Supplemental information

Suppl. Fig. 1 Structure-based sequence alignment of MvNei1 with human endonuclease VIII (Nei)like 1 (hNEIL1; PDB code 1TDH (29)), *Escherichia coli* Nei (EcoNei; PDB code 1K3W (23)), *Lactococcus lactis* Fpg (LlaFpg; PDB code 1XC8 (50)), *Bacillus stearothermophilus* Fpg (BstFpg; PDB code 1R2Y (48)), *Escherichia coli* Fpg (EcoFpg; PDB code 1K82 (22)) and *Thermus thermophilus* Fpg (TthFpg; PDB code 1EE8 (24)). Secondary structure elements are shown above the aligned sequences (β -strands as red arrows and α -helices as green tubes). Conserved residues are highlighted in blue, and the two DNA-binding motifs (helix-two-turns-helix and zinc finger) are boxed. The location of the α F- β 9/10 loop of Fpg and putative lesion recognition of MvNei1 is indicated by a dark green box. The conserved arginine residue at the tip of the zinc(less) finger (277 in MvNei1 and hNEIL1, and 252 in EcoNei) is indicated with an asterisk.

Suppl. Fig. 2 Void filling residues in Fpg/Nei enymes

(A) EcoNei-DNA (PDB code 1K3W; (23)). (B) MvNei1-DNA (PDB code 3A46). (C) LlaFpg-DNA (PDB code 1KFV; (53)). Void filling amino acid residues are shown in gray for EcoNei, in green for MvNei1 and in beige for LlaFpg. The deoxyribitol moiety (dRbl or PED: pentane-3,4-diol-5-phosphate), THF and Pr (1,3-propanediol) are shown in pink.

Supplemental Fig. 1

	_	αA		_	β1				(хΒ		β2		
MvNei1 hNEIL1 EcoNei LlaFpg BstFpg EcoFpg TthFpg	 MPEGPE MPEGPE - PEGPE - PELPE MPELPE - PELPE - PELPE	VALTADIL LHLASQF IRRAADNU VETVRREL VETIRRTU VETSRRG VETSRRG	EKYF /NEAC EAAI EKRI LPLI EPHL RPL\	FKGKTL CRALVF IKGKPL IVGQKI IVGKTI LVGATI LVGATI	- EYI GGCV TDVW - ISI - EDV - EDV - LHA - RQV	DFIS EKS- EAT- EAT- RIFW VVRN VH	GRYSK -SVSF -FPQL -YPRM PNIIF GRLRW -RDPA	(SEPE RN-PE - KT - 1VLTG (HPRD) (PV NPV	GYDD VP Y FEQL SEAF - SEE - SEE TA	FIA F-E QSQL KKEL AAR IYRL	NLPL SSAY LIGC LTGK MIGC LSDC AEGF	KVSNVD RI-SAS HVTHVE TIQGIS TVRGLE PVLSVO RILEVD		56 50 48 56 52 48
	<u> </u>	βЗ			β4		β	5			_	β6		
MvNei1 hNEIL1 EcoNei LlaFpg BstFpg EcoFpg TthFpg	 T <mark>KGK</mark> FL ARGKEL TRGKAL RRGKYL RRGKFL RRAKYL RRGKFL	WFELFDP RLILSPLF LTHFS IFEIG KFLLD LLELP LFALE	PGAQF	DKSNKV PQQEPL NDL DDF RE E(GG\	VYIWN LALVF TLYS FRLIS DALIS GWIII /E <mark>L</mark> VA	TFG <mark>L</mark> RFGM HNQL HLRM HLRM HLGM HLG <mark>M</mark>	TGMWS SGSFQ YGVWF EGKYP EGRYA SGSLF TGGFF	SL QL - VP VVDT VL - AT VASA VASA XL	F REE - GEEP LDAP L - EP E - LP E	EAK LPRI QTTI REKI PLEPI PEKI PTPI	YTRA HAHL RVLF HDHL HTHV HDHV HTRA	VLSFD- RFYTAF VKLQTA TMKFA- VFCFTD DLVMSN ALVLE-		103 105 96 99 102 98 89
	-	β7		β8		αC								
MvNei1 hNEIL1 EcoNei LlaFpg BstFpg EcoFpg TthFpg	 NELM PGPRLA DKT DGQ GSE GKV GRT	AYFSDMRI LCFVDIRF ILLYS LIYADVRI LRYRDVRI LRYRDPRF LYFHDPRF	NFGTF RFGRV - ASDI (FGTV (FGTN RFGAV RFGRL	FKFSNS VDLGG- LEMLRF VELIST VHVYAK VLWTKE LFGVRF	PEQLT TDQVL KEEAD ELEGH RGDYR	KR-K K PY-F RRPP NVLT EI-P	L-NEL WQPGF FLQRV LKKKI LA-EL HL LLLRL	GPDF GPCV GPDV GPEP GPEP GPEP	LK LQEY LD TY LSPA LSDD L S	QQFI PNL FSP FNGI EAF	N RE-S TPEV EDFD AVLA EYLH AFPO	IDDIDIS VLRN-L VKERLL EKLFRE ERAVKT IQKCAKK		148 152 144 147 155 149 139
		αD			_	αE						αF		
MvNei1 hNEIL1 EcoNei LlaFpg BstFpg EcoFpg TthFpg	 KIKKYK ADKAFD SPRFRN KLRKST K K KESA	QPIVALL RPICEAL RQFAGLL KKIKPYL RSVKALL TAIKPWL RPLKALL	IDQKK DQR DQA EQT DQT IDNK DQR H2	(IGS <mark>G</mark> L -FFNG] -FLAGL -LVAGL -VVAG -LVVG -LAAG TH moti	-GNYL (GNYL -GNYL -GNIY -GNIY /GNIY /GNIY	VAEI RAEI RVEI VDEV VDES ASES ADEA	YRAK YRLK WQVG WLAK FRAG FRAG	(IDPH (IPPF) (ITGN) (IHPE) (IHPE) (IHPD) (IHPD) (ISPF)	KLGS EKAP HKAK KETN RPAA RLAS RPAP	SNLTI SVLE DLN/ IQLIE SLSE SLSE SLSE	DQEI EALQ AAQL ESSI SKEI LAEC EEEA	QHRPSF		199 208 194 197 200 194 187
											αG			
MvNei1 hNEIL1 EcoNei LlaFpg BstFpg EcoFpg TthFpg	 ELTLSQ	KIRTKLQ	- ENLV NPDLL - DAL - HLL - ERL - ERL - ERL - RRL Y	VYWIKY LELCHS AHALLE HDSIIE HEEMV ARVIK (RALRE	YETKL SVPKE EIPRF EILQK ATIGE AVLLR EVLAE	AYDS VVQL SYAT AIKL AVMK SIEQ AVEL	-NHIG -GGR- -GGSS -GGST -GGTT - <u>GGST</u> Putat β9	SIR VR LKDFI LSDQ ive le	-YMV GY TYSA TYVN LQSD SYRC sion	IL - G IL - G ITQG G PDG recog	NESS SGEE ST EA KP LP 9niti 310	KIGRKN DF C GK C GT GG On /αF-β1	5	236 253 214 230 234 228 223
MvNei1 hNEIL1 EcoNei LlaFpg BstFpg EcoFpg TthFpg	YHPNIH VDENKH	PTEKEF AAFR/ HG - ALF MQN FQH FQ FQT	-DFL\ AWLR(-RFK\ NELQ\ HLLY\ QELQ\ FRHA\	/ <mark>Y</mark> RK - K C <mark>Y</mark> GMPC /FHRDC /YGKTC /YGRQC /YGRKC / <mark>Y</mark> GREC	(K D GEPCE GEKCS GNPCK GEPCR GLPCP	PNGN RCGS RCGA RCGT VCGT ACGR Zin	KVIAD -MSSL IIEKT EIQKI PIEKT PIEKT PVERP c-fing	OKIIG <u>QDRH</u> TLS- KVA- VVA- KHA- <u>VVA-</u>	SGKN	IKRT GRT GRG GRG QRA GRG K GRG K	TYWA IWFQ FYWC THFC THYC THFC	PAIQKL I <u>GDPG</u> PL PGCQH- PVCQQK PRCQR- RQCQK- PTCQGE		288 288 262 271 274 268 264

Imamura et al. supplemental page 1

Supplemental Fig. 2



Imamura et al. supplemental page 2