJ	Ligates broken ends in MMEJ. <sup>5</sup>
Lig3	MMEJ	Ligates broken ends in MMEJ. Also essential for mitochondrial DNA repair. <sup>5</sup>
Parp1	MMEJ	Binds to resected DNA ends and biases pathway choice towards MMEJ. <sup>6</sup>
Nbn	MMEJ	Part of the MRN complex which performs resection of DSBs, initiating MMEJ or HR. <sup>7</sup>
Polq	MMEJ	Encodes Polymerase Theta, which displaces RPA from ssDNA, anneals sequences with microhomology and adds untemplated and templated nucleotides to the ends of DSBs. <sup>8</sup>
Dclre1c	NHEJ	Encodes Artemis, an endo- and exonuclease active during NHEJ. <sup>9</sup>
Poll	NHEJ	Polymerase. Prefers substrates with complementary sequence between ends. Also active in MMEJ. <sup>10</sup>
Polm	NHEJ	Polymerase. Prefers downstream base as template. <sup>10</sup>
Trex1	NHEJ	A 3' to 5' exonuclease. Essential for ssDNA-mediated gene editing. <sup>11</sup>
Trp53bp1	NHEJ	Protects broken ends from resection by MRN complex and helps recruit other repair factors. <sup>12</sup>
Lig4	NHEJ	Primary ligase for NHEJ. Forms a synaptic complex with Xlf and Xrcc4 to hold broken ends together. <sup>13</sup>
Prkdc	NHEJ	Encodes DNA-PKcs, a kinase that binds to Ku complex and enables end processing by Artemis (product of Dclre1c). Regulates cell cycle arrest in response to DSBs. <sup>14</sup>
Xlf	NHEJ	Forms a synaptic complex with Lig4 and Xrcc5 to hold broken ends together. <sup>13</sup>
Xrcc5	NHEJ	Forms a synaptic complex with Lig4 and Xlf to hold broken ends together. Stimulates activity of Lig4 for ligation of broken ends. <sup>13</sup>

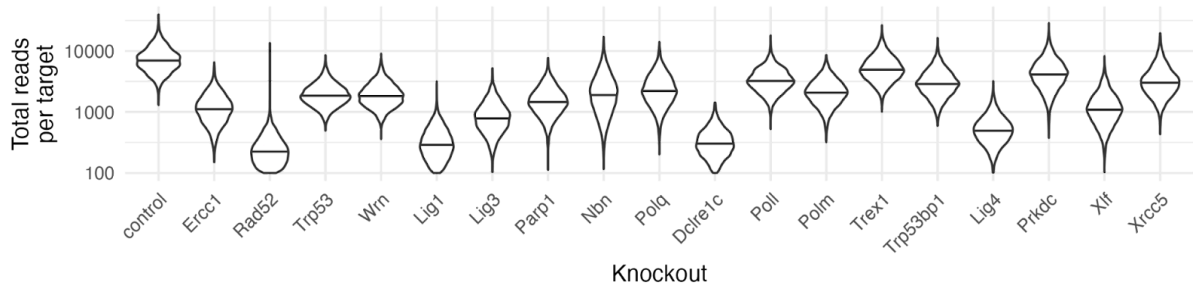
**Supplementary Table 1.** The panel of DNA repair gene knockouts used in this study, organised by pathway. Pathway assignments for each gene are based on the description of the function presented.

<b>Target</b>	<b>sgRNA</b>	<b>Forward primer for short-read sequencing</b>	<b>Reverse primer for short-read sequencing</b>
HPRT exon 2	TATACCTAATCATTATGCCG	GCAGATTAGCGATGATGAACC	ACACCACACACACACCCTCT
HPRT exon 3.1	ACTTGCTCGAGATGTCATGA	TGTGTCCTGTAAAAGTTTAATGTGTAA	AATCCAGCAGGTCAGCAAAG
HPRT exon 3.2	AGCCCCCTTGAGCACACAG	GGACTGAAAGACTTGCTCGAGAT	CAGTAGCTTTCAGTCTGATAAAA
HPRT exon 3.3	GCCCTCTGTGTGCTCAAGGG	GGACTGAAAGACTTGCTCGAGAT	CAGTAGCTTTCAGTCTGATAAAA
HPRT exon 3.4	AAAATCTACAGTCATAGGAA	AAGTTCTTTGCTGACCTGCTG	TATGTAGCATAGTTTGCAAGTT
HPRT exon 4	GATGATCTCTCAACTTTAAC	CACAAGAATTTGCATTTAAGTGATG	CCTCTAAGTAAGTGGTTGAAAGC
HPRT exon 7	TTAACAGCTTGCTGGTGAAA	GTAGTCTTTCCATATGCC	AACCCTAAGCCATTTTCATAGT

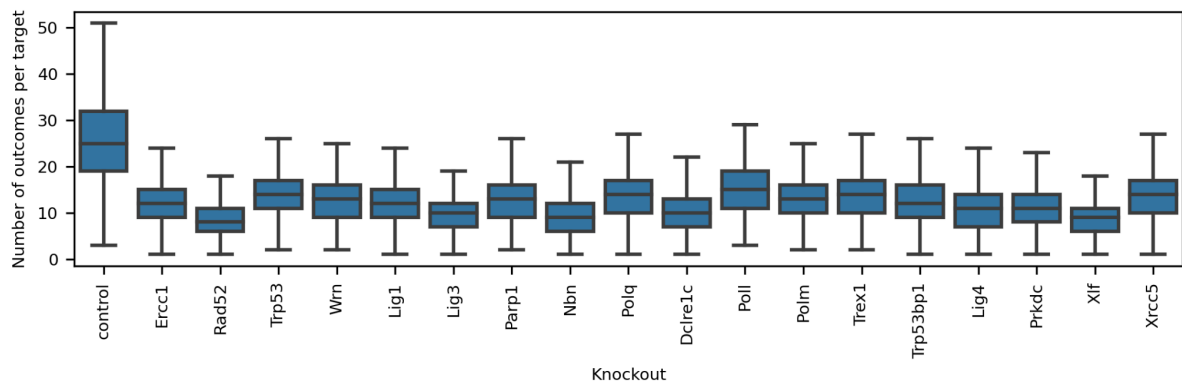
**Supplementary Table 2.** PCR primers used to amplify endogenously targeted sequences

# Supplementary Figures

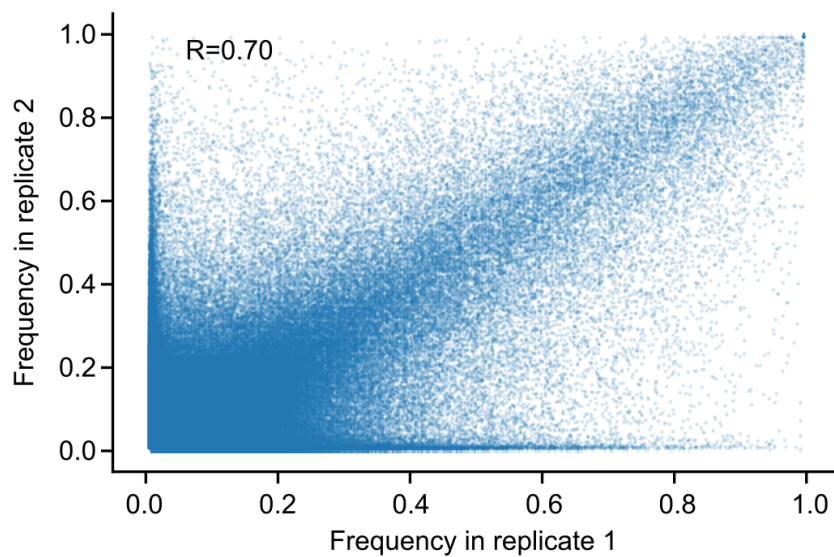
**a**

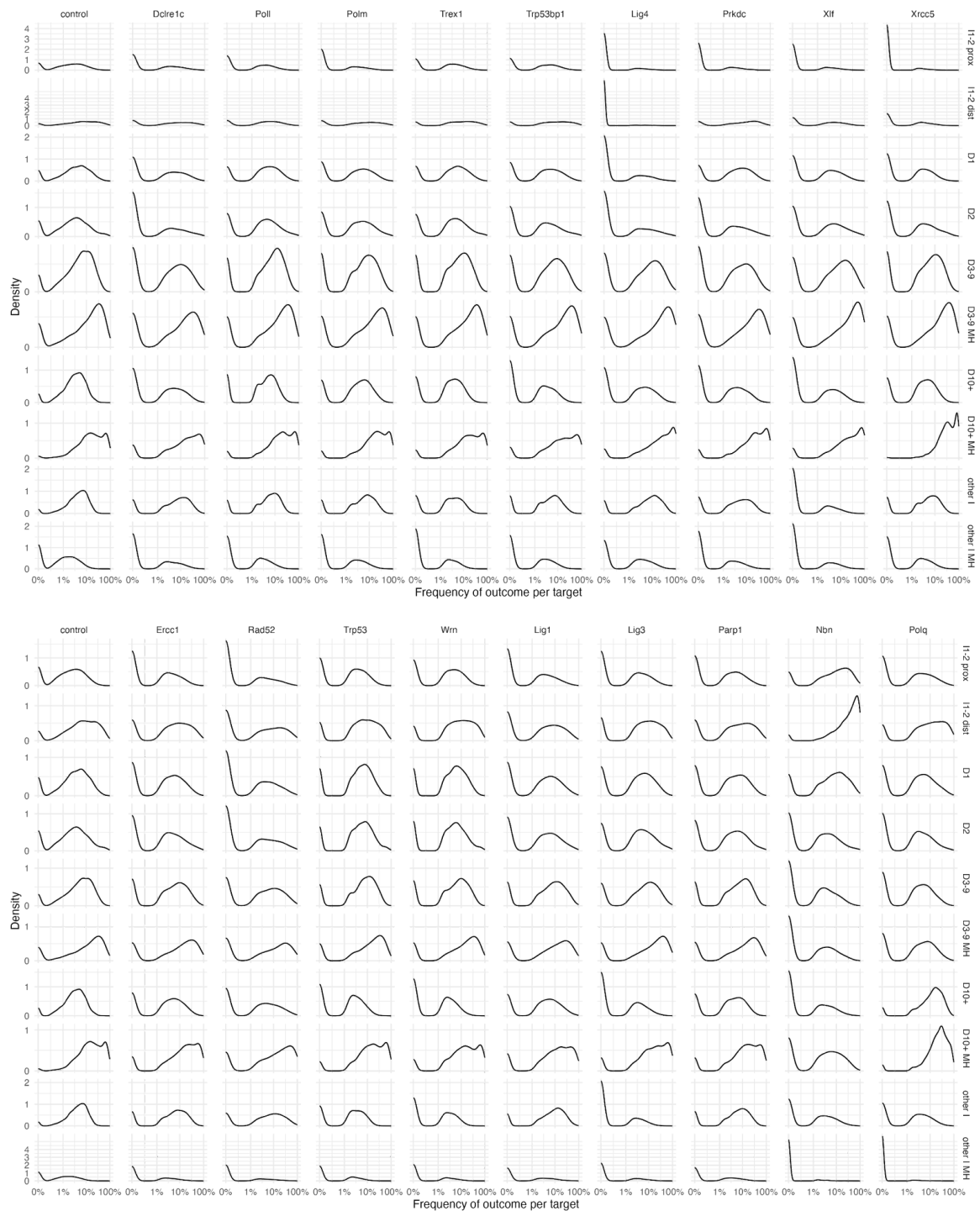


**b**

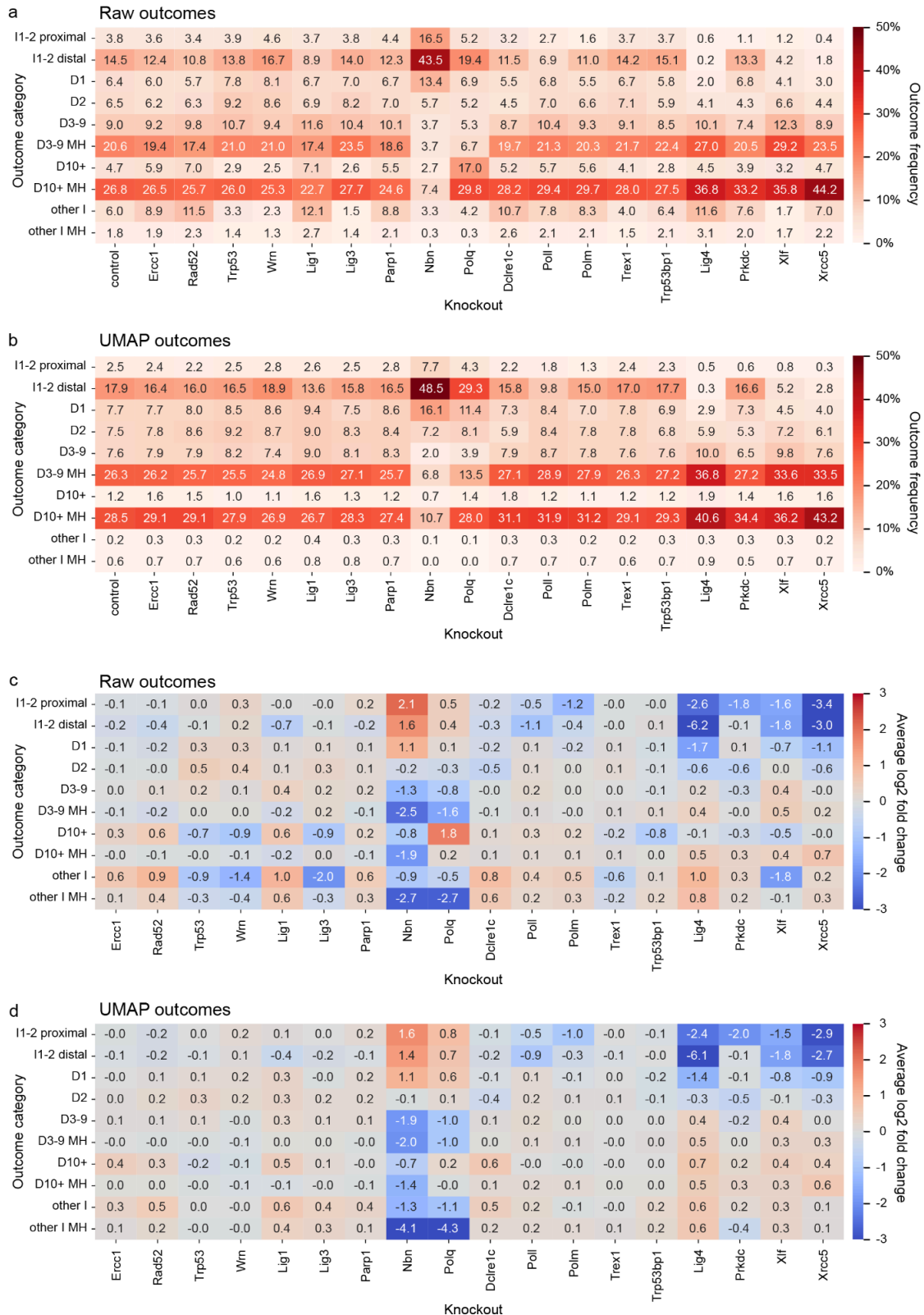


**c**

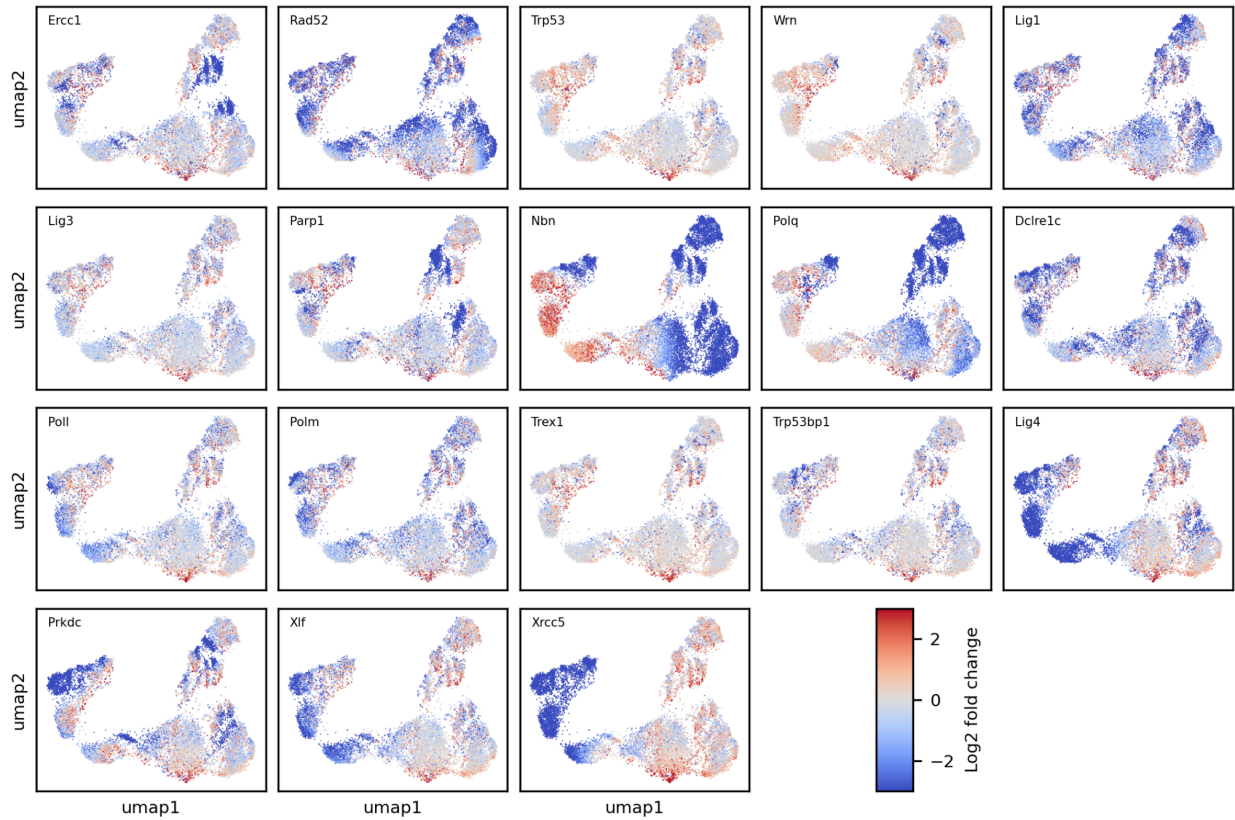


**d**

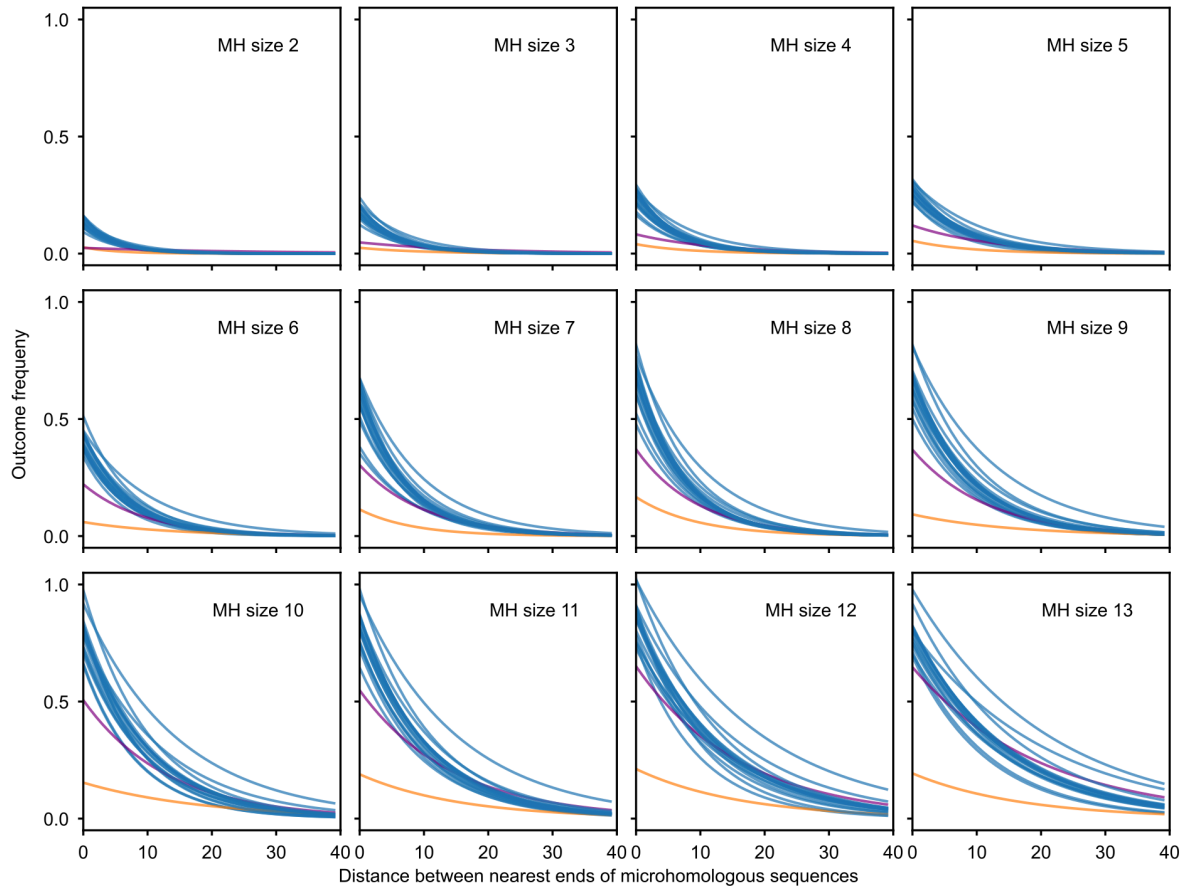
**Supplementary Figure 1. (a).** Density (x-axis) of number of reads per target (y-axis; logarithmic scale) for screens in mouse embryonic stem cell lines with different repair gene knockouts (x-axis violins). Horizontal bar: median. Every target had at least 100 reads in every knockout. **(b).** Number of outcomes per target (y-axis) for screens in mouse embryonic stem cell lines with different repair gene knockouts (x-axis). Box: median and quartiles; whiskers: 1.5x interquartile range. **(c).** Representative reproducibility of outcomes. Frequency of individual repair outcomes (markers) in replicate 1 (x-axis) and replicate 2 (y-axis) of one screen. Frequency of 1 = 100%. **(d).** Diversity of outcomes. Frequency across targets (y-axis) of repair category outcome frequency within target (x-axis, logarithmic scale) of different outcome categories (rows) in different knockouts (columns).



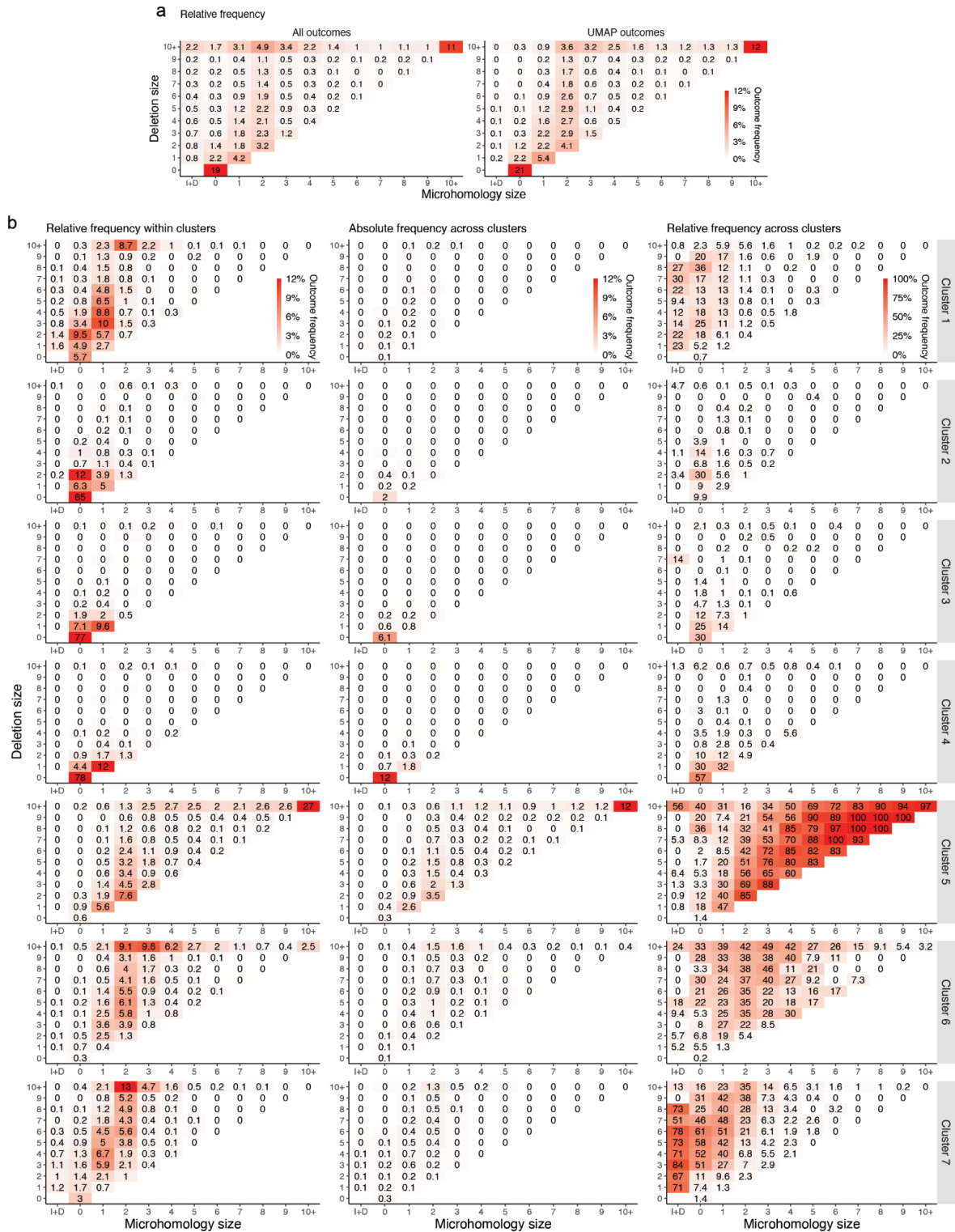
**Supplementary Figure 2. (a,b)** Average fraction of mutated reads across all targets (annotation, colour) of each outcome category (y-axis) observed in each knockout (x-axis). **(c,d)** Average of log-fold change across all targets (annotation, colour) of each outcome category (y-axis) observed in each knockout (x-axis). **(a,c)** Shows raw outcomes and **(b,d)** shows UMAP-filtered outcomes. D=deletion, I=insertion, MH=microhomology.



**Supplementary Figure 3.** UMAP projection of outcomes coloured by log-fold change in each knockout. UMAP projection of outcomes coloured by log-fold change (colour) of each outcome (dot) in each knockout (panels).

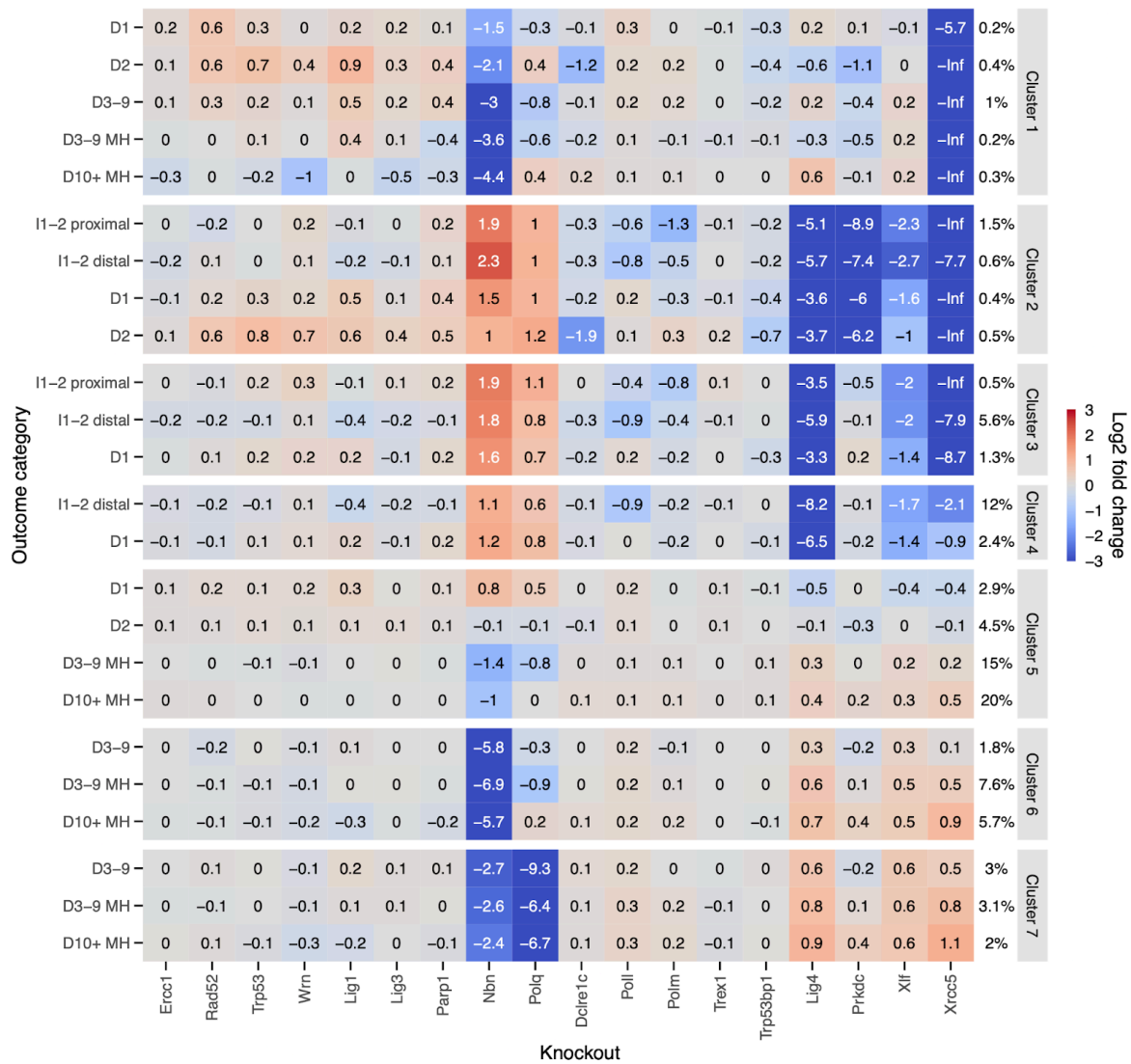


**Supplementary Figure 4.** The relationship between outcome frequency, size of microhomology sequence, and distance between microhomology sequences. Frequency of microhomology deletion (y-axis) vs the distance between microhomology sequence (x-axis) for a single length of microhomology (panels, labels) for each knockout (blue lines), with Nbn (orange) and Polq (purple) highlighted. Frequency of 1 = 100%.

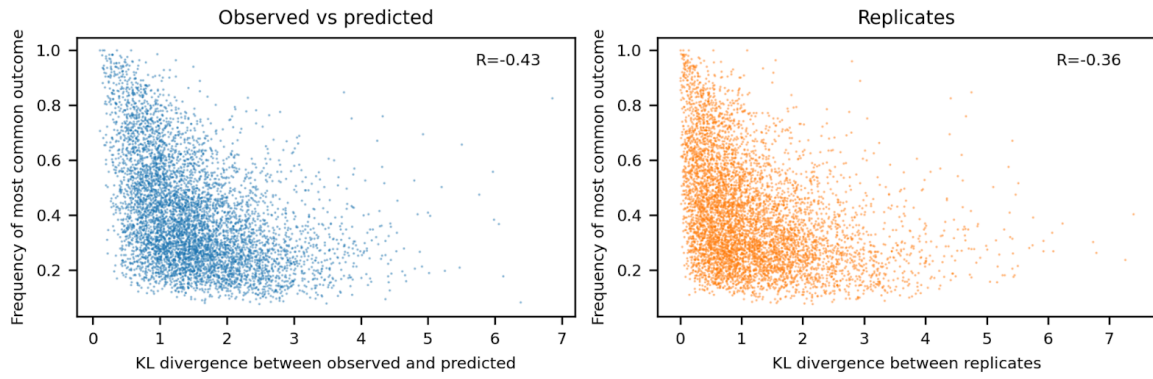


**Supplementary Figure 5.** Outcome distribution in terms of deletion size and microhomology. **(a)** Relative frequency of all outcomes or UMAP outcomes only. **(b)** Outcomes broken by UMAP cluster. Left column: relative frequency within cluster - each heatmap adds up to 100%. Middle column: absolute frequency distributed by cluster - same cell added up across clusters equals total UMAP frequency in (a). Right column: relative frequency distributed by cluster - same cell added up across clusters equals 100%. I+D = outcomes with insertions and deletions. Box at (0,0): pure insertions.

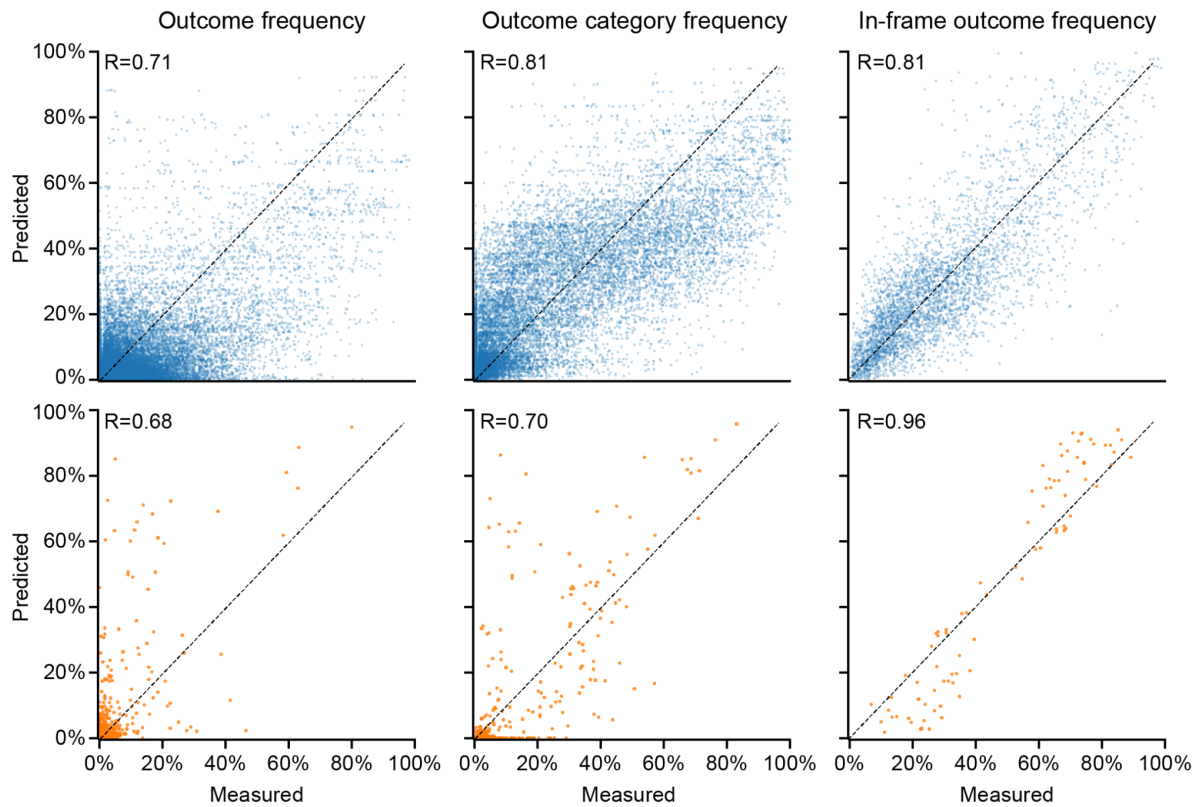




**Supplementary Figure 6.** Enrichment or depletion of outcomes categories within clusters across knockouts. Only outcome categories that made up >5% of the cluster depicted (a total of 7.2% outcomes not shown). D=deletion, I=insertion, MH=microhomology. Column next to cluster names indicates absolute average frequency of outcome category in a given cluster in control cells.



**Supplementary Figure 7.** Replicate concordance and prediction accuracy reflect frequency of most common outcome. KL divergence (x-axis) between observed and predicted outcome profiles (left panel, blue colour) or between replicates (right panel, orange colour) for one target (markers) contrasted against frequency of most common outcome for the target (y-axis). R = Pearson's R. Frequency of 1 = 100%.



**Supplementary Figure 8.** Measured (x-axis) and predicted (y-axis) frequencies of individual outcomes, outcome categories and in-frame outcomes (columns) in held-out targets in Lig4, Polq, Poll, Polm, Xrcc5 knockouts. Top: original screen (N=670 targets), bottom: endogenous validation screen (N=7 targets). R = Pearson's R.

## Supplementary References

1. Ahmad, A. *et al.* ERCC1-XPF endonuclease facilitates DNA double-strand break repair. *Mol. Cell. Biol.* **28**, 5082–5092 (2008).
2. Ismail, I. H. *et al.* The RNF138 E3 ligase displaces Ku to promote DNA end resection and regulate DNA repair pathway choice. *Nat. Cell Biol.* **17**, 1446–1457 (2015).
3. Williams, A. B. & Schumacher, B. p53 in the DNA-Damage-Repair Process. *Cold Spring Harb. Perspect. Med.* **6**, (2016).
4. Lu, H. & Davis, A. J. Human RecQ Helicases in DNA Double-Strand Break Repair. *Front Cell Dev Biol* **9**, 640755 (2021).
5. Lu, G. *et al.* Ligase I and ligase III mediate the DNA double-strand break ligation in alternative end-joining. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 1256–1260 (2016).
6. Wang, M. *et al.* PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res.* **34**, 6170–6182 (2006).
7. Truong, L. N. *et al.* Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7720–7725 (2013).
8. Wyatt, D. W. *et al.* Essential Roles for Polymerase  $\theta$ -Mediated End Joining in the Repair of Chromosome Breaks. *Mol. Cell* **63**, 662–673 (2016).
9. Chang, H. H. Y., Watanabe, G. & Lieber, M. R. Unifying the DNA End-processing Roles of the Artemis Nuclease: KU-DEPENDENT ARTEMIS RESECTION AT BLUNT DNA ENDS\*. *J. Biol. Chem.* **290**, 24036–24050 (2015).
10. Pryor, J. M. *et al.* Essential role for polymerase specialization in cellular nonhomologous end joining. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E4537–45 (2015).
11. Mazur, D. J. & Perrino, F. W. Identification and expression of the TREX1 and TREX2 cDNA sequences encoding mammalian 3'→5' exonucleases. *J. Biol. Chem.* **274**, 19655–19660 (1999).
12. Panier, S. & Boulton, S. J. Double-strand break repair: 53BP1 comes into focus. *Nat.*

*Rev. Mol. Cell Biol.* **15**, 7–18 (2014).

13. Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **18**, 495–506 (2017).
14. Goodarzi, A. A. *et al.* DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *EMBO J.* **25**, 3880–3889 (2006).