Transcriptional repression by a secondary DNA binding surface of DNA topoisomerase I safeguards against hypertranscription

Mei Sheng Lau^{*}, Zhenhua Hu, Xiaodan Zhao, Yaw Sing Tan, Jinyue Liu, Hua Huang, Clarisse Jingyi Yeo, Hwei Fen Leong, Oleg V. Grinchuk, Justin Kaixuan Chan, Jie Yan^{*}, Wee-Wei Tee^{*}

* Correspondence: M.S.L (<u>mslau@imcb.a-star.edu.sg</u>), J.Y (<u>phyyj@nus.edu.sg</u>), W-W.T (<u>wwtee@imcb.a-star.edu.sg</u>)

Supplementary information file contains: Supplementary Fig. 1-11

ubdomsing_HUMA HUMA HI KE GK DWGH-YE TARR KK CV DK-K-RATYYED TOPI_MOUSE_DNA_UGGSI KYII- MINPSS NIK GK KUMOH-YE TARR KK CV DK-K-RATYYED TOPI_SUNLA_DNA_UGGSI KYII- MINPSS NIK GK KUMOH-YE TARR KK CV DK-K-RATYYED TOPI_SUNLA_DNA_UGGSI KYII- MINPSS NIK GK KUMOH-YE TARR KK KK CV DK-K-RATYYED TOPI_SUNLA_DNA_UGGSI KYII- MINPSS NIK GK KUMOH-YE TARR KK KK CV DK-K-RATYYED TOPI_YACLADA MINPSK TYI- MINPSS NIK GK KUMOH-YE TARR KK KK CV DK-K-RATYYED TOPI_YACLADA KI KK KYYE RANNENSK KK GS DYK-F FK KK KK VV DE AK KK KY VYD TOPI_YACLADA KK KK KYRE RANNENSK KK GS DYK-F FK KK KK KK S TDS SG GK DYKEN TOPI_YACLADA KK KK KYRE RANNEN RANNEN FK KK KK KK S TDS SG GK DYKEN TOPI_YACLADA KK KK KYRE RANNEN RANNEN FK KK KK KK KK RANNEN RANNEN FK KK KK KK KK RANNEN SK KK KK KK RANNEN RANNEN FK KK KK KK RANNEN RANNEN FK KK KK KK KK RANNEN FK KK KK KK RANNEN RANNEN FK KK KK KK KK RANNEN RANNEN FK KK K						
Suddmains_nown QGS KY1 MK NP 55 N K GG KOWQ - EE TAREL K C D K M RUYNED TOP1_NOLE_DANOGS KY1 MK NP 55 N K GG KOWQ - EE TAREL K C D K M RUYNED TOP1_SRUE_DANVGQ KY1 MK NP 55 N K GG KOWQ - EE TAREL K C D K M RUYNED TOP1_SRUE_DANVGQ KY1 MK NP 55 N K GG KOWQ - EE TAREL K K C D K M RUYNED TOP1_SRUE_DANVGQ KY1 ML NP 55 N K GG KOPE N EE TAREL K K G G G G REPYTD TAREL K K G G G G REPYTD TOP1_VGCW_DANSKGR VY YG R RUM VONNAK REAL RAGE KEE EG G TADT VG C S R VEH N E Stata St	autodomosin 2 UUIMA					
Dorl, Joude, Dura, TOP J, Xohi, Zohi, M., Yung, Yun				GEKDWOK-YE		
TOPI_DROME_DNA_ TOPI_DROME_DNA_ TOPI_DROME_DNA_ TOPI_DROME_DNA_ TOPI_DROME_DNA_ TOPI_DROME_DNA_ TOPI_ACCW_DNA_ SIGNEY						
TOP1_CAEL_DAA, TOP1_CAEL_DAA, TOP1_CAEL_DAA, CARSTRV, TOP1_CAEL_DAA, SKGRROYFTC NKRVPSKK, KKKK, KKKK, SKGROYFTC KKKVFK, KKKKKKK, SGSOVKK, SKGROYFTC FKKVKKKK, KKKKKKK, SGSOVKK, SKGROYFTC FKKVKKKK, KKKKKKKK, SGSOVKK, SKGROYFTC FKKVKKKK, SKGROYFTC KKKVKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK	TOP1 DROME DNA					K- L RATYRDE
TOP1_FAST_DAA TOP1_VAST_DAA TOP1_VAST_DAA TOP1_VAST_DAA TOP1_VAST_DAA SIGROV_FYDU FK NEY K VV	TOP1 CAFEL DNA	VI GONKYL		GEKDEEK-YE	TARRIKKKIG	G- L REBYTDD
TOP_VACOUDA Consistency SK GRR Q YF KG MANY Q AR NAXK RDD F EY BY YN VMK RUNCE 08980 S0 0 N KDD F EY BY YN VMK RUNCE 08980 S0 0 N KDD F EY BY YN VMK RUNCE 08980 S0 0 N KDD F EY BY YN VMK RUNCE 08980 S0 0 N SDD F EY BY YN VMK RUNCE 08980 S0 0 N SDD F EY BY YN VMK RUNCE 08980 S0 0 N SDD F EY BY YN VMK RUNCE 08980 S0 0 SDD F EY BY YN VMK RUNCE 08980 S0 0 SDD F EY BY YN VMK RUNCE 08980 S0 0 SDD F EY BY YN VMK RUNCE 08980 S0 0 SDD F EY BY YN SDD F E E E E F YN F SD F E E E E F YN F SD F E E E E E F YN F SD F E E E E F YN F SD F E E E E F YN F S	TOP1 YEAST DNA	LENSEKYV		GOS DYKK- FE	KAROLKSYLD	A- L BRDYTRN
Consistency 5334207500 053345677 7877547098 6898865386 0 0447566 VIV V474 000 653345577 7877547098 694 000 514 subdomain3_HUMA MKS KERW VRQ RXVALYF DK LALRAGINE E EGET NDIVGC SUBVEHI NL TOP1_MOUSE_DNA_ WKS KERW VRQ RXVALYF DK LALRAGINE E EGET NDIVGC SUBVEHI NL TOP1_STALA WKS KERW VRQ RXVALYF DK LALRAGINE KE EGET NDIVGC SUBVEHI NL TOP1_ACAEL WKS KERW VRQ RXVALYF DK LALRAGINE KE EGET NDIVGC SUBVEHI NL TOP1_ACAEL WKS KERW RVQ RXVALYF DK LALRAGINE KE EGET NDIVGC SUBVEHI NL TOP1_ACAEL WKS KERW RVQ RXVALYF DK LALRAGINE KE EGET NDIVGC SUBVEHI NL TOP1_ACAEL MKS KERW RVQ RXVALYF DK LALRAGINE KE EGET NDIVGC SUBVEKIN EGEN NDIVGC SUBVEKIN EGEN NDIVGC SUBVEKIN EGEN NDIVGC EGEN NDIVGC EGEN NDIVGC EGEN ND	TOP1 VACCW DNA	SKGRROYFYG	KMHVONRNAK			KNI KKSSTDS
VVV V474 484 494 504 514 Subdomain3_HUMA WKSKEMKVRQ RAVALYF DK LALPARERE EGT NDTVGC CSERVEHINE TOPI_MOUSE_DNA WKSKEMKVRQ RAVALYF DK LALPARERE EGT NDTVGC CSERVEHINE TOPI_SCHA_DNA WKSKEMKVRQ RAVALYF DK LALPARERE EGT NDTVGC CSERVEHINE TOPI_SCHA_DNA WKSKEMKVRQ RAVALYF DK LALPARERE EGT NDTVGC CSERVEHINE TOPI_SCHA_DNA KKSKEMKVRQ RAVALYF DK LALPARERE EGT NDTVGC CSERVEHINE TOPI_ACCU FKSKEMRVRQ RAVALYF DK LALPARERE EGT NDTVGC CSERVEHINE TOPI_ACCU FKST FKST FKST FKST	Consistency	5 3 5 4 2 <mark>6 7 5 0 0</mark>	0 5 5 3 4 5 5 <mark>6 7 *</mark>	7 8 7 7 <mark>5 4 9 0</mark> 9 8	6 8 9 8 9 6 <mark>5 3</mark> 9 6	6 <mark>0 * 9 4 4</mark> 7 5 6 6
subdomain3_HUMA WKSKEMKVRQ TOPI_MOUSE_DNA_ WKSKEMKVRQ RAVALYFFDK LALRAGNERE EGETADIVGC SLRVEHINL TOPI_MOUSE_DNA_ WKSKEMKVRQ RAVALYFFDK LALRAGNERE EGETADIVGC SLRVEHINL TOPI_SCALADA WKSKEMRVRQ RAVALYFFDK LALRAGNERE EGETADIVGC SLRVEHINL TOPI_SCALADA WKSKEMRVRQ RAVALYFFDK LALRAGNERE EGETADIVGC SLRVEHINL TOPI_SCALADA WKSKEMRVRQ RAVALYFFDK LALRAGNERE EGETADIVGC SLRVEHINL TOPI_SCALADA WKSKEMRVRQ RAVALYFFDK LALRAGNERD VEGADIVGC SLRVEHINL TOPI_SCALADA FKSKEMRVRQ RAVALYFFDK LALRAGNERD VEGADIVGC SLRVEHINL TOPI_ACACU_DNA_ FKSKEMRVRQ RAVALYFFDK LALRAGNERD VEGADIVGC SLRVEHINL TOPI_ACCU_DNA_ FKSKEMRVRQ RAVALYFFDK LALRAGNERD VEGADIVGC SLRVEHINL TOPI_ACCU_DNA_ FKSKEMRVRQ RAVALYFFDK LALRAGNERD VEGADIVGC SLRVEHINL TOPI_YCAST_DNA_ FKSKEMRVRQ RAVALYFFDK LALRAGNERD VEGADIVGC SLRVEHINL TOPI_YCAST_DNA_ FKSKEMRVRQ RAVALYFFDK LALRAGNERD VEGADIVGC SLRVEHINL TOPI_YCAST_DNA_ FKSKEMRVRQ RAVALYFFDK LALRAGNERD VEGADIVGC SLRVEHINL TOPI_MOUSE_DNA_ HOF_ELD			494	494	504	E14
TOP1_MOUSE_DNA_ TOP1_KENLA_DNA_ WESKEBMYURO RAVALYF DK LALBAGINE E EGC TADTVCC CSURVEHIN TOP1_RENLA_DNA_ WESKEBMYURO TOP1_RENLA_DNA_ TOP1_RENLA_DNA_ TOP1_RENLA_DNA_ TOP1_RENLA_DNA_ TOP1_RENLA_DNA_ TOP1_RENLA_DNA_ LKSKVMURERO RAVALYF DK LALBAGINE D E-OODTVCC CSURVEHINL CSURVEHINL TOP1_CAEL_DNA_ TOP1_CAEL_DNA_ TOP1_CAEL_DNA_ TOP1_CAEL_DNA_ TOP1_STATUS CSURVEHINL LKENVERVC CSURVEHINL CSURVEHINL TOP1_VACCW_DNA_ TOP1_CAEL_DNA_ TOP1_CAEL_DNA_ TOP1_STATUS CSURVEHINL LKENVERVC CSURVEHINL CSURVEHINL TOP1_VACCW_DNA_ TOP1_STATUS CSURVEHINL TOP1_VACCW_DNA_ TOP1_STATUS CSURVEHINL CSURVEHINL CSURVEHINL TOP1_MOUSE_DNA_ TOP1_STATUS CSURVEHINL CSURVEHINL TOP1_VACCW_DNA_ SUBDOMAIN3_HUMA HF- ELDD	subdomain3 HUMA				FGETADTVGC	
TOP1_ERNL_DNA_ WK SKEMK VRO RAVALYF DK LALE AR NE E E. G. TADT VGC C SLE VENYL TOP1_DROME_DNA WK SKEMK VRO RAVALYF DK LALE AR NE K D VDL AADT VGC C SLE VENYL TOP1_CAEL_DNA_ FK SKEMK VRO RAVALYF DK LALE AR NE K D VDL AADT VGC C SLE VENYL TOP1_CAEL_DNA_ FK SKEMK VRO RAVALYF DK LALE AR NE K D VDL AADT VGC C SLE VENYL TOP1_ACCUMAL AF 87 70 57 78 7 86 88 98 75 5 77 7 67 5 5 38 37 8 4 7 7 7 88 68 98 95 subdomain3_HUMA HP - ELD - GOE YV VEF D L GKDSTRY YNK WP VEKRY FK NL QLF ME TOP1_MOUSD_AA, HP - ELD - GOE YV VEF D P GKDSTRY YNK WP VEKRY FK NL QLF ME TOP1_NOUSD_AA, HP - ELD - GOE YV VEF D P GKDSTRY YNK WP VEKRY FK NL QLF ME TOP1_PLOROME_DNA, HP - ELD - GOE YV VEF D P GKDSTRY YNK WP VEKRY FK NL QLF ME TOP1_PLOROME_DNA, HK - ELD - GOE YV VEF D P GKDSTRY YNK WP VEKRY FK NL QLF ME TOP1_SOLAL, AK & ELD - GOE YV VEF D P GKDSTRY <td>TOP1 MOUSE DNA</td> <td>WKSKEMKVRO</td> <td>RAVALYFIDK</td> <td>LALRAGNEKE</td> <td>EGETADTVGC</td> <td>CSLRVEHI NL</td>	TOP1 MOUSE DNA	WKSKEMKVRO	RAVALYFIDK	LALRAGNEKE	EGETADTVGC	CSLRVEHI NL
TOP1-DROME_DAA_ MKS KEMM VRQ RAVALYF DK LALRAENER D VD RAADTVGC CSLE VEHVL TOP1-CAEL_DNA_ LKS KEMM VRQ RATALYF DK LALRAENE D VD RAADTVGC CSLE VEHVL TOP1-YEAST_DNA_ LKS KEMM VRQ RATALYF DK LALRAENE D VD RAADTVGC CSLE VEHVL TOP1-YEAST_DNA_ LKS KEMM VRQ RATALYF DK LALRAENE D VD RAADTVGC CSLE VEHVL TOP1-YEAST_DNA_ LKS KEMM VRQ RATALYF DK LALRAENE D CSLE VEHVL LKENKHIE Consistency 47 87 87 95 78 78 68 89 87 77 77 67 75 S 38 37 88 77 77 Y <td< td=""><td>TOP1 XENLA DNA</td><td>WKSKEMKVRQ</td><td>RAVALYFIDK</td><td></td><td>EGETADTVGC</td><td>CSLRVEHI NL</td></td<>	TOP1 XENLA DNA	WKSKEMKVRQ	RAVALYFIDK		EGETADTVGC	CSLRVEHI NL
TOP1_CAEEL_DNA_ FK SK KE MA VR O RAT FLY FP DK LALK AG NE FD VDC AADT VGC CSL FV EH KL TOP1_YEAST_DNA_ LK SK VML ERO KAVAH YL DV FALB AG GS EDC ADT VGC CSL FV EH KL TOP1_YACCW_DNA_	TOP1_DROME_DNA_	WKSKEMRVRQ	RAVALYFIDK	LALRAGNEKD	E - DQADTVGC	CSLRVEHVQL
TOP1_YEAST_DNA_ L KS KV/MLERO KAVAI YL DV FALRAGGE SS EDE-ADTYGC CS RYERVT TOP1_VACCW_DNA_ NYQLAYFM LMETMFFI NYGLAYFM LKEN_FTYGL LKEN_FTYGL LK KN KHIE Consistency #78.795778 S8878 77 78.95898 75.5777 YK 547 Y subdomain3_HUMA HF - ELD- - GGE YVVEF D L GKDSI RY YNK PVEKRV FNLQLFME- TOP1_KNLA_DNA_ FO ELD- - GGE FVVEF D P GKDSI RY YNK PVEKRV FNLQLFME- TOP1_SENLA_DNA_ FO ELD- - GGE FVVEF D P GKDSI RY YNK P VEKRV FNLQLFME- TOP1_SENLA_DNA_ FO ELD- - GGE FVVEF D P GKDSI RY YNK P VEKRV FNLQLFME- TOP1_SENLA_DNA_ FO FO AKKINEDD KK KEFVVEF D L GKDSI RF YOE BVDKQV FNLQLFME- TOP1_YEAST_DNA_ F SE FNTV F D P GKDSI RF YOE BVDKQV FNL ELF ME TOP1_YACCW_DNA_ F SE FNTV F D P GKDSI RF YOE BVDKQV FNN KI F FME TOP1_YACCW_DNA_ FNTOFI FNTVF	TOP1_CAEEL_DNA_	F K S K E MR V R Q	RATALYFIDK	LALRAGNEK D	VDE AADT VGC	CSLRVEHI KL
TOP1_VACCW_DNA_ - NYOL AYF IN ET MFF - - IF & KMEY LKE N. ET VGL L TL KINKHI E Consistency 4787795778 7868889 75 577 767 538378177 78898959	TOP1_YEAST_DNA_	<mark>L</mark> K S K V <mark>ML</mark> E R Q	KAVAI YL <mark>I</mark> D <mark>V</mark>	FALRAGGEKS	EDE-ADTVGC	CSLRYEHVTL
Consistency 4787795778 7868898875 577 7 67 538 78 7 78<988999 subdomain3_HUMA HP - ELD	TOP1_VACCW_DNA_	NYQL <mark>A</mark> VFM	L ME T MF F I	<mark>R F G</mark> K M <mark>K</mark> Y	LKEN-ETVGL	LT <mark>LKN</mark> KHI <mark>E</mark> I
subdomain3_HUMA HP. ELD	Consistency	<mark>4</mark> 7 8 7 7 <mark>9 5</mark> 7 7 8	7 8 6 8 8 9 8 * 7 <mark>5</mark>	<mark>5</mark> 77 <mark>*</mark> 7 <mark>*</mark> 67 <mark>*</mark> 5	<mark>5 3 8 3</mark> 7 8 * * * 7	7 8 <mark>* 9 </mark> 5 8 * 9 <mark>5</mark> 9
bubdomain3_HUMA HP - ELD GQE YVEF D L BKDSI RY YNK M PVEKRY FKNL QLFME- TOP1_MOUSE_DNA_ HP - ELD GQE YVEF D P BKDSI RY YNK M PVEKRY FKNL QLFME- TOP1_SNUALDNA_ FQ - ELD GQE YVEF D P BKDSI RY YNK M PVEKRY FKNL QLFME- TOP1_SNUALDNA_ FQ - ELD GGE FVVEF D P BKDSI RY YNK M PVEKRY FKNL QLFME- TOP1_SNUALDNA_ FQ - ELD GGE FVVEF D P BKDSI RY YNK M PVEKRY FKNL QLFME- TOP1_SNUALDNA_ FQ - ELD GKE NVVF D P BKDSI RY YNK M PVEKRY FKNL QLFME- TOP1_SNUALDNA_ FQ - GKE NVVF D P BKOSI RY YNK M PVEKRY FKNL LFKKK TOP1_YACKUDNA_ FY DELVIK - V GKOSI RY YNK M PVEKRY FKNL LFKKK Consistency 4500343000 0034848957 7 4 * 18987 665 477888 877487760					VV	•
Subdomain3_HUMA_NEY_ELD	autodomosina. LILIAAA				· · · · · · · · · · · · · · · · · · ·	
TOPI, MOUSE, DNA, PRIVE ELD						
TOPI_DROME_DNA_ HK-ELN					Y NKWPVEKRV	
TOP1_CAEL_DNA_ TOP1_VEAST_DNA_ TOP1_VEAST_DNA_ Consistency F DSAKL NEDD KKKEK F F VV FF DFL GK DS IR F Y OVE V DK QV Y KN K I F ME- F K V H K S N T I F KR F TOP1_VACCW_DNA_ SP	TOP1_DROME_DNA				YNEVEKRV	EKNLELEME-
TOP1_YEAST_DNA_ TOP1_VACCW_DNA_ SP	TOP1 CAFEL DNA	E DSAKL NEDD	KKEKEEVVEE		ENRVSVEKRV	YKNIKIEME-
TOP_VACCV_DNA_ Consistency SP DEI VI K V 0KD KVSH 4500343000 EFV9HKSNRL 4500343000 VKP - LKLTD 665 477888 VKP - LKLTD 665 477888 Subdomain3_HUMA NCQPE DDLFD RLNTGI LNKH LQDL MEGLTA KVFRTYNASI VFRTYNASI TLQQLKE TT QQDLKE TCP1_KNL_DNA KQPE DDLFD TLQQLKE TLQQLKE T QQDLKE TCP1_RNA_NA KQPE DDLFD TLQQLKE T QQDLKE TCP1_RAMANA KQPE DDLFD TLQQLKE T QQDLKE T QQDLKE TCP1_RAMANA KQPE DDLFD TLQQLKE T QQDLKE T T QQDLKE T QQDLKE T T QQDLKE T QQDLKE T T QQDLKE T T QQDLKE T T QQDLKE T T T QQDLKE T T T T T T T T T T T T T T T T T T T	TOP1 YEAST DNA	KP		DFL GK DSI RF	YOEVEVDKOV	
Consistency 4 5 00 3 4 3 0 0 0 0 3 4 8 4 8 9 5 7 7 4 4 1 8 9 8 7 6 6 5 4 77888 8 1 7 4 8 7 7 6 0 subdomain3_HUMA NK OPE DD LFD RENTGI LNKH LQQ LKE LT KVFRTVNASI TLQQQLKE LT TOP1_MOUSE_DNA, NK OPE DD LFD RENTGI LNKH LQQ LME GLTA KVFRTVNASI TLQQQLKE LT TOP1_MOUSE_DNA, NK OPE DD LFD RENTGI LNKH LQQ LME GLTA KVFRTVNASI TLQQQLKE LT TOP1_XENLA_DNA, NK OPE DD LFD RENTGI LNKH LQQ LME GLTA KVFRTVNASI TLQQQLKE LT TOP1_XENLA_DNA, NK OPE DD LFD RENTGV NEH LKE LME GLTA KVFRTVNASI TLQQ QL LT TOP1_CAEEL_DNA, GKAPS DD LFD RLDTATLNDH LRS LMD GLTV KVFRTVNASI TLQE QL LK T TOP1_YEAST_DNA, GKAPS DD LFD RLDTATLNDH LRS LMD GLTV KVFRTVNASI TLQE QL LK T TOP1_YEAST_DNA, A 8 85 7 6 1 8 9 6 6 4 5 8 7 6 6 9 75 6 7 4 9 8 7 178 7 8 8 8 8 4 7 7 4 7 7 3 3 9 6 T. T. T. Subdomain3_HUMA AP DE NIP AKI LS YNRANRAV AL E CNHO T. T. <td< td=""><td>TOP1 VACCW DNA</td><td><mark>S P</mark></td><td> DEI VI K-</td><td>- F<mark>V</mark>GKDKVSH</td><td>E F V V H K S N R L</td><td>YKPL-LKLTD</td></td<>	TOP1 VACCW DNA	<mark>S P</mark>	DEI VI K-	- F <mark>V</mark> GKDKVSH	E F V V H K S N R L	YKPL-LKLTD
	Consistency	4 <mark>5 0 0 3 4 3</mark> 0 0 0	0034 <mark>8489</mark> 57	7 * <mark>4</mark> * * * 8 9 8 7	6 6 <mark>5 *</mark> 4 7 7 8 8 8	8 * 7 * <mark>4</mark> 8 7 7 6 <mark>0</mark>
SubdomainS_HUMA_NKQPEDLEFD REINTGIENKH LQDLMEGETA KVFRTYNASI TLQQQLKELT TOP1_XENLA_DNA_ KKQPEDLEFD REINTGVINEH LKELMEGETA KVFRTYNASI TLQQQLEET TOP1_XENLA_DNA_ KKQPEDLEFD REINTGVINEH LKELMEGETA KVFRTYNASI TLQQQLEET TOP1_DROME_DNA_ HKEGDLEFD REINTGVINEH LKELMEGETA KVFRTYNASI TLQQQLEET TOP1_DROME_DNA_ HKEGDLEFD REINTGVINEH LKELMEGETA KVFRTYNASI TLQQQLEET TOP1_YEAST_DNA_ PKQPGHQLEFD REIDTSTINKY LQNYMPGETA KVFRTYNASIK TLQQQLEKT TOP1_YEAST_DNA_ PKQPGHQLEFD REIDTSTINKY LQNYMPGETA KVFRTYNASIK TMQDQLDET P TOP1_VACCW_DNA_ DS SPEEFLEFN KESERKVYEC IKQF-GIRI KDLRIYGVNY TFLY Consistency 486857618 9166458766 975674987 78**8884 1774773396 616						
TOPI_SNIA_DNA_NKOPDOLED RENTSULNKH LOOL MADELAA KVFRIYNASL TLOOOLDLT TOPI_SNIA_DNA_KKEGODLFD RENTSULNKH LOOL KKEGUTA KVFRIYNASL TLOOOLDLT TOPI_CAEELDNA_GKAPSDDLFD REDTSILNKK LOON WPGLTA KVFRIYNASK TLOOOLDLT TOPI_CAEELDNA_GKAPSDDLFD REDTSILNKY LOON WPGLTA KVFRIYNASK TLOOOLDLIP TOPI_VACCW_DNA_DSSPEEFLFN KLSERKVYEC KOF GRI KOLRTYGVNY TFF					KVERTYNASI	
Indit JALDE_DNA_HK & GODLE D REWTQVLNEH LGOLMA GLANA KI LOSOLUTU TOP1_DROME_DNA_HK & GODLE D REWTQVLNEH LKEL MEGLTA KVFRTYNASK TLOSOLUT TOP1_CAEEL_DNA_GKAPSDDLFD REDTATLNDH LRSLMDGLTV KVFRTYNASK TLOSOLUT Consistency 486857618 9166458766 9756741987 78***8884 774773396 Subdomain3_HUMA APDE NIPARI LSYNRANRAV ALCNHO 78***8884 774773396 subdomain3_HUMA APDE NIPARI LSYNRANRAV ALCNHO 78***8884 774773396 Subdowain3_HUMA APDE NIPARI LSYNRANRAV ALCNHO 78***8884 7747773396 TOP1_SUNA_DNA NSDDNVPARI LSYNRANRAV ALCNHO 78***8884 7747773396 TOP1_YEAST_DNA_ NFKDNVAA					KVERTYNASI	
TOP_CAEL_DNA_ GKAPS DD_E D RLDTATLNDH LRSLMDGLTV KVFRTYNASL TLOEQLIK T TOP_CAEL_DNA_ PKOPGHOLFD RLDPSILNKY LQNYMPGLTA KVFRTYNASL TLOEQLIK T TOP_YACCW_DNA_ DSSPEEFLFN KLSERKVYEC IKQF.GIRI KVFRTYNASK TMQOQLDL P TOP_VACCW_DNA_ DSSPEEFLFN KLSERKVYEC IKQF.GIRI KVFRTYNASK TMQOQLDL P TOP_VACCW_DNA_ DSSPEEFLFN KLSERKVYEC IKQF.GIRI KVFRTYNASK TFLY Consistency 4868576*8 9*66458766 975674 987 78**8884 77747733966 	TOP1 DROME DNA				KVERTYNASK	
TOP1_YEAST_DNA_ TOP1_YEAST_DNA_ DS S PEFF FF N KLSERRVYEC Consistency 4868576 * 8 9*66458766 975674 987 78** 8884 TMODOLDL P TFLY 7747733 96 TOP1_VOLSE_DNA_ TOP1_MOUSE_DNA_ DP DE NVP AKI LS YNRANRAY ALL CNHO TOP1_XENLA_DNA_ DS DDNVP AKI LS YNRANRAY ALL CNHO TOP1_XENLA_DNA_ NS DDNVP AKI LS YNRANRAY ALL CNHO TOP1_YEAST_DNA_ N* KGS YAE KI LKYNAANRTY ALL CNHO TOP1_YACCW_DNA_ N* KGS YAE KI LKYNAANRTY ALL CNHO TOP1_YACCW_DNA_ TOP1_Y	TOP1 CAFEL DNA	GKAPSDDLED	REDTATINDH		KVERTYNASI	
TOP_VACCU_DNA_ Consistency DS SPEEFLERN KLSERKVYEC IKQFSIRI KDLRTYGVNY IFLY Subdomain3_HUMA APDE NIP AKI LS YNRANRAY AIL CONHO 7747733966 subdomain3_HUMA APDE NIP AKI LS YNRANRAY AIL CONHO TOP1_NOUSE_DNA APDE NIP AKI LS YNRANRAY AIL CONHO TOP1_NOUSE_DNA APDE NIP AKI LS YNRANRAY AIL CONHO TOP1_NOUSE_DNA APDE NIP PAKI LS YNRANRAY AIL CONHO TOP1_ROME_DNA DP SATUP EKL LA YNRANRAY AIL CONHO TOP1_CAEL_DNA N K GSVAERI LS YNRANRAY AIL CONHO TOP1_VACCW_DNA N K GSVAERI L K YNRANRAY AIL CONHO TOP1_VACCW_DNA N K GSVAERI L K YNRANRITY AIL CONHO TOP1_VACCW_DNA S3 44 7 9 44 7 9 777 6 8 7 9 5 8 8 8 7 7 7 Gonsistency S3 44 7 9 44 7 9 777 6 8 7 9 5 8 8 8 7 7 7 M	TOP1 YEAST DNA	PKOPGHOLFD	REDPSIENKY		KVERTYNASK	
Consistency 4868576 8966458766 975674 987 78** 8884 7747733 96 subdomain3_HUMA APDE NU PANI L S Y NRANRAY AI L C NHO APDE NU PANI L S Y NRANRAY AI L C NHO TOP1_NOUSE_DNA, APDE NU PANI L S Y NRANRAY AI L C NHO APDE NU PANI L S Y NRANRAY AI L C NHO TOP1_XENLA_DNA, NS DD NV PANI L S Y NRANRAY AI L C NHO C NHO C NHO TOP1_CAEEL_DNA, NS DD NV PANI L S Y NRANRAY AI L C NHO C NHO C NHO TOP1_YEAST_DNA, NK GS VAAE (I L AY NRANRAY AI L C NHO NK MT NV KS I S PL PS PKKL I AL TI KO Consistency S 34 4 7 9 4 4 7 9 7 77 7 8 7 9 5 8 8 8 7 7 7 M K M K MV A- KS I S PL PS PKKL I AL TI KO I C NHO I C NHO I C NHO Consistency S 34 4 7 9 4 4 7 9 7 77 7 8 7 9 5 8 8 8 7 7 7 I C NHO I C NHO MU Consistency S 34 4 7 9 4 4 7 9 7 7 7 6 8 7 9 5 8 8 8 7 7 7 I C NHO I C NHO	TOP1_VACCW_DNA_	DSSPEEFLFN	KLSERK VYEC	IKQFGIRI	K DL R T Y GV NY	Τ F
subdomain3_HUMA APDE NIP ARI LSYNRANRAY AILCNHO TOP1_MOUSE_DNA APDE NIP ARI LSYNRANRAY AILCNHO TOP1_XRNLA_DNA NSDDNY PARI LSYNRANRAY AILCNHO TOP1_DROME_DNA DSATVPERL LAYNRANRAY AILCNHO TOP1_CAEELDNA NFKDNYAARI LSYNRANRAY AILCNHO TOP1_YEAST_DNA N KGSWAEKI LKYNAANRTY AILCNHO TOP1_YEAST_DNA NF WT NY KS ISPLPSPKKL IALTIKC Consistency 534470/44 7 9777687958 888 7777	Consistency	4 868 5 76**8	0*66459766		* 7 0 * * * 0 0 0 4	* 7 7 4 7 7 2 2 0 6
subdomain3_HUMA APDE NI PARI L SYNRANRAY AILCNHO TOP1_MOUSE_DNA APDE NI PARI L SYNRANRAY AILCNHO TOP1_XRNLA_DNA NSDDNVPARI L SYNRANRAY AILCNHO TOP1_ZRNLA_DNA NSDDNVPARI L SYNRANRAY AILCNHO TOP1_CAEEL_DNA NPKDNVAAKI L SYNRANRAY AILCNHO TOP1_YEAST_DNA N KGSVAAKI L SYNRANRAY AILCNHO TOP1_YACCW_DNA NF WT NV KS I SPLPSPKKL I ALTIKO Consistency 5344794477 9777687958 8887777			9 00438700	<mark>975674</mark> *987	787788884	//4//5590
TOPI_MOUSE_DNA_APDENVPARI LSYNRANRAY ALLCNHO TOPI_XENLA_DNA_NSDDNVPARI LSYNRANRAY ALLCNHO TOPI_DROME_DNA_DPSATVPEKL LAYNRANRAY ALLCNHO TOPI_CAEL_DNA_NFKDNVARI LSYNRANRAY ALLCNHO TOPI_YEAST_DNA_N-KGSVAEKI LKYNAANRTY ALLCNHO TOPI_VACCV_DNA_NFWTNV-KS ISPLPSPKKL IALTIKO Consistency 53447944 7 9777687958 886 777		616	5 004 <u>5 8 7 0 0</u> 626	97 <mark>567<mark>4</mark>*987</mark>	7 0 0 0 0 0 0 4	// <mark>4//559</mark> 0
TOPI ZENLA DNA NS DDNV PARI LSYNRANRAV ALL CNHO TOPI DROME DNA DP S ATV PE RL LAYNRANRAV ALL CNHO TOPI CAELDNA NF KONVAARI LSYNRANRAV ALL CNHO TOPI YEAST DNA N- KGS VAE RL LKYNAANRTV ALL CNHO TOPI YACCW DNA NF WT NV - KS ISPLP SPKKL IALTIKO Consistency 5344794477 9777687958 887777	subdomain3 HUMA			975674*987	<mark>7 8 8 8 8 8</mark> 8 8 8 8	// 4// 3320
TOPI_DROME_DNA_DPSATUPE KL LAYNRANRAY AILCNHO TOPI_CAEELDNA_NPKDNWAAKI LSYNRANROV AILCNHO TOPIYEAST_DNA_N-KGSVAEKI LKYNAANRTY AILCNHO TOPIYEAST_DNA_N-KGSVAEKI LKYNAANRTY AILCNHO Consistency S3447944 7 977687958 88 777 d Tu	subdomain3_HUMA TOP1 MOUSE DNA			975674 987	/ 0 0 0 0 0 4	// 4// 5590
TOP1_CAEEL_DNA_ NPKDNVAAKI LSYNRANRQV AILCNHO TOP1_YEAST_DNA_ N-KGSVAEKI LKYNAANRTV AILCNHO TOP1_VACCW_DNA_ NFWTNVKS ISPLPSPKKL IAITIKO Consistency 534479447 7 9777687958 88 777	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA			975674987	/ 0 0 0 0 0 4	//4//3390
TOP1_YEAST_DNA_ TOP1_VACCW_DNA_ Consistency 53 44 79 44 77 97776 879 58 88 7777	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_		L SYNRANRAV L SYNRANRAV L SYNRANRAV L SYNRANRAV L AYNRANRAV	975674987 AILCNHQ AILCNHQ AILCNHQ AILCNHQ	/ 0 0 0 0 0 4	// 4 //3390
TOPL_VACCW_DNA_NFWTNV KS ISPLPSPKKL IALTIKO Consistency 53447944 7 9777687958 88 777	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_ TOP1_CAEEL_DNA_		626 LSYNRANRAV LSYNRANRAV LSYNRANRAV LAYNRANRAV LSYNRANRAV LSYNRANRAV	975674 987 AILCNHQ AILCNHQ AILCNHQ AILCNHQ AILCNHQ	/ 0 0004	// 4//3320
Consistency 53447944 7 9777687958 88 777	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_ TOP1_CAEEL_DNA_ TOP1_YEAST_DNA_	616 AP DE NI P AKI AP DE NVP AKI NS DDNVP AKI DP S AT VP E KL NPK DNV AAKI N- KGS VAE KI	626 C S Y NR ANR AV C S Y NR ANR AV C S Y NR ANR AV C A Y NR ANR AV C S Y NR ANR AV	975674 987 AILCNHQ AILCNHQ AILCNHQ AILCNHQ AILCNHQ AILCNHQ	7.0 0004	// •// 33 2 0
d T R	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_ TOP1_CAEEL_DNA_ TOP1_YEAST_DNA_ TOP1_VACCW_DNA_	AP DE NI PAKI AP DE NV PAKI NS DDNV PAKI DP SATV PEKL NPK DNV AAKI N- KGS VAEKI NF WT NV KS	626 C S Y NR ANR AV L S Y NR ANR AV L S Y NR ANR AV L A Y NR ANR AV L S Y NR ANR AV L S Y NR ANR TV I S P L P S P K K L	975674 987 AILCNHQ AILCNHQ AILCNHQ AILCNHQ AILCNHQ AILCNHQ IALTIKQ	7.0 0004	// 4// 3320
d T R	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_ TOP1_CAEEL_DNA_ TOP1_YEAST_DNA_ TOP1_YACCW_DNA_ Consistency	AP DE NI P AKI AP DE NV P AKI NS DDNV P AKI DP S AT VP E KL NP K DNV AAKI N- KGS V AE KI NF WT NV KS S 344 7 944 7	626 C S Y NR ANR AV L S Y NR ANR AV L S Y NR ANR AV L AY NR ANR AV L S Y NR ANR AV L S Y NR ANR AV L S Y NR ANR TV I S PL P S P K KL 97776 879 58	975674 987 AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC I ALTIKC 881 777	// 00004	// <u>//</u> /
	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_ TOP1_CAEEL_DNA_ TOP1_YEAST_DNA_ TOP1_YEAST_DNA_ Consistency		626 C S Y NR ANR AV L S Y NR ANR AV L S Y NR ANR AV L AY NR ANR AV L S Y NR ANR AV L S Y NR ANR AV L S Y NR ANR AV S Y NR ANR TV S PL P S P K KL 97776 879 58	975674 987 AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC B887777	/0 0004	// / 33 20
	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_ TOP1_CREL_DNA_ TOP1_YEAST_DNA_ TOP1_VACCW_DNA_ Consistency	APDE NI P AKI APDE NV P AKI NS DDNVP AKI DP S AT VP E KL NF KDNV AAKI N- KGS VAE KI NF WT NV KS S 3 4 4 7 9 4 4 7		975674 987 AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC B8777	/0 0004	
	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_XENLA_DNA_ TOP1_CAEL_DNA_ TOP1_CAEL_DNA_ TOP1_YEAST_DNA_ Consistency		C	975674 987 AILCNHO AILCNHO AILCNHO AILCNHO AILCNHO AILCNHO I ALTI KO 881777	/0 0004	d
	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_ TOP1_DROME_DNA_ TOP1_VAST_DNA_ TOP1_VAST_DNA_ Consistency		C	975674 987 AILCNHO AILCNHO AILCNHO AILCNHO AILCNHO I ALTIKO 88 777	/0 0004	d
	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_ TOP1_CAEEL_DNA_ TOP1_YAET_DNA_ TOP1_YACCW_DNA_ Consistency	AP DE NI P AKI AP DE NV P AKI NS DD VV P AKI DP S AT V P E KI NF KD VV AAKI N KG S V AE KI N F WT NV KS 5 3 4 4 7 9 4 4 7	C C	975674 987 AILCNHO	/0 0004	d
	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_DROME_DNA_ TOP1_CHEL_DNA_ TOP1_YEAST_DNA_ TOP1_VACCW_DNA_ Consistency		C	975674 987 AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC B8 777	/0 0004	d
	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_XENLA_DNA_ TOP1_CATEL_DNA_ TOP1_YEAST_DNA_ TOP1_YEAST_DNA_ Consistency		C	975674 987 AILCNHO AILCNHO AILCNHO AILCNHO AILCNHO AILCNHO I ALTIKO 887777	/0 0004	d
	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_XENLA_DNA_ TOP1_OROME_DNA_ TOP1_CATEL_DNA_ TOP1_VEAST_DNA_ Consistency		C C	975674 987 AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC AILCNHC I ALTIKC 88 777	/0 0004	d T
	subdomain3_HUMA TOP1_MOUSE_DNA_ TOP1_XENLA_DNA_ TOP1_ZENLA_DNA_ TOP1_DROME_DNA_ TOP1_YEAST_DNA_ TOP1_VACCW_DNA_ Consistency		C C C	975674 987 AI L CNHO AI L CNHO		d T

start of core subdomain 3

TOP1	∆G(kcal/mol)	
R546Q	-3.18 ± 1.97	
Wildtype	-8.99 ± 2.43	

TOP1(4K-A)

b

TOP1(R546Q)

Supplementary Fig. 1 DNA-interacting residues on secondary DNA binding surface are conserved across type1B topoisomerase across species and are predicted to disrupt DNA binding when mutated. (a) Protein sequence alignment of core subdomain III of human TOP1 (amino acids 434-633) with homologous regions of TOP1 from human, mouse, *Xenopus laevis, Drosophila melanogaster, Caenorhabditis elegans, Saccharomyces cerevisiae, and Vaccinia* virus (strain Western Reserve). Amino acid residues interacting with secondary DNA are marked with black arrow heads. Amino acid numbering is for human TOP1. (b-c) Superimposition of final MD trajectory structures of TOP1(4K-A)cc (b) and TOP1(R546Q)cc (c) (TOP1 in green, DNA in orange) in complex with DNA at secondary binding site from four replicate runs (yellow, magenta, cyan and white). (d) Computed binding free energies for the interaction of TOP1cc with secondary DNA. Values are averages ± SD obtained from the runs (two for R546Q, four for wildtype) in which secondary DNA remained stably associated with TOP1.







Figure S2 Overexpression of TOP1(WT), TOP1(R546Q) and TOP1(4K-A) in 293T cells.

N.A.

(a) Western blot showing protein level of exogenous HA-tagged TOP1 in lysates from equal number of sorted cells. The experiment is repeated at least twice with similar results. (b) Immunofluorescence images of TOP1(WT)-overexpressing cells stained for γ H2A.X. GFP signal marks positively transfected cells. '+TPT' are cells treated with 100 nM topotecan (TPT) for 24 hours and is used as positive control for punctate γ H2A.X signal that characterize DNA damage foci in genome. Scale bars: 10 μ m. (c) Western blot showing total TOP1 protein levels and γ H2A.X levels in lysates from cells transfected with indicated overexpression constructs. '+TPT' is lysate from cells treated with TPT as indicated in (b). Panels (b) and (c) are from independent experiments using orthogonal assays for determining DNA damage in cells with TOP1(WT) overexpression. (d) Live cell cycle analyses of TOP1(WT)-overexpressing cells and control cells.

+ TPT



Supplementary Fig. 3 Downregulation of RNA by TOP1 is dependent on amino acid residue on secondary DNA binding site. RNA content of 293T cells overexpressing TOP1 variant proteins with mutation on secondary DNA binding surface (a), or mutation of catalytic residue (b), in comparison with TOP1(WT) and vector control. Bar heights are median and error bars are interquartile range. n, number of biological replicates, each containing 100,000 sorted cells. Mann-Whitney test was performed; *p*-values (two-tailed) are indicated above the comparisons. Source data are provided as a Source Data file.



Supplementary Fig. 4 Increased transcription in mutant mESCs. Metaplots of elongating RNAP2 ('S2P') ChIP-seq signal over all genes in Mut.2 and WT mESCs.



Supplementary Fig. 5 Mutant and WT TOP1 occupancy on chromatin.

(**a-b**) Metaplot of TOP1 ChIP-seq signals over all genes (**a**), or genes classified according to gene expression levels (**b**). (**c**) Metaplots of TOP1 and elongating RNAP2 ('S2P) ChIP-seq signal of over genes classified according to expression levels. Note the difference in y-axis for 'High' genes compared to the other categories. (**d**) Western blots of TOP1 co-immunoprecipitation from WT and Mut.1 mESC nuclear lysates. AFF4 is a component of the P-TEFb and Super Elongation Complex (SEC) complexes. IP: immunoprecipitated; IB: immunoblot. The experiment is repeated twice with similar results. (**e**) Distribution of TOP1 ChIP-seq peaks in various annotated regions of the genomes.

6



Supplementary Fig. 6 Reduced negative supercoiling in genomes of mutant mESCs.

Representative immunofluorescence images of WT and mutant mESCs with psoralen staining. Bleomycintreated cells serves as negative control for psoralen staining. Scale bars: 20 µm. Quantification of signal in the images are presented in Fig. 3e. The experiment is repeated twice with similar results.



Supplementary Fig. 7 Plasmid relaxation activities of TOP1(WT) or TOP1(R546Q) recombinant proteins in varying salt conditions. n=4 independent experiments, mean \pm SD. Dotted lines are restricted cubic spline curves generated by calculating 52 points with X values ranging from 0 to 400. Source data are provided as a Source Data file.

Kidney Chromophobe



Head and Neck Squamous Cell Carcinoma (HPV-)



Pancreatic Adenocarcinoma







Stomach Adenocarcinoma



Supplementary Fig. 8 Spearman correlation plots between tumor-specific mRNA expression (TmS) and *TOP1* expression levels in indicated cancer types. Plots were generated at https://wwylab.github.io/TmS/articles/shinyapp.html using data from The Cancer Genome Atlas (TCGA) program.



Supplementary Fig. 9 Gating strategy on flow cytometry to sort for transfected 293T cells. The experiment was performed on Beckman-Coulter Mo-Flo Legacy Cell Sorter. FL2 and FL3 are blue lasers.



Supplementary Fig. 10 Gating strategy for cell cycle analysis of transfected 293T cells. The experiment was performed on BD Symphony A5 cell analyzer.



Supplementary Fig. 11 Purified recombinant TOP1(WT) and TOP1(R546Q) proteins. Coomassie blue staining of recombinant proteins and BSA standard.