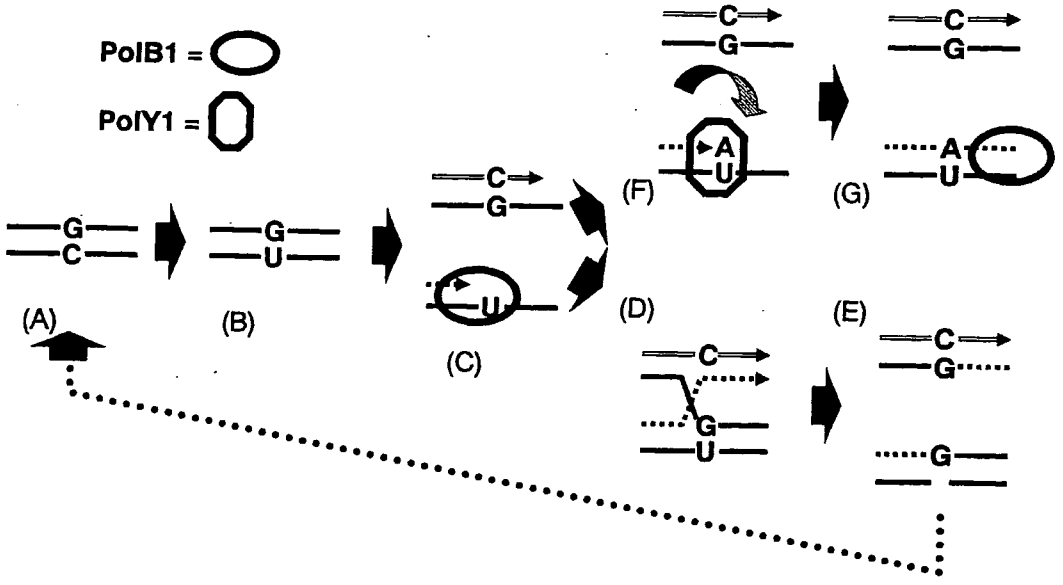

Figure 7.2 Daughter strand gaps by stalling of Pol B1 at template U (or HX) are sealed by HR and/or translesion bypass by Pol Y1. Cytosine in DNA (A) is deaminated by heat, thereby generating G:U mismatch (B). When the strand containing U is copied by Pol B1 before the U is removed, Pol B1 may stop before U (C). ssDNA region downstream of the U can be sealed by HR with sister chromatid (D). UDG may remove the U (E) and provide a chance to regenerate normal G:C basepair (A). However, if Pol B1 is switched to Pol Y1, the Y-family DNA Pol may bypass U by incorporating adenine opposite template U (F). Pol B1 extends the primer strand containing adenine opposite U, thereby generating a mutagenic mispair in DNA (G).

conventional DNA replicases. Whenever they encounter an unrepaired lesion in the template DNA strand, the default replication apparatus would halt. However, most living organisms are equipped with alternative Pols that can tolerate or even correctly accommodate certain abnormalities at the primer-template junction and thus relieve replication blocks. This process termed TLS is faster and energetically less expensive than other DNA reparation mechanisms that do not leave damaged parental strand behind. Depending on DNA lesion context, TLS can result in progeny mutations and that is probably why it is predominantly used under adverse and stressful conditions when an adaptive change in the genotype may be desirable.

According to their primary amino acid sequence similarities signifying common evolutionary origin, Pols are currently classified into seven families, namely A, B, C, D, E, X and Y (Ito and Braithwaite, 1991; Lipps *et al.*, 2003; Ohmori *et al.*, 2001). The enzymes from Y-family are exclusively engaged in the TLS process. They specialize in copying DNA at stalled replication forks, which arise *in vivo* from mismatched or misaligned primer ends usually due to the presence of damage on DNA template. They lack 3′–5′ exonuclease (proofreading) activity and are poorly processive and *error-prone* when copying undamaged DNA resulting often in untargeted mutagenesis (Kim *et al.*, 2001; Kim *et al.*, 1997). The archaeal members of the Y-family Pols have served as great models for solving the three-dimensional structure of these unique enzymes (Ling *et al.*, 2001; Silvian *et al.*, 2001). Their catalytic core consists of finger, palm, thumb and little finger (or wriest or PAD) domains, but the finger and thumb domains are much smaller than in high-fidelity polymerases. An active site cleft consisted of five sequence motifs common to all known Y-family Pols can accommodate DNA and nucleotide triphosphates with relaxed geometric constraints, which result in higher rates of misincorporation and low processivity (Fiala *et al.*, 2007; Ling *et al.*, 2003; Ling *et al.*, 2004b; Zang *et al.*, 2006). Some of these enzymes, however, accurately bypass certain DNA lesions. Examples include accurate bypass across cis-syn thymine-dimer by human Pol eta (Trincao *et al.*, 2001; Washington *et al.*, 2003) and correct

incorporation of cytosine opposite template N^2-guanine minor groove adducts (Choi *et al.*, 2006; Jarosz *et al.*, 2006) such as the steroid hormone derived DNA adduct by human Pol kappa (Suzuki *et al.*, 2004). REV1 and Pol zeta can bypass an AP site efficiently *in vitro* (Haracska *et al.*, 2002; Masuda *et al.*, 2002). Although not essential for 'short-term' survival, the Y-family Pols may play important roles in 'long-term' survival (or evolution) by counteracting DNA adducts resulting from radiation and genotoxic byproducts of metabolism in an energetically efficient way (Yeiser *et al.*, 2002).

The archaeal Y-family Pols have been studied mostly as models of their eukaryotic counterparts due to the availability of the three-dimensional structures and feasibility to do co-crystallization studies. The complexes of Pol Y1 (Dpo4) from *S. solfataricus* with benzo[*a*]pyrene (Chandani and Loechler, 2007; Ling *et al.*, 2004b), *cis-syn* thymine dimer (Ling *et al.*, 2003), 8-oxo-G (Zang *et al.*, 2006) and an AP site (Ling *et al.*, 2004a) have been described in great details. The crystal structures also serve as models for the prediction and manipulation of structural properties of other homologous enzymes from low temperature organisms such as the genetic model *E. coli* (Lee *et al.*, 2006). The three dimensional structure of Pol Y1 (Dpo4) was even used in molecular dynamics simulation study to elucidate the catalytic mechanism and kinetics of DNA lesion bypass at atomic resolution (Chandani and Loechler, 2007; Wang *et al.*, 2006). However, it remains to be determined by *in vivo* genetic experiments what the physiological relevance of the Y-family Pols is in Archaea. A pioneering work related to this topic is the disruption of the *dpo4* gene of *S. solfataricus* (P. Blum, unpublished results). The resulting disruptants display enhanced expression of genes involved in protection of oxidative stress, suggesting the roles of the Y-family DNA Pol in the tolerance against oxidative DNA damage.

Apart from the recently heavily studied *Sulfolobus* enzymes, the Y-family Pol homologues are present in numerous but interestingly not all archaeal species. This contrasts with the other two kingdoms of life, i.e. Eukarya and Bacteria, where at least one Y-family member is present. Of the completely sequenced 38 archaeal genomes, there are only 15 having a classified Y-family Pol homologue as described in the current version of the NCBI GeneBank Protein Sequence Database. We have used the program ClustalX (Thompson *et al.*, 1997) to construct an alignment of all the 18 archaeal protein sequences and generated their phylogenetic tree (Fig. 7.3). From this tree we can distinguish four distinct evolutionary branches representing the homologues found in euryarchaeal halobacteria and methanomicrobia as well as Crenarchaeota. Interestingly, the methanomicrobia group splits into two subgroups. One group including *Methanosarcina acetivorans* and *hungate* possesses only one Y-family homologue while another group including *Methanosarcina burtoni, mazei* and *barkeri* harbour two. Although *Picrophilus torridus* is an Euryarchaeota, its Y-family homologue resembles those of Crenarchaeota probably because it originated by a horizontal gene transfer, which is known to occur between this species and Crenarchaea (Futterer *et al.*, 2004). Because the temperate deepwater planktonic Crenarchaeote 4B7 homologue (Beja *et al.*, 2002) is standing apart from all the branches, it may represent a common evolutionary predecessor. It is likely that at least one of the *Methanosarcina mazei* representatives originates by a lateral gene transfer since this species is known to harbour numerous ORFs whose closest homologues are present in the bacterial domain (Deppenmeier *et al.*, 2002) and its natural habitat is rice paddies, a niche shared by other bacterial species such as *Bacillus subtillis* known to harbour numerous Y-family Pol genes (Sung *et al.*, 2003). One of the related species *Methanosarcina acetivorans* with the largest known archaeal genome of 5.8 MB (Galagan *et al.*, 2002) has only one Y-family Pol.

It still remains an open question of how old the Y-family Pols are on evolutionary scale and whether they precede the chromosomal replicases belonging to the B-family. Given the absence of the homologous genes or pseudogenes in about half of the sequenced archaeal genomes, it is tempting to speculate that Y-family Pols are late comers to compensate for genotoxic effects of an aggressive energy metabolism or life under higher radiation exposures. It is conceivable that the tolerated DNA lesions would become gradu-

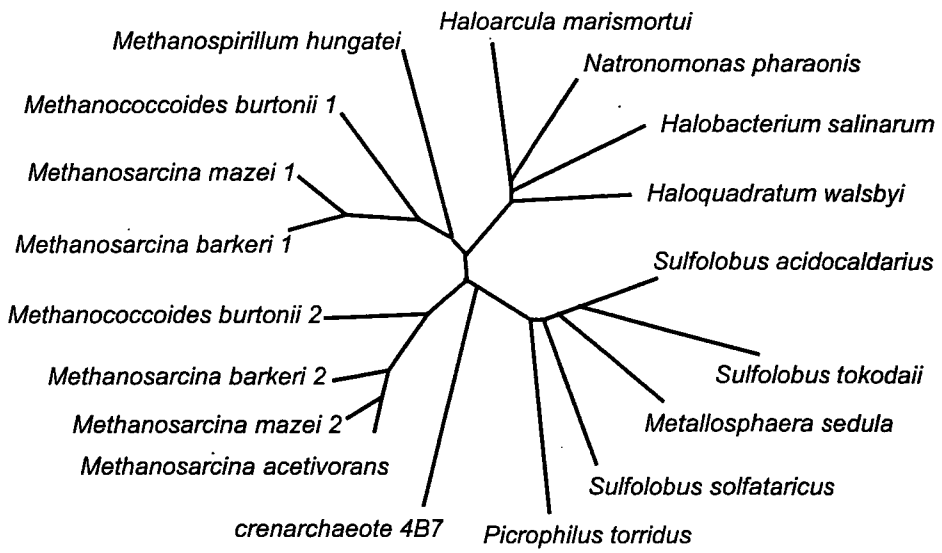

Figure 7.3 Tree showing the phylogeny of Y-family of DNA Pols in archeal species listed in Table 7.2. The tree was drawn using the drawtree program of the PHYLIP phylogeny package based on multiple amino acid sequence alignment generated by the ClustalX program (Thompson *et al.*, 1997). The amino acid sequence alignment is available upon request.

ally diluted under fast growing conditions when symmetric cell division prevails, rather than they would accumulate like in unrepaired genomes of non-dividing cells. In fact, evidence is presented recently that a mechanism may be in place to reduce mutagenesis in non-dividing resting cells by restricting TLS Pol expression to the G2/M phase of cell cycle (Waters and Walker, 2006). Such strategy, i.e. replication first and correctness second, may represent a selective advantage over the opposite strategy, i.e. correctness first and replication second, given that the living environment for Archaea may change rapidly and the most common lesions are bypassed in an *error-free* manner. The obligate symbiont *Nanoarchaeum equitans*, which is deeply branched on an evolutionary tree and probably represents a novel archaeal kingdom, lacks a genuine Y-family homologue despite having an extensive repertoire of other DNA repair enzymes (Waters *et al.*, 2003). Surprisingly, a RecA/RadA recombinase homologue in this species contains a short 53 amino acids segment sharing homology to human Pol iota, a representative of human Y-family Pols (GeneBank accession # AAR39271). The deep water archaeal species like *Pyrococcus* are also lacking Y-family Pols probably because they are well shielded from the surface UV radiation and oxygen. As the major components of archaeal cell membranes are isoprenoid ether-linked glyceroldiethers or tetraethers (De Rosa and Gambacorta, 1988; Itoh *et al.*, 2001), these organisms may be exposed to lower amounts of genotoxic lipid peroxides, which are generated from fatty acid-derived membranes of Bacteria and Eukarya by oxidative metabolism (Chung *et al.*, 2003a). One of the major lipid peroxidation products, trans-4-hydroxy-2-nonenal (Feng *et al.*, 2003; Feng *et al.*, 2004), forms a N^2-dG adduct which is known to be efficiently bypassed in an *error-free* manner by the TLS Y-family Pols (Wolfle *et al.*, 2006; Lone *et al.*, 2007). In fact, at least the DinB homologues, which are the most widespread among all Y-family members, are likely to function during an *error-free* bypass of the N^2-dG minor groove bulky DNA adducts (Jarosz *et al.*, 2006). All the 18 currently sequenced Y-family DNA Pols from Archaea seem to be enzymatically active because they harbour the important catalytic acidic residues in motifs I and III, which bind to the metal Mg^{2+} cofactor. Therefore TLS seems to play an important role in the physiology of certain archaeal species probably by letting them tolerate and become 'immune' against the genotoxic effects of specific DNA lesions resulting either from constant exposure to external radiation or internal reactive by-products of metabolism.

Table 7.2 The archaeal species containing Y-family DNA Pol homologues

Species	Classification	#	Habitat	Metabolism	Genome ref.
Sulfolobus acidocaldarius	*Crenarchaeota Thermoprotei*	1	Acid hot spring, USA Yellowstone Natl. Park, pH 2–3, 80°C	Aerobic, oxidation of elemental sulphur	(Chen *et al.*, 2005)
Sulfolobus tokodaii	*Crenarchaeota Thermoprotei*	1	Beppu hotsprings, Kyushu, pH 2–3, 80°C	Aerobic thermoacidophilic	(Kawarabayasi *et al.*, 2001)
Sulfolobus solfataricus	*Crenarchaeota Thermoprotei*	1	Continental solfatara, Naples, 87°C	Strictly aerobic, produces sulfuric acid	(She *et al.*, 2001)
Crenarchaeote 4B7	*Crenarchaeota Thermoprotei*	1	Temperate water plankton, 200 m deep near the coast of Oregon		(Beja *et al.*, 2002)
Metallosphaera sedula	*Crenarchaeota; Thermoprotei*	1			Unpublished
Picrophilus torridus	*Euryarchaeota; Thermoplasmata*	1	Dry solfataric field in northern Japan, 65°C	Most thermoacidophilic organism, growing on organic substrates, extracellular pH 0, intracellular pH 4.6	(Futterer *et al.*, 2004)
Methanosarcina mazei	*Euryarchaeota; Methanomicrobia*	2	Rice paddies	Fermenting acetate, methylamines and methanol to methane, carbon dioxide and ammonia, metabolically versatile methanogen	(Deppenmeier *et al.*, 2002)
Methanosarcina barkeri	*Euryarchaeota; Methanomicrobia*	2	Freshwater coastal lagoon west of Naples, Italy	Obligate CO_2 reduction with H_2, methyl reduction with H_2, aceticlastic fermentation of acetate or methylotrophic catabolism of methanol, methylated amines and dimethylsulphide, also grows nonmethanogenically with CO, fixes molecular nitrogen, metabolic diversity	(Maeder *et al.*, 2006)
Methanococcoides burtonii	*Euryarchaeota; Methanomicrobia*	2	anoxic and methanesaturated hypolimnion of Ace Lake (Antarctica), 1.7–29.5°C	strictly anaerobic, slightly halophilic, psychrophilic and methylotrophic	Unpublished

Methanosarcina acetivorans	*Euryarchaeota; Methanomicrobia*	1		acetate-utilizing methanogen	(Galagan *et al.*, 2002)
Methanospirillum hungatei	*Euryarchaeota; Methanomicrobia*	1	Anaerobic, freeze-dried granular sludge obtained from an upflow anaerobic sludge bed reactor treating sugar beet wastewater		Unpublished
Halobacterium salinarum	*Euryarchaeota; Halobacteria*	1	High salinity (10 times that of sea water)	Aerobic halophilic chemoorganotroph growing on the degradation products of less halophilic organisms as the salinity reaches near saturation	(Ng *et al.*, 2000)
Haloquadratum walsbyi	*Euryarchaeota; Halobacteria*	1	Spanish solar saltern, thalassic NaCl-saturated environments, sub-lethal conditions of an extremely high $MgCl_2$ concentration and high solar irradiance	Pigmented, synthesize sialic acid and poly-gamma-glutamate capsule, lives in near anoxic conditions due to the extremely high salinity	(Bolhuis *et al.*, 2006)
Natronomonas pharaonis	*Euryarchaeota; Halobacteria*	1	Salt-saturated lakes of pH 11	Grows optimally at pH 8.5 in 3.5 M NaCl, lacks key enzymes of glycolytic pathways but has functional respiration and oxidative phosphorylation, lipoproteins make up one-third of the predicted secretome, harbours blue-light photoreceptor	(Falb *et al.*, 2005)
Haloarcula marismortui	*Euryarchaeota; Halobacteria*	1	Dead Sea	physiologically capable of exploiting diverse environments, large number of environmental response regulators	(Baliga *et al.*, 2004c)

The results are based on the NCBI GeneBank protein sequence database as of March 2007. The number of homologues in each species is indidicated under # and references are shown for each corresponding genome where available. The information about habitat and metabolism is taken from the EMBL-EBI Integr8 database or directly from the specified literature.

Future perspectives

Hyperthermophilic Archaea are interesting model organisms in terms of finding novel molecular strategies of how cells survive in harsh environments. As has been reviewed here, the replicative DNA Pols have unique molecular features that halt DNA synthesis when they encounter deaminated bases, i.e. U or HX, in template DNA. We expect that more unique mechanisms in DNA repair and damage tolerance will be revealed in Archeaa given the wealth of genome sequence information and ease of crystallization of proteins involved in the processing of DNA damage. However, there are several technical limitations that may hamper the progress of research in this area. The first problem may be the difficulty of cultivation. Many hyperthermophilic Archaea should be cultivated at extremely high temperatures and often require strictly anaerobic conditions and inorganic energy sources. Therefore, the actual genome stabilities such as mutation rates of most of hyperthermophilic Archaea are not experimentally determined. In addition, because of the difficulty to culture, information about chemical analysis of DNA lesions is very rare. We suppose that hyperthermophilic Archaea should have massively damaged bases in DNA such as U, 8-OH-G and AP sites. However, quantitative data concerning the DNA damage are scare and thus it is not sure that hyperthermophilic Archaea protect their genome stability mainly via chemical prevention such as using antioxidants to prevent oxidation of DNA or via efficient DNA repair and damage avoidance mechanisms discussed in this chapter. Collaboration with analytical chemists with experts of Archaea is essential to solve the problem. The second problem may be the lack of genetic approaches in particular gene disruption methods. Only a limited number of Archaea, such as *Halobacterium* species, *Sulfolobus solfataricus*, *Pyrococcus Kodakaraensis*, can be utilized for genetic studies. It is hard to evaluate the physiological roles of the mechanisms even if the proteins involved are crystallized and the fine three-dimensional structures are solved. In this regard, it is desirable to establish the gene knockout and knockin approaches to identify the roles of critical amino acids predicted by structural studies. Nevertheless, recent developments, such as measurement of mutation rates based on DNA instead of phenotypes of living organisms (Bielas and Loeb, 2005), high resolution mass spectrometry equipped with high performance liquid chromatography (Zhu *et al.*, 2003) and RNA-based gene knockdown (Collins and Cheng, 2005), will enable hyperthermophilic Archaea to be better model organisms to establish the relationships between genome integrity and a variety of environmental stresses.

Acknowledgements

This work was supported by grants-in-aid for scientific research from the Ministry of Education, Culture, Sports, Science and Technology, Japan, the Ministry of Health, Labour and Welfare, Japan and the Japan Health Science Foundation. Part of the wok was financially supported by the Budget for Nuclear Research of the Ministry of Education, Culture, Sports, Science and Technology, based on the screening and counselling by the Atomic Energy Commission.

References

Aboussekhra, A., Biggerstaff, M., Shivji, M. K., Vilpo, J. A., Moncollin, V., Podust, V. N., Protic, M., Hubscher, U., Egly, J. M., and Wood, R. D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. Cell *80*, 859–868.

Akiba, T., Ishii, N., Rashid, N., Morikawa, M., Imanaka, T., and Harata, K. (2005). Structure of RadB recombinase from a hyperthermophilic archaeon, Thermococcus kodakaraensis KOD1: an implication for the formation of a near-7-fold helical assembly. Nucleic Acids Res. 33, 3412–3423.

Araujo, S. J. and Wood, R. D. (1999). Protein complexes in nucleotide excision repair. Mutat. Res. *435*, 23–33.

Aravind, L., Makarova, K. S., and Koonin, E. V. (2000). SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. Nucleic Acids Res. 28, 3417–3432.

Aravind, L., Walker, D. R., and Koonin, E. V. (1999). Conserved domains in DNA repair proteins and evolution of repair systems. Nucleic Acids Res. *27*, 1223–1242.

Asami, Y., Murakami, M., Shimizu, M., Pisani, F. M., Hayata, I., and Nohmi, T. (2006). Visualization of the interaction between archaeal DNA polymerase and uracil-containing DNA by atomic force microscopy. Genes Cells *11*, 3–11.

Bae, K. W., Baek, K. W., Cho, C. S., Hwang, K. Y., Kim, H. R., Sung, H. C., and Cho, Y. (1999). Expression, purification, characterization and crystallization of flap endonuclease-1 from *Methanococcus jannaschii*. Mol. Cells *9*, 45–48.

Baliga, N. S., Bjork, S. J., Bonneau, R., Pan, M., Iloanusi, C., Kottemann, M. C., Hood, L., and DiRuggiero, J. (2004a). Systems level insights into the stress response to UV radiation in the halophilic archaeon Halobacterium NRC-1. Genome Res. *14*, 1025–1035.

Baliga, N. S., Bjork, S. J., Bonneau, R., Pan, M., Iloanusi, C., Kottemann, M. C., Hood, L., and DiRuggiero, J. (2004b). Systems level insights into the stress response to UV radiation in the halophilic archaeon Halobacterium NRC-1. Genome Res. *14*, 1025–1035.

Baliga, N. S., Bonneau, R., Facciotti, M. T., Pan, M., Glusman, G., Deutsch, E. W., Shannon, P., Chiu, Y., Weng, R. S., Gan, R. R., Hung, P., Date, S. V., Marcotte, E., Hood, L., and Ng, W. V. (2004c). Genome sequence of Haloarcula marismortui: a halophilic archaeon from the Dead Sea. Genome Res. *14*, 2221–2234.

Beja, O., Koonin, E. V., Aravind, L., Taylor, L. T., Seitz, H., Stein, J. L., Bensen, D. C., Feldman, R. A., Swanson, R. V., and DeLong, E. F. (2002). Comparative genomic analysis of archaeal genotypic variants in a single population and in two different oceanic provinces. Appl. Environ. Microbiol. *68*, 335–345.

Bielas, J. H. and Loeb, L. A. (2005). Quantification of random genomic mutations. Nat. Methods *2*, 285–290.

Bolhuis, H., Palm, P., Wende, A., Falb, M., Rampp, M., Rodriguez-Valera, F., Pfeiffer, F., and Oesterhelt, D. (2006). The genome of the square archaeon Haloquadratum walsbyi : life at the limits of water activity. BMC. Genomics *7*, 169.

Bolt, E. L., Lloyd, R. G., and Sharples, G. J. (2001). Genetic analysis of an archaeal Holliday junction resolvase in Escherichia coli. J. Mol. Biol. *310*, 577–589.

Boudsocq, F., Kokoska, R. J., Plosky, B. S., Vaisman, A., Ling, H., Kunkel, T. A., Yang, W., and Woodgate, R. (2004). Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. J. Biol. Chem. *279*, 32932–32940.

Cappelli, E., Taylor, R., Cevasco, M., Abbondandolo, A., Caldecott, K., and Frosina, G. (1997). Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. J. Biol. Chem. *272*, 23970–23975.

Chandani, S. and Loechler, E. L. (2007). Molecular modeling benzo[a]pyrene N2-dG adducts in the two overlapping active sites of the Y-family DNA polymerase Dpo4. J. Mol. Graph. Model. *25*, 658–670.

Chapados, B. R., Hosfield, D. J., Han, S., Qiu, J., Yelent, B., Shen, B., and Tainer, J. A. (2004). Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. Cell *116*, 39–50.

Chen, L., Brugger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H. P., and Garrett, R. A. (2005). The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota. J. Bacteriol. *187*, 4992–4999.

Choi, J. Y., Angel, K. C., and Guengerich, F. P. (2006). Translesion synthesis across bulky N2-alkyl guanine DNA adducts by human DNA polymerase kappa. J. Biol. Chem. *281*, 21062–21072.

Chung, F. L., Pan, J., Choudhury, S., Roy, R., Hu, W., and Tang, M. S. (2003a). Formation of trans-4-hydroxy-2-nonenal- and other enal-derived cyclic DNA adducts from omega-3 and omega-6 polyunsaturated fatty acids and their roles in DNA repair and human p53 gene mutation. Mutat. Res. *531*, 25–36.

Chung, J. H., Im, E. K., Park, H. Y., Kwon, J. H., Lee, S., Oh, J., Hwang, K. C., Lee, J. H., and Jang, Y. (2003b). A novel uracil-DNA glycosylase family related to the helix-hairpin-helix DNA glycosylase superfamily. Nucleic Acids Res. *31*, 2045–2055.

Chung, J. H., Suh, M. J., Park, Y. I., Tainer, J. A., and Han, Y. S. (2001). Repair activities of 8-oxoguanine DNA glycosylase from *Archaeoglobus fulgidus*, a hyperthermophilic archaeon. Mutat. Res. *486*, 99–111.

Cohen, G. N., Barbe, V., Flament, D., Galperin, M., Heilig, R., Lecompte, O., Poch, O., Prieur, D., Querellou, J., Ripp, R., Thierry, J. C., Van der, O. J., Weissenbach, J., Zivanovic, Y., and Forterre, P. (2003). An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*. Mol. Microbiol. *47*, 1495–1512.

Collins, B. K., Tomanicek, S. J., Lyamicheva, N., Kaiser, M. W., and Mueser, T. C. (2004). A preliminary solubility screen used to improve crystallization trials: crystallization and preliminary X-ray structure determination of Aeropyrum pernix flap endonuclease-1. Acta Crystallogr. D. Biol. Crystallogr. *60*, 1674–1678.

Collins, R. E. and Cheng, X. (2005). Structural domains in RNAi. FEBS Lett. *579*, 5841–5849.

Connolly, B. A., Fogg, M. J., Shuttleworth, G., and Wilson, B. T. (2003). Uracil recognition by archaeal family B DNA polymerases. Biochem. Soc. Trans. *31*, 699–702.

De Felice, M., Medagli, B., Esposito, L., De Falco, M., Pucci, B., Rossi, M., Gruz, P., Nohmi, T., and Pisani, F. M. (2007). Biochemical evidence of a physical interaction between *Sulfolobus solfataricus* B-family and Y-family DNA polymerases. Extremophiles. *11*, 277–282.

De Rosa, M. and Gambacorta, A. (1988). The lipids of archaebacteria. Prog. Lipid Res. *27*, 153–175.

Demple, B. and Harrison, L. (1994). Repair of oxidative damage to DNA: enzymology and biology. Annu. Rev. Biochem. *63*, 915–948.

Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R. A., Martinez-Arias, R., Henne, A., Wiezer, A., Baumer, S., Jacobi, C., Bruggemann, H., Lienard, T., Christmann, A., Bomeke, M., Steckel, S., Bhattacharyya, A., Lykidis, A., Overbeek, R., Klenk, H. P., Gunsalus, R. P., Fritz, H. J., and Gottschalk, G. (2002). The genome of Methanosarcina mazei: evidence for lateral gene transfer between Bacteria and Archaea. J. Mol. Microbiol. Biotechnol. *4*, 453–461.

Dianov, G., Price, A., and Lindahl, T. (1992). Generation of single-nucleotide repair patches following excision of uracil residues from DNA. Mol. Cell Biol. *12*, 1605–1612.

Dionne, I., Nookala, R. K., Jackson, S. P., Doherty, A. J., and Bell, S. D. (2003). A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Mol. Cell *11*, 275–282.

DiRuggiero, J., Santangelo, N., Nackerdien, Z., Ravel, J., and Robb, F. T. (1997). Repair of extensive ionizing-radiation DNA damage at 95 degrees C in the hyperthermophilic archaeon *Pyrococcus furiosus*. J. Bacteriol. *179*, 4643–4645.

Eisen, J. A. (1998). A phylogenomic study of the MutS family of proteins. Nucleic Acids Res. *26*, 4291–4300.

Falb, M., Pfeiffer, F., Palm, P., Rodewald, K., Hickmann, V., Tittor, J., and Oesterhelt, D. (2005). Living with two extremes: conclusions from the genome sequence of Natronomonas pharaonis. Genome Res. *15*, 1336-1343.

Fan, L., Arvai, A. S., Cooper, P. K., Iwai, S., Hanaoka, F., and Tainer, J. A. (2006). Conserved XPB core structure and motifs for DNA unwinding: implications for pathway selection of transcription or excision repair. Mol. Cell *22*, 27–37.

Feng, Z., Hu, W., Amin, S., and Tang, M. S. (2003). Mutational spectrum and genotoxicity of the major lipid peroxidation product, trans-4-hydroxy-2-nonenal, induced DNA adducts in nucleotide excision repair-proficient and -deficient human cells. Biochemistry *42*, 7848–7854.

Feng, Z., Hu, W., and Tang, M. S. (2004). Trans-4-hydroxy-2-nonenal inhibits nucleotide excision repair in human cells: a possible mechanism for lipid peroxidation-induced carcinogenesis. Proc. Natl. Acad. Sci. U.S. A *101*, 8598–8602.

Fiala, K. A., Brown, J. A., Ling, H., Kshetry, A. K., Zhang, J., Taylor, J. S., Yang, W., and Suo, Z. (2007). Mechanism of template-independent nucleotide incorporation catalyzed by a template-dependent DNA polymerase. J. Mol. Biol. *365*, 590–602.

Fitz-Gibbon, S. T., Ladner, H., Kim, U. J., Stetter, K. O., Simon, M. I., and Miller, J. H. (2002). Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum aerophilum. Proc. Natl. Acad. Sci. U.S. A *99*, 984–989.

Fogg, M. J., Pearl, L. H., and Connolly, B. A. (2002). Structural basis for uracil recognition by archaeal family B DNA polymerases. Nat. Struct. Biol. *9*, 922–927.

Fondufe-Mittendorf, Y. N., Harer, C., Kramer, W., and Fritz, H. J. (2002). Two amino acid replacements change the substrate preference of DNA mismatch glycosylase Mig.MthI from T/G to A/G. Nucleic Acids Res. *30*, 614–621.

Frosina, G., Fortini, P., Rossi, O., Carrozzino, F., Raspaglio, G., Cox, L. S., Lane, D. P., Abbondandolo, A., and Dogliotti, E. (1996). Two pathways for base excision repair in mammalian cells. J. Biol. Chem. *271*, 9573–9578.

Fujihashi, M., Numoto, N., Kobayashi, Y., Mizushima, A., Tsujimura, M., Nakamura, A., Kawarabayasi, Y., and Miki, K. (2007). Crystal structure of archaeal photolyase from *Sulfolobus tokodaii* with two FAD molecules: implication of a novel light-harvesting cofactor. J. Mol. Biol. *365*, 903–910.

Fujikane, R., Komori, K., Shinagawa, H., and Ishino, Y. (2005). Identification of a novel helicase activity unwinding branched DNAs from the hyperthermophilic archaeon, *Pyrococcus furiosus*. J. Biol. Chem. *280*, 12351–12358.

Futterer, O., Angelov, A., Liesegang, H., Gottschalk, G., Schleper, C., Schepers, B., Dock, C., Antranikian, G., and Liebl, W. (2004). Genome sequence of Picrophilus torridus and its implications for life around pH 0. Proc. Natl. Acad. Sci. U.S. A *101*, 9091–9096.

Gadsden, M. H., McIntosh, E. M., Game, J. C., Wilson, P. J., and Haynes, R. H. (1993). dUTP pyrophosphatase is an essential enzyme in Saccharomyces cerevisiae. EMBO J. *12*, 4425–4431.

Galagan, J. E., Nusbaum, C., Roy, A., Endrizzi, M. G., Macdonald, P., FitzHugh, W., Calvo, S., Engels, R., Smirnov, S., Atnoor, D., Brown, A., Allen, N., Naylor, J., Stange-Thomann, N., DeArellano, K., Johnson, R., Linton, L., McEwan, P., McKernan, K., Talamas, J., Tirrell, A., Ye, W., Zimmer, A., Barber, R. D., Cann, I., Graham, D. E., Grahame, D. A., Guss, A. M., Hedderich, R., Ingram-Smith, C., Kuettner, H. C., Krzycki, J. A., Leigh, J. A., Li, W., Liu, J., Mukhopadhyay, B., Reeve, J. N., Smith, K., Springer, T. A., Umayam, L. A., White, O., White, R. H., Conway, d. M., Ferry, J. G., Jarrell, K. F., Jing, H., Macario, A. J., Paulsen, I., Pritchett, M., Sowers, K. R., Swanson, R. V., Zinder, S. H., Lander, E., Metcalf, W. W., and Birren, B. (2002). The genome of M. acetivorans reveals extensive metabolic and physiological diversity. Genome Res. *12*, 532–542.

Gao, Y. G., Su, S. Y., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H. (1998). The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA. Nat. Struct. Biol. *5*, 782–786.

Gary, R., Kim, K., Cornelius, H. L., Park, M. S., and Matsumoto, Y. (1999). Proliferating cell nuclear antigen facilitates excision in long-patch base excision repair. J. Biol. Chem. *274*, 4354–4363.

Gogos, A. and Clarke, N. D. (1999). Characterization of an 8-oxoguanine DNA glycosylase from *Methanococcus jannaschii*. J. Biol. Chem. *274*, 30447–30450.

Greagg, M. A., Fogg, M. J., Panayotou, G., Evans, S. J., Connolly, B. A., and Pearl, L. H. (1999). A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. Proc. Natl. Acad. Sci. U.S. A *96*, 9045–9050.

Grogan, D. W. (2000). The question of DNA repair in hyperthermophilic Archaea. Trends Microbiol. *8*, 180–185.

Grogan, D. W., Carver, G. T., and Drake, J. W. (2001). Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon Sulfolobus acidocaldarius. Proc. Natl. Acad. Sci. U.S. A *98*, 7928–7933.

Gruz, P., Pisani, F. M., Shimizu, M., Yamada, M., Hayashi, I., Morikawa, K., and Nohmi, T. (2001). Synthetic activity of Sso DNA polymerase Y1, an archaeal DinB-like DNA polymerase, is stimulated by processivity factors proliferating cell nuclear an-

tigen and replication factor C.J. Biol. Chem. *276*, 47394–47401.

Gruz, P., Shimizu, M., Pisani, F. M., De Felice, M., Kanke, Y., and Nohmi, T. (2003). Processing of DNA lesions by archaeal DNA polymerases from *Sulfolobus solfataricus*. Nucleic Acids Res. *31*, 4024–4030.

Haracska, L., Prakash, S., and Prakash, L. (2002). Yeast Rev1 protein is a G template-specific DNA polymerase. J. Biol. Chem. *277*, 15546-15551.

Hayashi, I., Morikawa, K., and Ishino, Y. (1999). Specific interaction between DNA polymerase II (PolD) and RadB, a Rad51/Dmc1 homolog, in *Pyrococcus furiosus*. Nucleic Acids Res. *27*, 4695–4702.

Hinks, J. A., Evans, M. C., De Miguel, Y., Sartori, A. A., Jiricny, J., and Pearl, L. H. (2002). An iron–sulfur cluster in the family 4 uracil-DNA glycosylases. J. Biol. Chem. *277*, 16936-16940.

Hogrefe, H. H., Hansen, C. J., Scott, B. R., and Nielson, K. B. (2002). Archaeal dUTPase enhances PCR amplifications with archaeal DNA polymerases by preventing dUTP incorporation. Proc. Natl. Acad. Sci. U.S. A *99*, 596–601.

Hopfner, K. P., Karcher, A., Craig, L., Woo, T. T., Carney, J. P., and Tainer, J. A. (2001). Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. Cell *105*, 473–485.

Hopfner, K. P., Karcher, A., Shin, D. S., Craig, L., Arthur, L. M., Carney, J. P., and Tainer, J. A. (2000). Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. Cell *101*, 789–800.

Horst, J. P. and Fritz, H. J. (1996). Counteracting the mutagenic effect of hydrolytic deamination of DNA 5-methylcytosine residues at high temperature: DNA mismatch N-glycosylase Mig.Mth of the thermophilic archaeon *Methanobacterium thermoautotrophicum* THF. EMBO J. *15*, 5459–5469.

Hosfield, D. J., Mol, C. D., Shen, B., and Tainer, J. A. (1998). Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. Cell *95*, 135–146.

Hwang, K. Y., Baek, K., Kim, H. Y., and Cho, Y. (1998). The crystal structure of flap endonuclease-1 from *Methanococcus jannaschii*. Nat. Struct. Biol. *5*, 707–713.

Ito, J. and Braithwaite, D. K. (1991). Compilation and alignment of DNA polymerase sequences. Nucleic Acids Res. *19*, 4045–4057.

Itoh, Y. H., Sugai, A., Uda, I., and Itoh, T. (2001). The evolution of lipids. Adv. Space Res. *28*, 719–724.

Jarosz, D. F., Godoy, V. G., Delaney, J. C., Essigmann, J. M., and Walker, G. C. (2006). A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. Nature *439*, 225–228.

Kadyrov, F. A., Dzantiev, L., Constantin, N., and Modrich, P. (2006). Endonucleolytic function of MutLalpha in human mismatch repair. Cell *126*, 297–308.

Kaiser, M. W., Lyamicheva, N., Ma, W., Miller, C., Neri, B., Fors, L., and Lyamichev, V. I. (1999). A comparison of eubacterial and archaeal structure-specific 5′-exonucleases. J. Biol. Chem. *274*, 21387–21394.

Kanugula, S., Pauly, G. T., Moschel, R. C., and Pegg, A. E. (2005). A bifunctional DNA repair protein from *Ferroplasma acidarmanus* exhibits O6-alkylguanine-DNA alkyltransferase and endonuclease V activities. Proc. Natl. Acad. Sci. U.S. A *102*, 3617–3622.

Kanugula, S. and Pegg, A. E. (2003). Alkylation damage repair protein O6-alkylguanine-DNA alkyltransferase from the hyperthermophiles Aquifex aeolicus and *Archaeoglobus fulgidus*. Biochem. J. *375*, 449–455.

Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Otsuka, R., Nakazawa, H., Takamiya, M., Kato, Y., Yoshizawa, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Masuda, S., Yanagii, M., Nishimura, M., Yamagishi, A., Oshima, T., and Kikuchi, H. (2001). Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain7. DNA Res. *8*, 123–140.

Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., and Kikuchi, H. (1998). Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. DNA Res. *5*, 55–76.

Kawashima, T., Amano, N., Koike, H., Makino, S., Higuchi, S., Kawashima-Ohya, Y., Watanabe, K., Yamazaki, M., Kanehori, K., Kawamoto, T., Nunoshiba, T., Yamamoto, Y., Aramaki, H., Makino, K., and Suzuki, M. (2000). Archaeal adaptation to higher temperatures revealed by genomic sequence of Thermoplasma volcanium. Proc. Natl. Acad. Sci. USA *97*, 14257–14262.

Kelner, A. (1949). Effect of Visible Light on the Recovery of Streptomyces Griseus Conidia from Ultra-violet Irradiation Injury. Proc. Natl. Acad. Sci. U.S. A *35*, 73–79.

Kiener, A., Husain, I., Sancar, A., and Walsh, C. (1989). Purification and properties of *Methanobacterium thermoautotrophicum* DNA photolyase. J. Biol. Chem. *264*, 13880–13887.

Kim, K., Biade, S., and Matsumoto, Y. (1998). Involvement of flap endonuclease 1 in base excision DNA repair. J. Biol. Chem. *273*, 8842–8848.

Kim, S. R., Maenhaut-Michel, G., Yamada, M., Yamamoto, Y., Matsui, K., Sofuni, T., Nohmi, T., and Ohmori, H. (1997). Multiple pathways for SOS-induced mutagenesis in Escherichia coli: an overexpression of dinB/dinP results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. Proc. Natl. Acad. Sci. U.S. A *94*, 13792–13797.

Kim, S. R., Matsui, K., Yamada, M., Gruz, P., and Nohmi, T. (2001). Roles of chromosomal and episomal dinB genes encoding DNA pol IV in targeted and untar-

geted mutagenesis in Escherichia coli. Mol. Genet. Genomics *266*, 207–215.

Klungland, A. and Lindahl, T. (1997). Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). EMBO J. *16*, 3341–3348.

Komori, K. and Ishino, Y. (2001). Replication protein A in *Pyrococcus furiosus* is involved in homologous DNA recombination. J. Biol. Chem. *276*, 25654–25660.

Komori, K., Miyata, T., Daiyasu, H., Toh, H., Shinagawa, H., and Ishino, Y. (2000a). Domain analysis of an archaeal RadA protein for the strand exchange activity. J. Biol. Chem. *275*, 33791–33797.

Komori, K., Miyata, T., DiRuggiero, J., Holley-Shanks, R., Hayashi, I., Cann, I. K., Mayanagi, K., Shinagawa, H., and Ishino, Y. (2000b). Both RadA and RadB are involved in homologous recombination in *Pyrococcus furiosus*. J. Biol. Chem. *275*, 33782–33790.

Komori, K., Sakae, S., Daiyasu, H., Toh, H., Morikawa, K., Shinagawa, H., and Ishino, Y. (2000c). Mutational analysis of the *Pyrococcus furiosus* holliday junction resolvase hjc revealed functionally important residues for dimer formation, junction DNA binding, and cleavage activities. J. Biol. Chem. *275*, 40385–40391.

Komori, K., Sakae, S., Shinagawa, H., Morikawa, K., and Ishino, Y. (1999). A Holliday junction resolvase from *Pyrococcus furiosus*: functional similarity to Escherichia coli RuvC provides evidence for conserved mechanism of homologous recombination in Bacteria, Eukarya, and Archaea. Proc. Natl. Acad. Sci. U.S. A *96*, 8873–8878.

Koulis, A., Cowan, D. A., Pearl, L. H., and Savva, R. (1996). Uracil-DNA glycosylase activities in hyperthermophilic micro-organisms. FEMS Microbiol. Lett. *143*, 267–271.

Kvaratskhelia, M., Wardleworth, B. N., Norman, D. G., and White, M. F. (2000). A conserved nuclease domain in the archaeal Holliday junction resolving enzyme Hjc. J. Biol. Chem. *275*, 25540–25546.

Kvaratskhelia, M. and White, M. F. (2000a). An archaeal Holliday junction resolving enzyme from *Sulfolobus solfataricus* exhibits unique properties. J. Mol. Biol. *295*, 193–202.

Kvaratskhelia, M. and White, M. F. (2000b). Two Holliday junction resolving enzymes in *Sulfolobus solfataricus*. J. Mol. Biol. *297*, 923–932.

Lasken, R. S., Schuster, D. M., and Rashtchian, A. (1996). Archaebacterial DNA polymerases tightly bind uracil-containing DNA. J. Biol. Chem. *271*, 17692–17696.

Leclere, M. M., Nishioka, M., Yuasa, T., Fujiwara, S., Takagi, M., and Imanaka, T. (1998). The O6-methylguanine-DNA methyltransferase from the hyperthermophilic archaeon Pyrococcus sp. KOD1: a thermostable repair enzyme. Mol. Gen. Genet. *258*, 69–77.

Lee, C. H., Chandani, S., and Loechler, E. L. (2006). Homology modeling of four Y-family, lesion-bypass DNA polymerases: the case that E. coli Pol IV and human Pol kappa are orthologs, and E. coli Pol V and human Pol eta are orthologs. J. Mol. Graph. Model. *25*, 87–102.

Lilley, D. M. and White, M. F. (2000). Resolving the relationships of resolving enzymes. Proc. Natl. Acad. Sci. U.S. A *97*, 9351–9353.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362*, 709–715.

Lindahl, T. and Nyberg, B. (1974). Heat-induced deamination of cytosine residues in deoxyribonucleic acid. Biochemistry *13*, 3405–3410.

Lindahl, T., Sedgwick, B., Sekiguchi, M., and Nakabeppu, Y. (1988). Regulation and expression of the adaptive response to alkylating agents. Annu. Rev. Biochem. *57*, 133–157.

Lindahl, T. and Wood, R. D. (1999). Quality control by DNA repair. Science *286*, 1897–1905.

Ling, H., Boudsocq, F., Plosky, B. S., Woodgate, R., and Yang, W. (2003). Replication of a cis-syn thymine dimer at atomic resolution. Nature *424*, 1083–1087.

Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001). Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. Cell *107*, 91–102.

Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2004a). Snapshots of replication through an abasic lesion; structural basis for base substitutions and frameshifts. Mol. Cell *13*, 751–762.

Ling, H., Sayer, J. M., Plosky, B. S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D. M., and Yang, W. (2004b). Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase. Proc. Natl. Acad. Sci. U.S. A *101*, 2265–2269.

Lipps, G., Rother, S., Hart, C., and Krauss, G. (2003). A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. EMBO J. *22*, 2516-2525.

Lone, S., Townson, S. A., Uljon, S. N., Johnson, R. E., Brahma, A., Nair, D. T., Prakash, S., Prakash, L., and Aggarwal, A. K. (2007). Human DNA polymerase kappa encircles DNA: implications for mismatch extension and lesion bypass. Mol. Cell *25*, 601–614.

Lu, A. L., Li, X., Gu, Y., Wright, P. M., and Chang, D. Y. (2001). Repair of oxidative DNA damage: mechanisms and functions. Cell Biochem. Biophys. *35*, 141–170.

Maeder, D. L., Anderson, I., Brettin, T. S., Bruce, D. C., Gilna, P., Han, C. S., Lapidus, A., Metcalf, W. W., Saunders, E., Tapia, R., and Sowers, K. R. (2006). The Methanosarcina barkeri genome: comparative analysis with *Methanosarcina acetivorans* and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes. J. Bacteriol. *188*, 7922–7931.

Margison, G. P., Povey, A. C., Kaina, B., and Santibanez Koref, M. F. (2003). Variability and regulation of O6-alkylguanine-DNA alkyltransferase. Carcinogenesis *24*, 625–635.

Marti, T. M., Kunz, C., and Fleck, O. (2002). DNA mismatch repair and mutation avoidance pathways. J. Cell Physiol *191*, 28–41.

Masuda, Y., Takahashi, M., Fukuda, S., Sumii, M., and Kamiya, K. (2002). Mechanisms of dCMP transferase reactions catalyzed by mouse Rev1 protein. J. Biol. Chem. *277*, 3040–3046.

Matsui, E., Kawasaki, S., Ishida, H., Ishikawa, K., Kosugi, Y., Kikuchi, H., Kawarabayashi, Y., and Matsui, I.

(1999). Thermostable flap endonuclease from the archaeon, *Pyrococcus horikoshii*, cleaves the replication fork-like structure endo/exonucleolytically. J. Biol. Chem. *274*, 18297–18309.

Matsui, E., Musti, K. V., Abe, J., Yamasaki, K., Matsui, I., and Harata, K. (2002). Molecular structure and novel DNA binding sites located in loops of flap endonuclease-1 from *Pyrococcus horikoshii*. J. Biol. Chem. *277*, 37840–37847.

McCready, S. and Marcello, L. (2003). Repair of UV damage in Halobacterium salinarum. Biochem. Soc. Trans. *31*, 694–698.

McGlynn, P. and Lloyd, R. G. (2000). Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression. Cell *101*, 35–45.

Modrich, P. and Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu. Rev. Biochem. *65*, 101–133.

Moncalian, G., Lengsfeld, B., Bhaskara, V., Hopfner, K. P., Karcher, A., Alden, E., Tainer, J. A., and Paull, T. T. (2004). The rad50 signature motif: essential to ATP binding and biological function. J. Mol. Biol. *335*, 937–951.

Nakamura, J., Walker, V. E., Upton, P. B., Chiang, S. Y., Kow, Y. W., and Swenberg, J. A. (1998). Highly sensitive apurinic/apyrimidinic site assay can detect spontaneous and chemically induced depurination under physiological conditions. Cancer Res. *58*, 222–225.

Newman, M., Murray-Rust, J., Lally, J., Rudolf, J., Fadden, A., Knowles, P. P., White, M. F., and McDonald, N. Q. (2005). Structure of an XPF endonuclease with and without DNA suggests a model for substrate recognition. EMBO J. *24*, 895–905.

Ng, W. V., Kennedy, S. P., Mahairas, G. G., Berquist, B., Pan, M., Shukla, H. D., Lasky, S. R., Baliga, N. S., Thorsson, V., Sbrogna, J., Swartzell, S., Weir, D., Hall, J., Dahl, T. A., Welti, R., Goo, Y. A., Leithauser, B., Keller, K., Cruz, R., Danson, M. J., Hough, D. W., Maddocks, D. G., Jablonski, P. E., Krebs, M. P., Angevine, C. M., Dale, H., Isenbarger, T. A., Peck, R. F., Pohlschroder, M., Spudich, J. L., Jung, K. W., Alam, M., Freitas, T., Hou, S., Daniels, C. J., Dennis, P. P., Omer, A. D., Ebhardt, H., Lowe, T. M., Liang, P., Riley, M., Hood, L., and DasSarma, S. (2000). Genome sequence of Halobacterium species NRC-1. Proc. Natl. Acad. Sci. U.S. A *97*, 12176-12181.

Nishino, T., Komori, K., Tsuchiya, D., Ishino, Y., and Morikawa, K. (2001). Crystal structure of the archaeal holliday junction resolvase Hjc and implications for DNA recognition. Structure. *9*, 197–204.

Nishino, T., Komori, K., Tsuchiya, D., Ishino, Y., and Morikawa, K. (2005). Crystal structure and functional implications of *Pyrococcus furiosus* hef helicase domain involved in branched DNA processing. Structure. *13*, 143–153.

Nohmi, T. (2006). Environmental stress and lesion-bypass DNA polymerases. Annu. Rev. Microbiol. *60*, 231–253.

Ogrunc, M., Becker, D. F., Ragsdale, S. W., and Sancar, A. (1998). Nucleotide excision repair in the third kingdom. J. Bacteriol. *180*, 5796–5798.

Ohmori, H., Friedberg, E. C., Fuchs, R. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z., and Woodgate, R. (2001). The Y-family of DNA polymerases. Mol. Cell *8*, 7–8.

Olsen, G. J. and Woese, C. R. (1996). Lessons from an Archaeal genome: what are we learning from *Methanococcus jannaschii*? Trends Genet. *12*, 377–379.

Pascucci, B., Stucki, M., Jonsson, Z. O., Dogliotti, E., and Hubscher, U. (1999). Long patch base excision repair with purified human proteins. DNA ligase I as patch size mediator for DNA polymerases delta and epsilon. J. Biol. Chem. *274*, 33696-33702.

Pearl, L. H. (2000). Structure and function in the uracil-DNA glycosylase superfamily. Mutat. Res. *460*, 165–181.

Percival, K. J., Klein, M. B., and Burgers, P. M. (1989). Molecular cloning and primary structure of the uracil-DNA-glycosylase gene from Saccharomyces cerevisiae. J. Biol. Chem. *264*, 2593–2598.

Petit, C. and Sancar, A. (1999). Nucleotide excision repair: from E. coli to man. Biochimie *81*, 15–25.

Podlutsky, A. J., Dianova, I. I., Podust, V. N., Bohr, V. A., and Dianov, G. L. (2001). Human DNA polymerase beta initiates DNA synthesis during long-patch repair of reduced AP sites in DNA. EMBO J. *20*, 1477–1482.

Prakash, S., Johnson, R. E., and Prakash, L. (2005). Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. Annu. Rev. Biochem. *74*, 317–353.

Qian, X., Wu, Y., He, Y., and Luo, Y. (2005). Crystal structure of Methanococcus voltae RadA in complex with ADP: hydrolysis-induced conformational change. Biochemistry *44*, 13753–13761.

Rao, H. G., Rosenfeld, A., and Wetmur, J. G. (1998). *Methanococcus jannaschii* flap endonuclease: expression, purification, and substrate requirements. J. Bacteriol. *180*, 5406–5412.

Rashid, N., Morikawa, M., and Imanaka, T. (1996). A RecA/RAD51 homologue from a hyperthermophilic archaeon retains the major RecA domain only. Mol. Gen. Genet. *253*, 397–400.

Reich, C. I., McNeil, L. K., Brace, J. L., Brucker, J. K., and Olsen, G. J. (2001). Archaeal RecA homologues: different response to DNA-damaging agents in mesophilic and thermophilic Archaea. Extremophiles. *5*, 265–275.

Robb, F. T., Maeder, D. L., Brown, J. R., DiRuggiero, J., Stump, M. D., Yeh, R. K., Weiss, R. B., and Dunn, D. M. (2001). Genomic sequence of hyperthermophile, *Pyrococcus furiosus*: implications for physiology and enzymology. Methods Enzymol. *330*, 134–157.

Roberts, J. A., Bell, S. D., and White, M. F. (2003). An archaeal XPF repair endonuclease dependent on a heterotrimeric PCNA. Mol. Microbiol. *48*, 361–371.

Robinson, H., Gao, Y. G., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H. (1998). The hyperthermophile chromosomal protein Sac7d sharply kinks DNA. Nature *392*, 202–205.

Ruepp, A., Graml, W., Santos-Martinez, M. L., Koretke, K. K., Volker, C., Mewes, H. W., Frishman, D., Stocker, S., Lupas, A. N., and Baumeister, W. (2000). The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum. Nature *407*, 508–513.

Salerno, V., Napoli, A., White, M. F., Rossi, M., and Ciaramella, M. (2003). Transcriptional response to DNA damage in the archaeon *Sulfolobus solfataricus*. Nucleic Acids Res. *31*, 6127–6138.

Sancar, A. (1994). Structure and function of DNA photolyase. Biochemistry *33*, 2–9.

Sancar, A. (1996). DNA excision repair. Annu. Rev. Biochem. *65*, 43–81.

Sancar, G. B. (1990). DNA photolyases: physical properties, action mechanism, and roles in dark repair. Mutat. Res. *236*, 147–160.

Sandigursky, M. and Franklin, W. A. (2000). Uracil-DNA glycosylase in the extreme thermophile *Archaeoglobus fulgidus*. J. Biol. Chem. *275*, 19146-19149.

Sandler, S. J., Hugenholtz, P., Schleper, C., DeLong, E. F., Pace, N. R., and Clark, A. J. (1999). Diversity of radA genes from cultured and uncultured Archaea: comparative analysis of putative RadA proteins and their use as a phylogenetic marker. J. Bacteriol. *181*, 907–915.

Sandler, S. J., Satin, L. H., Samra, H. S., and Clark, A. J. (1996). recA-like genes from three archaean species with putative protein products similar to Rad51 and Dmc1 proteins of the yeast Saccharomyces cerevisiae. Nucleic Acids Res. *24*, 2125–2132.

Sartori, A. A., Fitz-Gibbon, S., Yang, H., Miller, J. H., and Jiricny, J. (2002). A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. EMBO J. *21*, 3182–3191.

Sartori, A. A. and Jiricny, J. (2003). Enzymology of base excision repair in the hyperthermophilic archaeon Pyrobaculum aerophilum. J. Biol. Chem. *278*, 24563–24576.

Savino, C., Federici, L., Johnson, K. A., Vallone, B., Nastopoulos, V., Rossi, M., Pisani, F. M., and Tsernoglou, D. (2004). Insights into DNA replication: the crystal structure of DNA polymerase B1 from the archaeon *Sulfolobus solfataricus*. Structure. *12*, 2001–2008.

Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L. (1995). The structural basis of specific base-excision repair by uracil-DNA glycosylase. Nature *373*, 487–493.

Sayers, J. R. and Artymiuk, P. J. (1998). Flexible loops and helical arches. Nat. Struct. Biol. *5*, 668–670.

Sedgwick, B. (2004). Repairing DNA-methylation damage. Nat. Rev. Mol. Cell Biol. *5*, 148–157.

Seigneur, M., Bidnenko, V., Ehrlich, S. D., and Michel, B. (1998). RuvAB acts at arrested replication forks. Cell *95*, 419–430.

Seitz, E. M., Brockman, J. P., Sandler, S. J., Clark, A. J., and Kowalczykowski, S. C. (1998). RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. Genes Dev. *12*, 1248–1253.

She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P. M., Heikamp-de Jong, I., Jeffries, A. C., Kozera, C. J., Medina, N., Peng, X., Thi-Ngoc, H. P., Redder, P., Schenk, M. E., Theriault, C., Tolstrup, N., Charlebois, R. L., Doolittle, W. F., Duguet, M., Gaasterland, T., Garrett, R. A., Ragan, M. A., Sensen, C. W., and Van der, O. J. (2001). The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc. Natl. Acad. Sci. U.S. A *98*, 7835–7840.

Shuttleworth, G., Fogg, M. J., Kurpiewski, M. R., Jen-Jacobson, L., and Connolly, B. A. (2004). Recognition of the pro-mutagenic base uracil by family B DNA polymerases from Archaea. J. Mol. Biol. *337*, 621–634.

Silvian, L. F., Toth, E. A., Pham, P., Goodman, M. F., and Ellenberger, T. (2001). Crystal structure of a DinB family error-prone DNA polymerase from *Sulfolobus solfataricus*. Nat. Struct. Biol. *8*, 984–989.

Skorvaga, M., Raven, N. D., and Margison, G. P. (1998). Thermostable archaeal O6-alkylguanine-DNA alkyltransferases. Proc. Natl. Acad. Sci. U.S. A *95*, 6711–6715.

Sleeth, K. M., Robson, R. L., and Dianov, G. L. (2004). Exchangeability of mammalian DNA ligases between base excision repair pathways. Biochemistry *43*, 12924–12930.

Slesarev, A. I., Mezhevaya, K. V., Makarova, K. S., Polushin, N. N., Shcherbinina, O. V., Shakhova, V. V., Belova, G. I., Aravind, L., Natale, D. A., Rogozin, I. B., Tatusov, R. L., Wolf, Y. I., Stetter, K. O., Malykh, A. G., Koonin, E. V., and Kozyavkin, S. A. (2002). The complete genome of hyperthermophile Methanopyrus kandleri AV19 and monophyly of archaeal methanogens. Proc. Natl. Acad. Sci. U.S. A *99*, 4644–4649.

Smith, D. R., Doucette-Stamm, L. A., Deloughery, C., Lee, H., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison, D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D., Spadafora, R., Vicaire, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D., Reeve, J. N., and. (1997). Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics. J. Bacteriol. *179*, 7135–7155.

Spies, M., Kil, Y., Masui, R., Kato, R., Kujo, C., Ohshima, T., Kuramitsu, S., and Lanzov, V. (2000). The RadA protein from a hyperthermophilic archaeon *Pyrobaculum islandicum* is a DNA-dependent ATPase that exhibits two disparate catalytic modes, with a transition temperature at 75 degrees C. Eur. J. Biochem. *267*, 1125–1137.

Sung, H. M., Yeamans, G., Ross, C. A., and Yasbin, R. E. (2003). Roles of YqjH and YqjW, homologs of the Escherichia coli UmuC/DinB or Y superfamily of DNA polymerases, in stationary-phase mutagenesis and UV-induced mutagenesis of Bacillus subtilis. J. Bacteriol. *185*, 2153–2160.

Suzuki, N., Itoh, S., Poon, K., Masutani, C., Hanaoka, F., Ohmori, H., Yoshizawa, I., and Shibutani, S. (2004). Translesion synthesis past estrogen-derived DNA adducts by human DNA polymerases eta and kappa. Biochemistry *43*, 6304–6311.

Tanaka, K., Miura, N., Satokata, I., Miyamoto, I., Yoshida, M. C., Satoh, Y., Kondo, S., Yasui, A., Okayama, H., and Okada, Y. (1990). Analysis of a human DNA excision repair gene involved in group A xeroderma pigmentosum and containing a zinc-finger domain. Nature *348*, 73–76.

Tashiro, R., Wang, A. H., and Sugiyama, H. (2006). Photoreactivation of DNA by an archaeal nucleoprotein Sso7d. Proc. Natl. Acad. Sci. U.S. A *103*, 16655–16659.

Taverna, P. and Sedgwick, B. (1996). Generation of an endogenous DNA-methylating agent by nitrosation in Escherichia coli. J. Bacteriol. *178*, 5105–5111.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *25*, 4876–4882.

Trincao, J., Johnson, R. E., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2001). Structure of the catalytic core of S. cerevisiae DNA polymerase eta: implications for translesion DNA synthesis. Mol. Cell *8*, 417–426.

Uemori, T., Sato, Y., Kato, I., Doi, H., and Ishino, Y. (1997). A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*: gene cloning, expression, and characterization. Genes Cells *2*, 499–512.

Vijayvargia, R. and Biswas, I. (2002). MutS2 family protein from *Pyrococcus furiosus*. Curr. Microbiol. *44*, 224–228.

Wang, Y., Arora, K., and Schlick, T. (2006). Subtle but variable conformational rearrangements in the replication cycle of *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4) may accommodate lesion bypass. Protein Sci. *15*, 135–151.

Washington, M. T., Prakash, L., and Prakash, S. (2003). Mechanism of nucleotide incorporation opposite a thymine-thymine dimer by yeast DNA polymerase eta. Proc. Natl. Acad. Sci. U.S. A *100*, 12093–12098.

Waters, E., Hohn, M. J., Ahel, I., Graham, D. E., Adams, M. D., Barnstead, M., Beeson, K. Y., Bibbs, L., Bolanos, R., Keller, M., Kretz, K., Lin, X., Mathur, E., Ni, J., Podar, M., Richardson, T., Sutton, G. G., Simon, M., Soll, D., Stetter, K. O., Short, J. M., and Noordewier, M. (2003). The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. Proc. Natl. Acad. Sci. U.S. A *100*, 12984–12988.

Waters, L. S. and Walker, G. C. (2006). The critical mutagenic translesion DNA polymerase Rev1 is highly expressed during G(2)/M phase rather than S phase. Proc. Natl. Acad. Sci. U.S. A *103*, 8971–8976.

White, M. F. (2003). Archaeal DNA repair: paradigms and puzzles. Biochem. Soc. Trans. *31*, 690–693.

Wolfle, W. T., Johnson, R. E., Minko, I. G., Lloyd, R. S., Prakash, S., and Prakash, L. (2006). Replication past a trans-4-hydroxynonenal minor-groove adduct by the sequential action of human DNA polymerases iota and kappa. Mol. Cell Biol. *26*, 381–386.

Wood, E. R., Ghane, F., and Grogan, D. W. (1997). Genetic responses of the thermophilic archaeon Sulfolobus acidocaldarius to short-wavelength UV light. J. Bacteriol. *179*, 5693–5698.

Woods, W. G. and Dyall-Smith, M. L. (1997). Construction and analysis of a recombination-deficient (radA) mutant of *Haloferax volcanii*. Mol. Microbiol. *23*, 791–797.

Yang, H., Fitz-Gibbon, S., Marcotte, E. M., Tai, J. H., Hyman, E. C., and Miller, J. H. (2000). Characterization of a thermostable DNA glycosylase specific for U/G and T/G mismatches from the hyperthermophilic archaeon Pyrobaculum aerophilum. J. Bacteriol. *182*, 1272–1279.

Yasui, A., Eker, A. P., Yasuhira, S., Yajima, H., Kobayashi, T., Takao, M., and Oikawa, A. (1994). A new class of DNA photolyases present in various organisms including aplacental mammals. EMBO J. *13*, 6143–6151.

Yasui, A. and McCready, S. J. (1998). Alternative repair pathways for UV-induced DNA damage. Bioessays *20*, 291–297.

Yeiser, B., Pepper, E. D., Goodman, M. F., and Finkel, S. E. (2002). SOS-induced DNA polymerases enhance long-term survival and evolutionary fitness. Proc. Natl. Acad. Sci. USA *99*, 8737–8741.

Zang, H., Irimia, A., Choi, J. Y., Angel, K. C., Loukachevitch, L. V., Egli, M., and Guengerich, F. P. (2006). Efficient and high fidelity incorporation of dCTP opposite 7, 8-dihydro-8-oxodeoxyguanosine by *Sulfolobus solfataricus* DNA polymerase Dpo4. J. Biol. Chem. *281*, 2358–2372.

Zhu, H., Bilgin, M., and Snyder, M. (2003). Proteomics. Annu. Rev. Biochem. *72*, 783–812.

DNA REPAIR 6 (2007) 781–788

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/dnarepair

Non-homologous end-joining for repairing I-SceI-induced DNA double strand breaks in human cells

Masamitsu Honma, Mayumi Sakuraba, Tomoko Koizumi, Yoshio Takashima, Hiroko Sakamoto, Makoto Hayashi*

Division of Genetics and Mutagenesis, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

ARTICLE INFO

Article history:
Received 12 July 2006
Received in revised form
4 December 2006
Accepted 4 January 2007
Published on line 12 February 2007

Keywords:
DNA double strand break (DSB)
Non-homologous end-joining (NHEJ)
Homologous recombination (HR)
I-SceI
Deletion
Genomic integrity

ABSTRACT

DNA double strand breaks (DSBs) are usually repaired through either non-homologous end-joining (NHEJ) or homologous recombination (HR). While HR is basically error-free repair, NHEJ is a mutagenic pathway that leads to deletion. NHEJ must be precisely regulated to maintain genomic integrity. To clarify the role of NHEJ, we investigated the genetic consequences of NHEJ repair of DSBs in human cells. Human lymphoblastoid cell lines TSCE5 and TSCE105 have, respectively, single and double I-SceI endonuclease sites in the endogenous thymidine kinase gene (*TK*) located on chromosome 17q. I-SceI expression generated DSBs at the TK gene. We used the novel transfection system (Amaxa Nucleofector) to introduce an I-SceI expression vector into the cells and randomly isolated clones. We found mutations involved in the DSBs in the TK gene in 3% of TSCE5 cells and 30% of TSCE105 cell clones. Most of the mutations in TSCE5 were small (1–30 bp) deletions with a 0–4 bp microhomology at the junction. The others consisted of large (>60) bp deletions, an insertion, and a rearrangement. Mutants resulting from interallelic HR also occurred, but infrequently. Most of the mutations in TSCE105, on the other hand, were deletions that encompassed the two I-SceI sites generated by NHEJ at DSBs. The sequence joint was similar to that found in TSCE5 mutants. Interestingly, some mutants formed a new I-SceI site by perfectly joining the two original I-SceI sites without deletion of the broken-ends. These results support the idea that NHEJ for repairing I-SceI-induced DSBs mainly results in small or no deletions. Thus, NHEJ must help maintain genomic integrity in mammalian cells by repairing DSBs as well as by preventing many deleterious alterations.

1. Introduction

DNA double strand breaks (DSBs) are the most dangerous form of DNA damage. They can be caused by ionizing radiation (IR) or radiometric chemicals, and they can occur spontaneously during DNA replication. Other DNA damage, such as single strand breaks, easily convert to DSBs when a replication fork encounters them [1,2]. The non- or misrepair of DSBs can cause cell death or neoplastic transformation [3,4], so the accurate repair of DSBs is important for maintaining genomic integrity [5]. DSBs are generally repaired through non-homologous end-joining (NHEJ) or homologous recombination (HR) [6,7]. NHEJ joins sequences at the broken ends, which have little or no homology, in a non-conservative manner, and some genetic information is lost. HR, on the other hand, requires extensive tracts of sequence homology and is

* *Corresponding author.* Tel.: +81 3 3700 9847; fax: +81 3 3700 2348.
E-mail address: honma@nihs.go.jp (M. Honma).
1568-7864/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.dnarep.2007.01.004

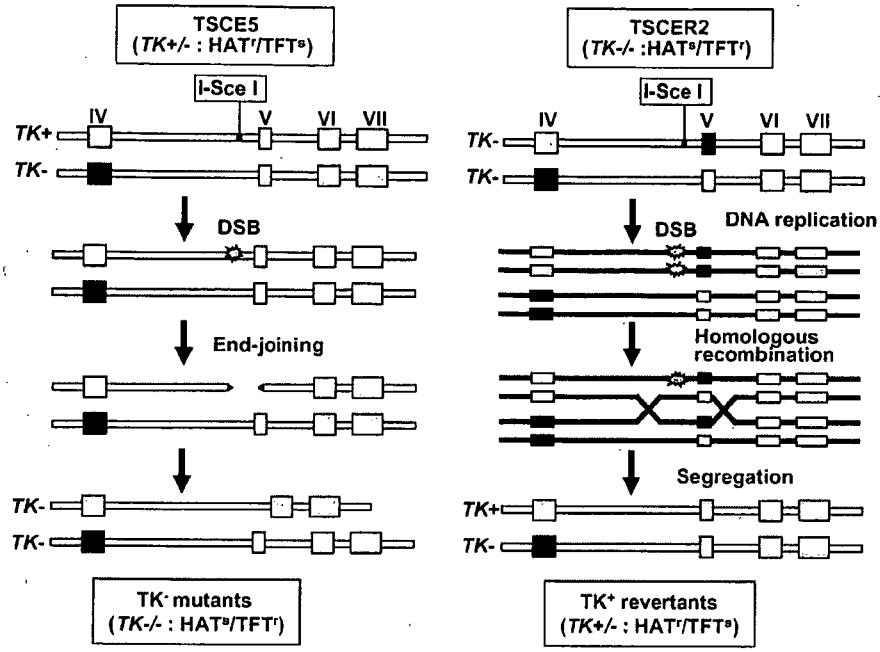

Fig. 1 – Schematic representation of the experimental system. Shadowed and closed rectangles represent the wild type and mutant exons of the *TK* gene, respectively. In TSCE5 cells, when a DSB at the I-SceI site is repaired by NHEJ and causes an exon 5 deletion, TK-deficient mutants are selected in TFT medium. In TSCER2 cells, when a DSB at the I-SceI site is repaired by HR, TK-proficient revertants are selected in HAT medium.

basically error-free [8]. HR is the primary DSB repair pathway in yeast and prokaryotes, but NHEJ is believed to be the primary pathway in mammalian cells [9]. HR is preferable to NHEJ because it is error-free, but NHEJ may have a different way to maintain genomic integrity.

We previously developed a human cell system to trace the fate of a DSB occurring in an endogenous single copy gene (Fig. 1) [10]. The human lymphoblastoid cell line, TSCE5, is heterozygous (+/–) and TSCER2 is compound heterozygous (–/–) for the thymidine kinase gene (*TK*), and both have an I-SceI endonuclease site in intron 4. DSBs can be generated at the I-SceI site by the introduction of an I-SceI enzyme expression vector. When DSBs occur at the *TK* locus, NHEJ in TSCE5 cells produces TK-deficient mutants, while HR between the alleles produces TK-proficient revertants in TSCER2 cells. Positive–negative drug selection for the TK phenotypes permits the distinction between NHEJ and HR repair mechanisms. Using the same system, we previously found that almost all I-SceI-induced DSBs in human cells are repaired by NHEJ and result in mainly 100–4000 bp deletions [10]. Drug selection, however, does not recover cells with genetic changes that are too small to influence TK function, and the resulting spectrum of mutations and reversions may be biased quantitatively as well as qualitatively.

To better understand the fate of DSBs in human cells, we randomly isolated non-selected clones after introducing DSBs and directly analyzed their DNA. A novel transfection system (Amaxa Nucleofector™) can introduce the I-SceI expression vector into most of cell population [11] and efficiently produces DSBs at the *TK* gene. With this improved method, we were able to detect cells with deletions at DSBs without drug selection and to trace the fate of DSBs without bias. We also developed a new cell line that has two I-SceI sites in the *TK* gene and can be used as a model for clustering DSBs. DNA sequence analysis of the mutants in this strain revealed that both single and double DSBs were repaired predominantly by NHEJ, producing only small genetic changes, or none. We discuss how NHEJ maintains genomic integrity.

2. Materials and methods

2.1. *Human cell lines for detecting NHEJ and HR induced by a single DSB*

Human lymphoblastoid cell lines TSCE5 and TSCER2 were previously created from TK6 cells [10], which are heterozygous for a point mutation in exon 4 of the *TK* gene (TK+/–) (Fig. 1). TSCE5 has a 31bp DNA flagment containing the 18 bp I-SceI site inserted 75bp upstream of exon 4 of the TK+ allele and retains TK function. TSCER2 is a TK-deficient mutant spontaneously arising from TSCE5. It has a point mutation (G:A transition) at 23bp of exon 5 of the TK+ allele of TSCE5. TSCER2 is compound heterozygote (TK–/–) for the *TK* gene. NHEJ for a DSB occurring at the I-SceI site results in TK-deficient mutants in TSCE5 cells, while HR between the alleles produces TK-proficient revertants in TSCER2 cells.

2.2. *I-SceI expression and isolation of mutant clones*

We introduced the I-SceI expression vector (pCBASce) by suspending 5×10^6 cells in 0.1 ml Nucleofector solution V (Amaxa

Biosystem, Koeln, Germany) with 50 μg of uncut pCBASce vector (or without the vector as a control), following the manufacturer's recommendations. We then plated the cells into 96-microwell plates at 1 cell/well. Two weeks later, we randomly isolated single colonies and independently expanded them for DNA analysis.

We maintained the cell culture for 3 days and then seeded them into 96-microwell plates in the presence of 2.0 μg/ml trifluorothymidine (TFT) for isolating TK-deficient mutants or HAT (200 μM hypoxanthine, 0.1 μM aminopterin, 17.5 μM thymidine) for isolating TK-proficient revertants. We counted the drug-resistant colonies 2 or 3 weeks later [12] and calculated the mutation and revertant frequencies according to the Poisson distribution [13].

2.3. *Creating a cell line containing two I-SceI sites*

The targeting vector, pTK10, which we had used to make TSCE5 cells, consists of about 6 kb of the original TK gene encompassing exons 5, 6, and 7 and an I-SceI site in intron 4 [10]. We constructed pTK13 by inserting an additional 21 bp DNA fragment containing the 18 bp I-SceI sequence into pTK10 at the Nco I site in intron 5 (152 bp down stream of exon 5) using site-directed mutagenesis (GeneTailer, Invitrogen) (Fig. 4a). To obtain TK-revertant clones with two I-SceI sites in the TK gene, we transfected TSCER2 cells (5×10^6) with 20 μg of linearized pTK13 vector using the Nucleofector system. After 72 h, we seeded the cells into 96-microwell plates containing HAT. We identified one revertant clone, TSCE105, as correctly targeted and confirmed its molecular structure by DNA sequencing.

2.4. *DNA analysis*

To analyze mutations in the isolated TSCE5 and TSCE105 clones, we amplified the part of the TK gene containing the I-SceI sites by PCR, labeling forward primers with a fluorescent dye. We used the following primers for the I-SceI site in intron 4: forward (166F), 5′-TGG GAG AAT TAA GAG TTA CTC C-3′; reverse (196R), 5′-AGC TTC CAC CCC AGC AGC AGC T-3′. We used the following for the I-SceI site in intron 5: forward (251F), 5′-GGA TGG GCA CAG AGA CAC CA-3′; reverse (241R), 5′-CTG ATT CAC AAG CAC TGA AG-3′. For TSCE105 clones, we used 166F and 241R to amplify the regions containing both I-SceI sites. Amplification was performed by denaturation at 96 °C for 5 min, followed by 25 cycles of 96 °C for 30 s, 57 °C for 30 s, 72 °C for 30 s, and extension at 72 °C for 10 min. We analyzed the PCR products using an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) and sequenced them with an ABI 310 genetic analyzer (Applied Biosystems, Foster City, CA).

3. Results and discussions

3.1. *Efficiency of the system for detecting NHEJ and HR repair of chromosomal DSBs using Amaxa nucleofection*

The lymphoblastoid cell lines, TSCE5 and TSCER2, which we previously developed, can trace the genetic consequences of chromosomal DSBs in the human genome. NHEJ for a DSB occurring at the I-SceI site results in TK-deficient mutants in TSCE5 cells, while HR between the alleles produces TK-proficient revertants in TSCER2 cells (Fig. 1) [10]. To introduce the I-SceI expression vector into the cells, we now used the Amaxa nucleofection system. The Amaxa Nucleofector™ can directly transfer DNA into the nucleus of the cells at high efficiency. It was designed for primary cells and hard-to-transfected cell lines such as the human B-cell lymphoblastoid [11,14]. Twenty-four hours after the nucleofection, approximately 65% of the transfected TSCE5 cells expressed the I-SceI enzyme, suggesting that DSBs were efficiently introduced into the cells (data not shown; Takashima et al., under submission).

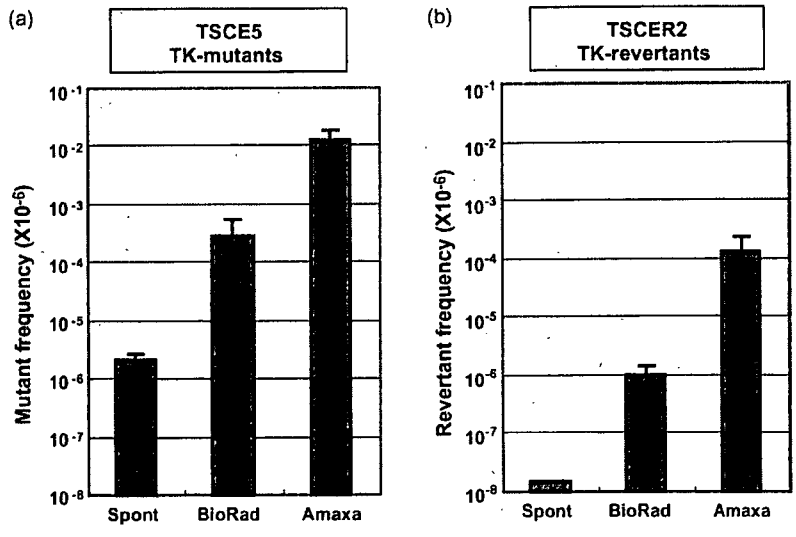

Fig. 2 – Detection of NHEJ and HR repaired DSBs using Amaxa nucleofection or BioRad electroporation. (a) Transfection of TSCE5 with the I-SceI expression vector using Amaxa nucleofection increased the TK-deficient mutant frequency more than 40-fold compared with BioRad electroporation. (b) Transfection of TSCER2 with the I-SceI expression vector using Amaxa nucleofection increased the TK-proficient revertant frequency more than 100-fold compared with BioRad electroporation.

Following Amaxa nucleofection, the mean TK mutant frequency in TSCE5 cells was 1.21%, which was more than 40-fold higher than the frequency we observed with the transfection system we had used previously (BioRad electroporation) (Fig. 2a), and the mean TK-proficient revertant frequency in TSCER2 cells was 1.22×10^{-4}, which was more than 100-fold higher than we observed previously (Fig. 2b). These results demonstrate that the Amaxa nucleofection system effciently introduced the expression vector and generated DSBs with high efficiency in the TSCE5 and TSCER2 cell lines. The relative contribution of NHEJ and HR for reparing the DSBs was 100:1. The value may be biased, however, because the drug selection assay recovers certain classes of NHEJ and HR.

3.2. *Genetic consequences of a chromosomal DSB in non-selected clones*

Because the I-SceI site is inserted into intron 4 of the functional *TK* allele 75 bp upstream of exon 5, any small deletions caused by NHEJ that do not affect TK function will not be recovered as TFT-resistant mutants in the TSCE5 assay. Similarly, in the TSCER2 assay, short tract gene conversion events that do not extend to exon 5 will not be recovered as TK revertants. Thus, recovery of TK mutants and revertants by drug selection may be biased. Because nucleofection can efficiently generate DSBs at the I-SceI site, however, the system enables detection of deletions and recombination in the *TK* gene without drug selection. We randomly isolated 926 transfected clones without TFT selection and directly analyzed DNAs from them. We observed that 29 (3.13%) of them had an I-SceI mutation; these were usually small (<60 bp) deletions, insertions, or rearrangements (Table 1). Fig. 3 shows the DNA sequences of 21 mutants with small genetic changes. Three of them (1659, 1841, and 1893) contained a 1 bp deletion at a CCC tract within the I-SceI site. Others had mostly 0–4 bp microhomologies at the junction, suggesting that the NHEJ machinery was involved. The mutant that had a 1 bp insertion at a TT tract within the I-SceI site (2018) might have been generated by misalignment of the cohesive ends. The mutant that exhibited a complicated DNA rearrangement involving a 50 bp deletion combined with a 9 bp inverted sequence that was a part of deleted sequence (1614) was probably the result of sister chromatid fusion and breakage after DNA replication, as described previously [10]. Five of the mutants showed large deletions (17.2%). This fraction may correspond to the *TK* mutants in the drug selection assay. The large deletions which were commonly detected in the drug selection assay ranged from 1070 to 4030 bp, and had 4–7 bp microhomology at their junctions (data not shown) [10].

One mutant was the product of gene conversion between homologous alleles. It had lost the I-SceI site and retained

Table 1 – Analysis of non-selected TSCE5 clones after I-SceI expression

Total clones	Mutant clones	Mutants (%)
926	29 (Total)	3.13
	23 (Small deletion, insertion, rearrangement; <60 bp)	2.48 (79.3)
	5 (Large deletion; >60 bp)	0.54 (17.2)
	1 (Gene conversion)	0.11 (3.4)

Clone	DNA sequence of TSCE5 mutants around I-SceI site	deletion Size (bp)
ori.	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTATCCCTA CTCTCGAGGATCTGGCAG	
1659	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTAT-CCTA CTCTCGAGGATCTGGCAG	-1
1841	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTAT-CCTA CTCTCGAGGATCTGGCAG	-1
1893	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTAT-CCTA CTCTCGAGGATCTGGCAG	-1
1875	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT—TATCCCTA CTCTCGAGGATCTGGCAG	-2
2099	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGT—CCCTA CTCTCGAGGATCTGGCAG	-3
2399	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTAT——A CTCTCGAGGATCTGGCAG	-4
1573	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTAT——A CTCTCGAGGATCTGGCAG	-4
2182	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTA————TCCCTA CTCTCGAGGATCTGGCAG	-8
2238	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTA————TCCCTA CTCTCGAGGATCTGGCAG	-8
1678	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A CTCTCGAGGATCTGGCAG	-9
1878	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A CTCTCGAGGATCTGGCAG	-9
1907	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A CTCTCGAGGATCTGGCAG	-9
2003	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A CTCTCGAGGATCTGGCAG	-9
2083	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A CTCTCGAGGATCTGGCAG	-9
2183	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A CTCTCGAGGATCTGGCAG	-9
2070	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTT———— ——CTCGAGGATCTGGCAG	-9
1657	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC AT————CCCTA CTCTCGAGGATCTGGCAG	-11
2078	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC AT———————— ————CTGGCAG	-27
1627	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATT———————— ————TGGCAG	-27
2018	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTTATCCCTA CTCTCGAGGATCTGGCAG	+1
1614	TCCGGCCATTTGGC———————— ———————— ———AGGATCTGGCAG	-50+9

Fig. 3 – DNA sequences at the repair junction of 21 of the 26 non-selected I-SceI mutants with small (<60 bp) genetic changes in TSCE5 cells ("ori." is original sequence). The I-SceI recognition site is highlighted in orange. Arrows indicate I-SceI cleavage sites. The 1 bp deletion in the CCC tract is shown in blue and the 1 bp insertion in the TT tract is shown in green. Microhomologous sequences at junctions are shown in red. The sequence in yellow with a left arrow indicates an inverted sequence from part of a deleted sequence. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

intron 4 of the *TK* gene that had been originally connected to the I-*Sce*I site. The appearance of HR mutants was infrequent in the non-biased assay, too, suggesting that I-SceI-induced DSBs are mainly repaired by NHEJ, resulting in small deletions [10,15–17]. This does not mean that HR rarely works for DSBs, however, because our I-SceI system does not cover all HR events.

Most I-SceI systems have been developed using artificial reporter substrates based on exogenous drug-resistance or fluorescence genes and are biased in favor of detecting certain classes of deletions and recombination events [18–20]. In the present system, however, we conducted a survey of DSBs occurring in the endogenous single-copy gene, and investigated the consequences of the DSB without selection bias. We first demonstrated the mutational spectrum induced by I-SceI endonuclease in the human genome. However, it does not necessarily reflect the fate of DSBs occurring spontaneously or induced by irradiation, because our I-SceI system does not monitor sister chromatid HR, which must be the major HR pathway in mammalian cells. Other I-SceI systems setting up two tandem copies of the selective gene on the same chromosome can not also evaluate sister chromatid HR quantitatively, because both chromatids are theoretically cleaved during S/G2 phase. We may underestimate the contribution of HR in the I-SceI system.

Although the I-SceI expression vector was introduced into about 65% of the cells, the frequency of mutants at the I-*Sce*I site in the non-selection assay was still only 3.1%. Three possibilities could explain this: (1) only a small proportion of TSCE5 cells expressing the I-SceI vector may undergo a DSB, (2) most cells with DSBs may undergo apoptosis, and (3) some DSBs may go back to their original sequence by perfect joining. The last possibility would be important to the maintenance of genomic integrity following DSB repair, but its demonstration would be difficult because it is impossible to distinguish between non-cleaved and perfectly repaired I-*Sce*I sites.

3.3. *Genetic consequences of two closely separated DSBs*

To efficiently generate DSBs in the genome, we developed a cell line containing two I-*Sce*I sites in the *TK* gene. We constructed a targeting vector, pTK13, consisting of 6kb of original *TK* gene including exon 5, 6, 7 and two I-*Sce*I sites flank-

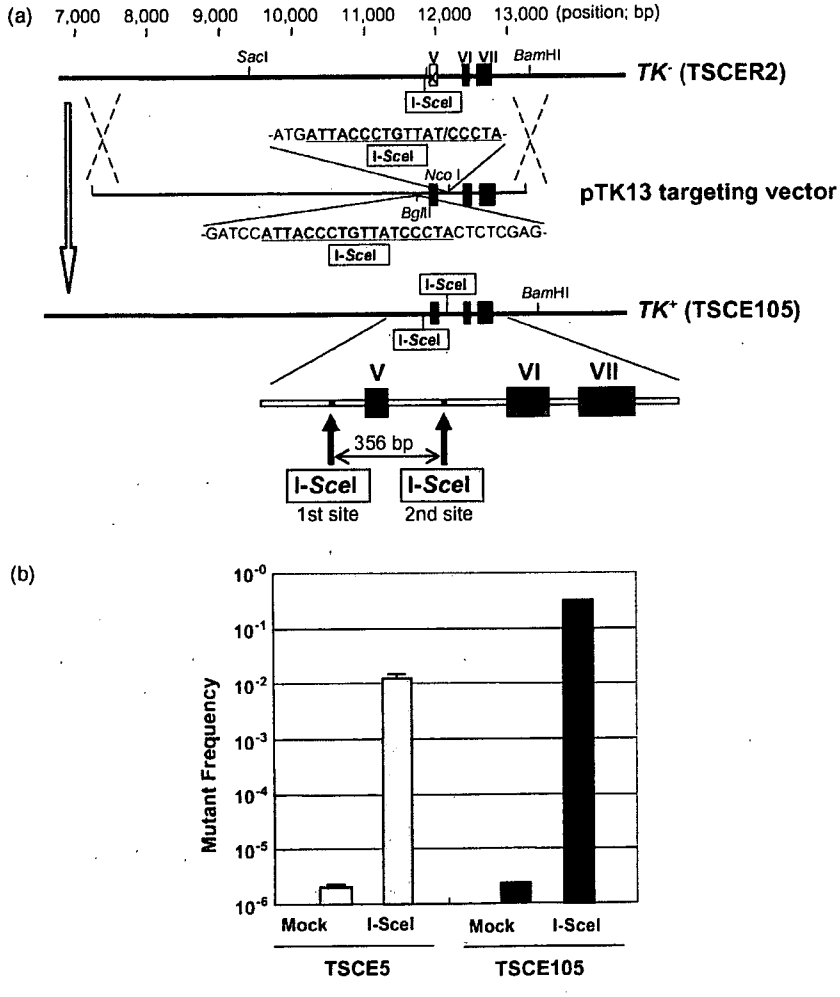

Fig. 4 – (a) Creating the TSCE105 cell line with two I-*Sce*I sites. The functional *TK* allele in the TSCE105 cell line has two I-*Sce*I recognition sites flanking exon 5, 356 bp apart. (b) The TK-deficient mutant frequency in the TSCE105 cells after introduction of DSBs by Amaxa nucleofection. The mutant frequency was 30-fold higher in TSCE105 than in TSCE5.

Table 2 – Analysis of TSCE105 mutants after I-SceI expression

Total Mutants	Type of Mutation	356bp 1st 2nd	Number of Mutants(%)
125	Only 1st I-SceI		1 (0.8)
	Only 2nd I-SceI		47 (38)
	Both, independent		6 (4.8)
	Both, combined		70 (56)
	Perfect joining		4
	Joining with small deletion (<60bp)		31
	Joining with large deletion (>60bp)		29
	Joining with rearrangement		6
	Recombination		1 (0.8)

ing exon 5, and transfected it to TSCER2 cells (Fig. 4a). One HAT-resistant recombinant, TSCE105, had another I-SceI site at intron 5 of the *TK* gene in addtion to the original I-SceI site in TSCE5. The two I-SceI sites are 356 bp apart, flanking exon 5 (Fig. 4a). TSCE105 was also a *TK* heterozygote and was TFT-sensitive. When we nucleofected the I-SceI expressing vector into TSCE105, the TK-deficient mutant frequency by the TFT selection assay, surprisingly, was extremely high (31.3%) (Fig. 4b). We also examined non-selectecd clones after nucleofection. Among 283 non-selected clones, 83 (29.3%) of them had a deletion mutation involving one or both I-SceI sites. This mutation frequency was about the same as the TK-deficient mutation frequency in the TFT selection assay, suggesting that most mutations in TSCE105 were deletions involving coding sequence of the *TK* gene.

To investigate the genetic changes induced by the two DSBs, we analyzed 125 mutants (42 TFT-selected and 83 non-selected) and classified them into 4 types depending on whether they occurred (1) only at the first I-SceI site, (2) only at the second I-SceI site, (3) independenly at both I-SceI sites, or (4) at the combined first and second I-SceI sites (Table 2). The majority (56%) were the last type. Interestingly, four of them joined the two I-SceI sites perfectly, creating a new I-SceI site. Fig. 5 shows the DNA sequences around the joint sites of 26 of the 31 mutants that had small deletions. Almost all of them had a 0–4 bp microhomology at the junction, and the sequences were similar to those found around single DSB repair sites (Fig. 3).

While a single DSB in TSCE5 cells caused predominantly small deletions, two closely occuring DSBs in TSCE105 cells were not repaired independently and caused large deletions involving the two I-SceI sites, indicating that multiple DSBs enhance genetic changes qualitatively as well as quantitatively. Mammalian cells may have difficulty retaining small DNA fragments generated by multiple DSBs. High doses of ionizing irradiation, too, not only increase mutation frequency but also change the mutation type to predominantly large deletions [21,22]. The genomic changes observed in TSCE5 and TSCE105 may reflect a dosage-effect, bringing about different numbers of DSBs. In both cases, however, NHEJ is involved and injury is minimized.

The mutants with perfect joining were generated by NHEJ without exonuclease processing in which the cleaved two flanking I-SceI ends simply join. Most of I-SceI-induced DSBs in TSCE5 and TSCE105 cells may be perfectly joined and create a new I-SceI site. Because the I-SceI enzyme is continuously expressed for at least 48 h after nucleofection (Takashima et al., under submission), the new I-SceI sites generated by perfect joining are cleaved again and again. When the DSBs are occasionally joined after exonuclease processing, they accumulate as deletional mutations and are not cleaved any more (Fig. 6). Thus, the perfect joining by NHEJ is important for repairing DSBs, at least endonuclease-induced DSBs. The perfect joining by NHEJ was also reported in other I-SceI-induced DSB systems [23,24]. Van Heemst et al. demonstrated that a blunt DSB induced by the *E. coli* transposon Tn5 were repaired without loss of nucleotides in Chinese hamster cell lines, suggesting that compatible ends precisely join without deletions [25]. The efficiency or accuracy of precise NHEJ was reduced in Ku80, DNA-PK, XRCC4, or p53 deficient cells [23–26].

NHEJ in mammalian cells involves seven components—Ku70, Ku80, DNA-PKcs, Artemis, XRCC4, Cernunnos/XLF, and Ligase IV [4,7,27–29]. Although the exact role of these proteins remains unknown, three steps have been suggested: (1) end-binding, (2) terminal processing, and (3) ligation [9]. Karanjawala et al demonstrated that defects in Artemis and DNA-PKcs, which are key components in step 2 and possess substantial nucleolytic activity, do not cause severe phenotypes or genomic instability [30]. On the other hand, deficiency of Ku (step 1) or Ligase IV (step 3) confers severe radiosensitivity or lethality [30]. Thus, the second step may not be essential in NHEJ of DSBs, especially of endonuclease-induced DSBs, because the cleaved DNA ends are ligatable

Clone	DNA sequence of TSCE105 mutants around junction site	Deletion Size (bp)
	1st site 2nd site	
perfect	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTATCCCTA GGTCTGTGCAGACTGC	
2412	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTATCCCTA GGTCTGTGCAAACTGC	-356 (0)
2429	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTATCCCTA GGTCTGTGCAAACTGC	-356 (0)
2445	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTATCCCTA GGTCTGTGCAAACTGC	-356 (0)
2465	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTATCCCTA GGTCTGTGCAAACTGC	-356 (0)
2703	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTAT-CCTA GGTCTGTGCAAACTGC	-357 (-1)
2650	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGT——CCCTA GGTCTGTGCAAACTGC	-359 (-3)
2393	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTG——CCCTA GGTCTGTGCAAACTGC	-360 (-4)
2453	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTA————TCCCTA GGTCTGTGCAAACTGC	-364 (-8)
2434	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC AT————ATCCCTA GGTCTGTGCAAACTGC	-365 (-9)
2345	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A GGTCTCTGCAAACGGC	-365 (-9)
2689	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A GGTCTCTGCAAACGGC	-365 (-9)
2714	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A GGTCTCTGCAAACGGC	-365 (-9)
2764	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A GGTCTCTGCAAACGGC	-365 (-9)
2444	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A GGTCTCTGCAAACGGC	-365 (-9)
2446	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A GGTCTGTGCAAACTGC	-365 (-9)
2424	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCT————A GGTCTGTGCAAACTGC	-365 (-9)
2304	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC AT————CCCTA GGTCTGTGCAAACTGC	-367 (-11)
2442	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGT———— ————GCAAACTGC	-372 (-16)
2443	TCCGGGCCAAATGTCCGGAGTTGTCAGATCC ————TA GGTCTGTGCAAACTGC	-372 (-16)
2402	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTAT—— ————AACTGC	-372 (-16)
2425	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTTATC—— ————TGC	-374 (-18)
2435	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ATTACCCTGTT———— ————GC	-378 (-22)
2713	TCCGGGCCAAATGGCCGGAGTTGTC——— ———— ——TGTGCAAACTGC	-384 (-28)
2735	TCCGGGCCAAATGGCCGGAGTTGTCAGATCC ———— ————CTGC	-378 (-22)
2405	TCCGGGCCAAATGGCCG———— ———— —TCTGTGCAAACCGC	-392 (-36)
2437	TCCGGGC———— ———— ————AAACTGC	-409 (-53)

Fig. 5 – DNA sequences at the NHEJ repair junction around the I-SceI junction site in TSCE5 cells. "Perfect" is the DNA sequence when two I-SceI sites join perfectly and create a new I-SceI site (highlighted in orange). Sequences in black are upstream of the first I-SceI site and those in blue are downstream of the second I-SceI site. A total of 26 TSCE105 mutants with deletions combining two I-SceI sites are shown. Underlining indicates a new I-SceI recognition sequence produced by error-free NHEJ. Red indicates microhomologous sequences at junctions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

and do not require terminal processing. Perfect joining by NHEJ probably skips the second step. Naturally occurring DSBs produced by oxidative stress, ionizing radiation, and DNA-damaging agents, however, do not have directly ligatable DNA ends and need some form of nucleolytic processing [7,9]. Their repair by NHEJ results in deletions, even if it works properly. In the present study, the size of the deletions caused by NHEJ, however, were relatively small. No recovered TSCE5 or TSCE105 mutants exhibited large deletions or translocations similar to those frequently observed

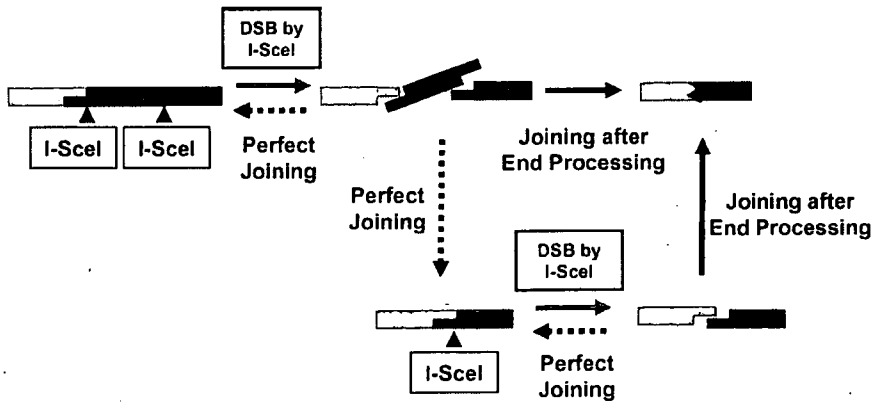

Fig. 6 – A model for NHEJ generating deletions in TSCE105 cells. When a DSB is repaired by perfect joining, an I-SceI site newly generates and is cleaved again. The rare DSB that is joined after exonuclease processing converts to a deletional mutation and accumulates in the cell population.