Published online 30 August 2010 Nucleic Acids Research, 2011, Vol. 39, No. 1 1–18
doi:10.1093/nar/gkq742

SURVEY AND SUMMARY

Natural and engineered nicking endonucleases—from cleavage mechanism to engineering of strand-specificity

Siu-Hong Chan[1,*], Barry L. Stoddard[2] and Shuang-yong Xu[1]

[1]New England Biolabs, Inc. 240 County Road, Ipswich, MA 01938 and [2]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue N. A3-025, Seattle, WA 98109, USA

Received July 15, 2010; Revised July 30, 2010; Accepted August 5, 2010

ABSTRACT

Restriction endonucleases (REases) are highly specific DNA scissors that have facilitated the development of modern molecular biology. Intensive studies of double strand (ds) cleavage activity of Type IIP REases, which recognize 4–8 bp palindromic sequences, have revealed a variety of mechanisms of molecular recognition and catalysis. Less well-studied are REases which cleave only one of the strands of dsDNA, creating a nick instead of a ds break. Naturally occurring nicking endonucleases (NEases) range from frequent cutters such as Nt.CviPII (^CCD; ^ denotes the cleavage site) to rare-cutting homing endonucleases (HEases) such as I-HmuI. In addition to these *bona fida* NEases, individual subunits of some heterodimeric Type IIS REases have recently been shown to be natural NEases. The discovery and characterization of more REases that recognize asymmetric sequences, particularly Types IIS and IIA REases, has revealed recognition and cleavage mechanisms drastically different from the canonical Type IIP mechanisms, and has allowed researchers to engineer highly strand-specific NEases. Monomeric LAGLIDADG HEases use two separate catalytic sites for cleavage. Exploitation of this characteristic has also resulted in useful nicking HEases. This review aims at providing an overview of the cleavage mechanisms of Types IIS and IIA REases and LAGLIDADG HEases, the engineering of their nicking variants, and the applications of NEases and nicking HEases.

A BRIEF INTRODUCTION TO RESTRICTION ENDONUCLEASES

Restriction endonucleases (REases) are an indispensable tool for DNA manipulation in molecular biology. From routine DNA cloning to detection of DNA sequence polymorphism (1,2) to CpG methylation profiling (3), analyses of gene expression in serial analysis of gene expression (SAGE) (4,5) and most recently sample preparation for high-throughput DNA sequencing (6,7), molecular biologists need REases. The discovery of these useful enzymes can be traced back to the research into the physiology of bacteriophage infection in the 1950s where bacteriophage isolated from one *Escherichea coli* strain were 'restricted' from infecting another (8,9). Many of the first REases to be discovered, such as EcoKI and EcoBI, have not become widely used in the manipulation of DNA because they are Type I REases that recognize bipartite sequences separated by 6–8 nt and cleave anywhere near the recognition site to thousands of nucleotides away. Instead, it was the characterization of first few Type IIP REases such as HpaI, HpaII (10–12), EcoRI (13) and HindIII (14) in the early 1970s that allowed the cleavage of DNA in a predictable and consistent manner. The commercial availability of more REases that followed jump-started the modern era of molecular biology (15).

The number of characterized REases of all types has increased greatly since the 1980s in the midst of advances in DNA sequencing technologies and exogenous expression systems in *E. coli* (Figure 1). Many endonucleases are now classified as REases based on their specific recognition sequences and their catalytic activity, without verification of their phage-restricting activity in the host cells. When experimental characterization is absent, ORFs that contains conserved REases catalytic motifs and are located in the vicinity of an ORF that shows homology to

*To whom correspondence should be addressed. Tel: +1 978 380 7267; Fax: +1 978 921 1350; Email: chan@neb.com

© The Author(s) 2010. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

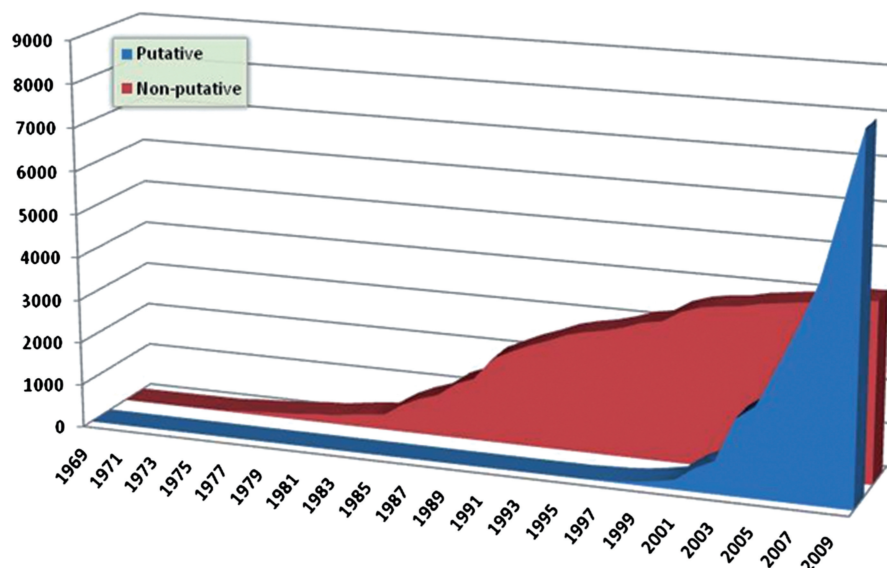

Figure 1. Cumulative number of putative and non-putative R and fused RM genes from 1969 to 2009. The red shaded region indicates the cumulative number of non-putative R and fused RM genes; the blue shaded region indicates the cumulative number of putative genes.

characterized MTases (for Type II REases) or that contain conserved motifs (such as specificity domain of Types I and III) are annotated as putative REases in REBASE (16).

Almost every newly sequenced microbial genome contains ORFs that show notable homology to existing R-M systems. REases are found to have different molecular architectures, sequence requirements for recognition, modes of cleavage and cofactor requirements. They are classified into four main types: Type I through IV (17). Type IIP is the best studied and most widely used type of REases because of their palindromic DNA-recognition sequences and their ability to generate consistent product ends that can be modified or ligated together. There are presently 304 known, distinct Type II specificities (as of June, 2010) represented by 3800 characterized ORFs registered in REBASE. More than 3000 putative genes are yet to be characterized [(16); http://rebase.neb.com/cgi-bin/statlist].

Type IIP REases recognize and cleave within palindromic dsDNA sequences that are usually 4–8 bp in length. These enzymes are usually homodimers where each of two subunits recognizes a single half-site of the DNA target. Therefore, Type IIP REases always recognize palindromic sequences. Related Type IIF enzymes are tetramers (i.e. dimers of dimers) that display similar binding of individual DNA half-sites by individual protein subunits and cleavage of palindromes. Unlike Type IIP enzymes that bind one site at a time, these tetrameric REases usually bind two target sites, cleaving one or both in an allosteric or cooperative manner. In either case, residues that are responsible for target sequence recognition are embedded within the DNA cleavage site, such that cleavage always occurs within the recognition sequence [reviewed extensively in (15,18–22)].

There are, however, naturally occurring endonucleases that make single-strand breaks. These nicking endonucleases (NEases) invariably recognize non-palindromes. They can be *bona fide* nicking enzymes, such as frequent cutter Nt.CviPII (^CCD; D = A, G or T; ^ denotes the cleavage site) (23) and Nt.CviQII (R^AG; R = A or G) (24), or rare-cutting homing endonucleases (HEases) I-BasI and I-HmuI, both of which recognize a degenerate 24-bp sequence (25–27). As well, isolated large subunits of heterodimeric Type IIS REases such as BtsI, BsrDI (28) and BstNBI/BspD6I (29–32) display nicking activity.

NEases have historically been less well-studied and exploited than their ds-cleaving cousins, because they can escape detection by conventional methodologies for screening REases using long but linear DNA substrates (such as λ DNA). Nevertheless, there has been continuous progress in the discovery of natural NEases and engineering of NEases from Types IIS and IIA REases and LAGLIDADG HEases. That has also led to discoveries of different cleavage mechanisms of these interesting enzymes.

TYPES IIS AND IIA REases

The discovery of Type II REases that cleave outside of their recognition sites led to the designation of the 'S' subtype which refers to 'shifted cleavage' (33). The definition was later revised to 'asymmetric recognition and cleavage sites' to highlight the asymmetric nature of the recognition sites. There are, however, Type IIA REases that recognize asymmetric sequence but cleave symmetrically within the recognition sequences (BbvCI cleaves 2 nt downstream of the 5′-end of each strand of CCTCAGC) (17). Studies of these REases that recognize asymmetric sequences have revealed cleavage mechanisms drastically different from those of Type IIP REases. Many Type IIS REases contain separated DNA-binding domain and DNA-cleavage domain, as has been illustrated in the case of FokI (34,35), BfiI (36,37) and MnlI (38). Unlike most Type IIP REases that possess the canonical PD-(D/

E)XK catalytic motif, the catalytic site of Type IIS REases comes in a variety of flavors: ββα-metal motifs similar to those found in non-specific nucleases such as colicins and Holliday junction resolving T4 endonuclease VII, a range of variants of the canonical PD-(D/E)XK catalytic motif and the phospholipase D motif that cleaves DNA without the requirement of divalent metal ions. Whereas Type IIP REases dimerize and bring two catalytic subunits together to make a cut on each DNA strand close enough to result in a ds break, Type IIS REases adopt different approaches for DNA binding and cleavage. For example, the cleavage of the two strands of the ds target sequences of some Type IIS REases has been shown to be strictly sequential (e.g. BfiI, Mva1269I and FokI). The following classification of Type II REases that recognize asymmetric sequences is made based on their molecular architecture and their mode of ds cleavage (Figure 2; Table 1).

Class I: one catalytic site per protein chain; act as homodimers

Examples of this class of enzymes are Type IIS REases FokI, MlyI and AlwI. FokI is the prototypic Type IIS REases due to its early discovery and extensive characterization. FokI possesses a C-terminal PD-(D/E)XK catalytic motif and an N-terminal DNA-binding domain. FokI binds to the recognition site GGATG as a monomer but transiently dimerizes via its DNA-cleavage domain in order to make a ds break 9- and 13-bp downstream at the top and bottom strand, respectively (39–41). In this mode of cleavage, the DNA recognition domain of one of the monomers is bound to its recognition sequence, while the recruited monomer does not have to be (Figure 2). Evident suggested that free FokI molecules in solution display a higher affinity toward DNA-bound FokI monomers than towards unbound recognition sites, so that a bound FokI monomer efficiently recruits a second monomer to make a ds break even at a low enzyme to substrate ratio (42). On the other hand, the cleavage rate of a two-site substrate is over 10-fold faster than a one-site substrate, suggesting that two FokI monomers bound to separate sites in the same DNA molecule associate with each other even more readily than one bound and one free monomer (42). In a two-site substrate, one of the sites is cleaved more efficiently than the other, suggesting a regulatory role of the slow cleavage site (42).

By mutating residues in the DNA-binding domain and the DNA-cleavage domain, Halford and colleagues recently showed that the monomer bounds to the top strand of the target sequence cleaves the bottom strand and the recruited monomer cleaves the top strand (43). When Asn13 of the DNA cleavage domain is mutated to tyrosine and Asp450 of the DNA-binding domain to alanine, a mixture of this N13Y/D450A mutant with a stochiometric ratio of wild-type FokI becomes more efficient in nicking the bottom-strand than making a ds break (Figure 2). The top strand is still cleaved at a lower rate, because homodimerization of the wild-type FokI leads to cleavage of the second strand. On the other hand, when the DNA-cleavage domain mutant (N13Y) was mixed with the DNA-binding domain mutant (D450A), only the top strand is cleaved (43), generating a useful top-strand nicking enzyme (Figure 2).

Disruption of dimerization can also result in nicking variants of Class I enzymes. Existing as monomer in the absence of target DNA, MlyI dimerizes on the target DNA to make a blunt end ds cleavage 5-nt downstream of GAGTC (30). Mutations that abolish the dimerization of MlyI results in a variant that nicks the top strand of the substrate (44). Similarly, impairing dimerization of AlwI (GGATC 4/5) by swapping the C-terminal amino acid sequence with that of Nt.BstNBI results in a top-strand nicking AlwI variant (45).

Class II: two catalytic sites per protein chain; act as monomers

Type IIS REases Mva1269I, BsrI, BsmI and BtsCI are examples of this class of enzymes. Mva1269I (GAATGC 1/−1) contains a canonical PD-(D/E)XK and a non-canonical PD-X_n-E-X_{12}-QR catalytic site and exists as a monomer in both solution and DNA-bound form. Mutagenesis experiments have identified two catalytic sites where the N-terminal non-canonical catalytic site cleaves the bottom strand and the C-terminal canonical catalytic site cleaves the top strand of the target sequence (Figure 2). In addition, cleavage of the bottom strand is a prerequisite for the cleavage of the top strand, as substitution of the scissile phosphodiester bond at the bottom strand with phosphorothioate linkage attenuates the cleavage of the top strand but not *vice versa* (46). This nick-induced cleavage of the opposite strand has also been reported in Holliday junction-resolving T7 endonuclease I (47,48) and BfiI (ACTGGG 5/4) (49,50), although little homology is found between Mva1269I and T7 endonuclease I, and a completely different catalytic mechanism is employed by BfiI (see below).

Sequence alignment has revealed the homologous catalytic sites BsrI (ACTGG 1/-1), BsmI (GAATGC 1/–1) and BtsCI (GGATG 2/0). By optimization of mutations at each of the catalytic sites, BtsCI mutants that efficiently nick the top or the bottom strand specifically have been created (51).

Three thermostable Type IIS enzymes that recognize related sequences [BsaI (GGTCTC 1/5), BsmAI GTCTC (1/5) and BsmBI CGTCTC (1/5)] have also been proposed to contain two separate catalytic sites (52). By a combination of random mutagenesis and selection, top-strand (R236D) and bottom-strand nicking variants (N441D/R442G) of BsaI have been created with negligible ds-cleavage activity (52). Equivalent mutations have been incorporated into BsmAI and BsmBI. For BsmBI, variant R233D is a top-stranded nicking enzyme with a low level of ds-cleavage activity. Variant R438D nicks the bottom strand with no observable ds-cleavage activity. For BsmAI, variant R221D is a top-strand nicking without observable ds-cleavage activity. However, the bottom-nicking BsmAI variant created by replicating the bottom-strand nicking mutations of BsaI still exhibited substantial ds-cleavage activity. Additional mutagenesis had only yielded top-strand nicking enzyme variants. On

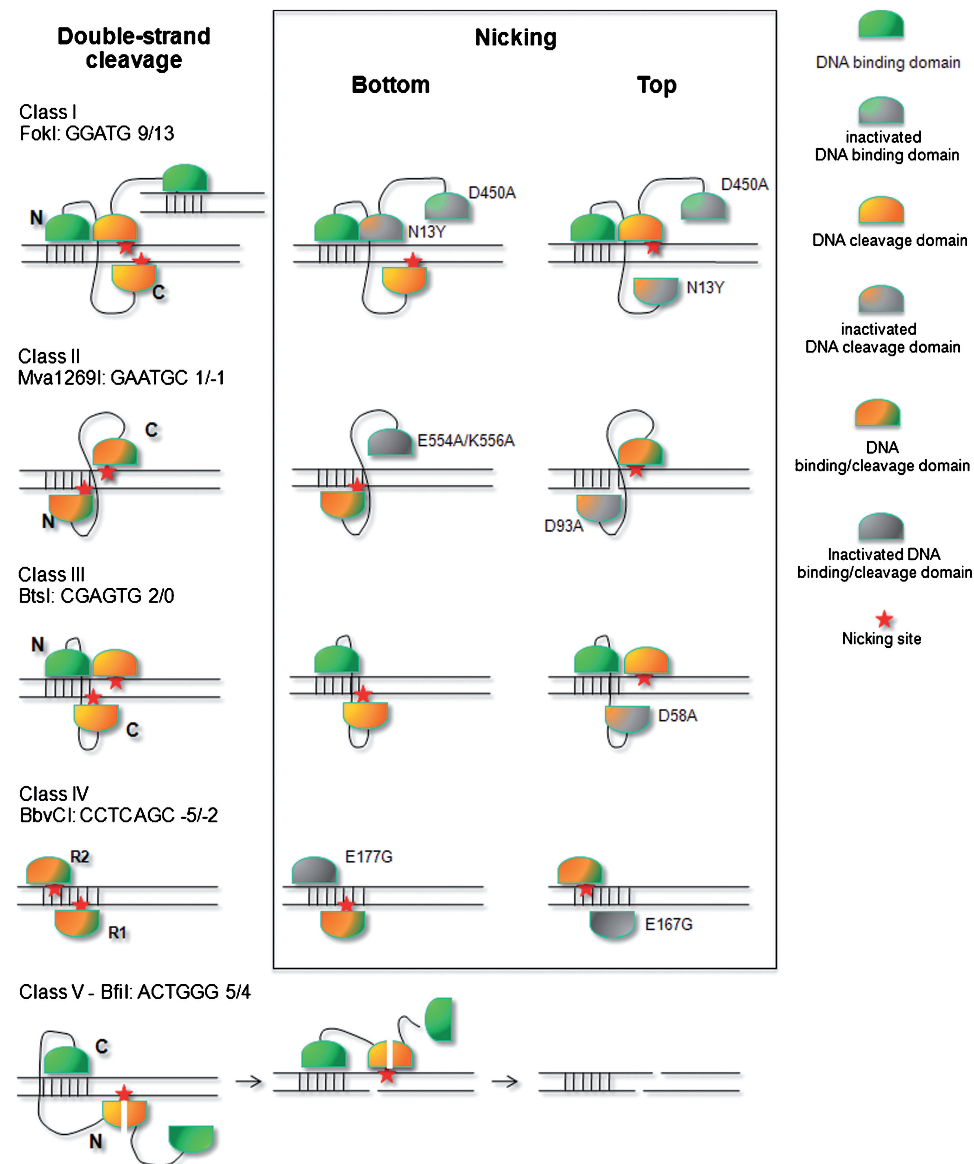

Figure 2. Five classes of restriction endonucleases that recognize asymmetric sequences and their nicking variants. Class I: homodimeric with one catalytic site per protein chain. Class II: monomeric with two catalytic sites per protein chain. Class III: heterodimeric with one catalytic site per protein chain; the large subunit is active in the absence of the small subunit. Class IV: heterodimeric with one catalytic site per protein chain; the two subunits are of comparable size and none of the monomers is active in the absence of the partner. Class V: homodimeric with a half catalytic site (phospholipase D-type) per protein chain. In Classes II and IV, information on the segregation of the DNA-binding domain and the DNA-cleavage domain is not available. Therefore, single semicircles are used to represent a combined DNA binding/cleavage domain. In all cases, mutations that inactivate the cleavage activity are indicated.

the hand, by using a randomized mutagenesis approach in combination with a stand-specific nicking activity selection methodology, top- and bottom-strand nicking variants of SapI (GCTCTTC 1/4) have been identified (see below) (53). Similar amino acid substitutions introduced into BspQI, a thermostable isoschizomer of SapI, have resulted in BspQI nicking variants (54).

Class III: one catalytic site per protein chain; act as heterodimers; the large subunit is active in the absence of the small subunit

Biochemically characterized examples include Type IIS enzymes BtsI, BsrDI, BspD6I and BstNBI. BtsI (CGAGTG 2/0) and BsrDI (GCAATG 2/0) recognize similar target sequences and cleave at the same position. They possess the same catalytic sites as the Class II enzymes Mva1269I, BsrI and BsmI, but in separate protein subunits (Figure 2). The large subunits contain a canonical PD-(D/E)XK catalytic motif similar to the C-terminal domain of Mva1269I, whereas the small subunits contain a non-canonical PD-X_n-E-X_{12}-QR catalytic motif that resembles the N-terminal catalytic site of Mva1269I (28). These Type IIS REases are also classified as Type IIT because of their heterodimeric architecture. In the heterodimeric state, BsrDI and BtsI cleave their target sequences on both strands. Intriguingly, the isolated large

Table 1. Classes of Type II restriction endonucleases that recognize asymmetric sequences and examples

Class I (one catalytic site per protein chain; act as homodimers)	
FokI	GGATG 9/13
MlyI	GAGTC 5/5
AlwI	GGATC 4/5
Class II (two catalytic sites per protein chain; act as monomers)	
Mva1269I	GAATGC 1/–1
BsrI	ACTGG 1/–1
BsmI	GAATGC 1/–1
BtsCI	GGATG 2/0
Class III (one catalytic site per protein chain; act as heterodimers; the large subunit is active in the absence of the small subunit)	
BtsI	GCAGTG 2/0
BsrDI	GCAATG 2/0
BstNBI/BspD6I	GAGTC 4/6
Class IV (one catalytic site per protein chain; act as heterodimers; neither of the monomers is active in the absence of the other subunit)	
BbvCI	CCTCAGC –5/–2
Bpu10I	CCTNAGC –5/–2
Class V [half of a catalytic site (phospholipase D-type) per protein chain; act as homodimers on each strand sequentially]	
BfiI/BmrI	ACTGGG 5/4

subunits of both enzymes nick the bottom strands of their target sites efficiently (28). The large subunits of BtsI and BsrDI are therefore named Nb.BtsI and Nb.BsrDI, respectively. Interestingly, the isolated small subunits of both enzymes do not display any nicking activity (28). In an attempt to show that the small subunits actually carry an active DNA cleavage site and to develop top strand-nicking enzymes, the catalytic site of the large subunits of both BtsI and BsrDI were inactivated through mutagenesis. The reconstituted heterodimers, consisting of the wild-type small subunits and cleavage-deficient large subunits (L^-S^+) nick the top strand specifically, resulting in top strand-nicking Nt.BsrDI (GCAATG 2/−) and Nt.BtsI (GCAGTG 2/−) (28). It also shows that each of the large and small subunits of BtsI and BsrDI carry an active DNA-cleavage site that cleaves the bottom and the top strand, respectively.

A similar molecular architecture has been reported in the heterodimeric BstNBI/BspD6I (GAGTC 4/6) REases, which are identical proteins described by two independent groups. In both cases, the large subunits were initially identified as top-strand nicking enzymes (29–31), while later reports indicated that the presence of a smaller protein encoded by the adjacent ORF resulted in ds cleavage activity (32) (G. Wilson and H. Kong, personal communication). Similar to BsrDI/BtsI, the small subunit possesses a catalytic site that is only functional when complexed with the large-nicking subunit. The crystal structures of Nt.BspD6I and its small subunit (ssBspD6I) show that although overall sequence similarity is low, the DNA-binding domain of Nt.BspD6I is structurally homologous to that of FokI (55). More interestingly, the ssBspD6I has relatively high-sequence similarity and structure homology to the catalytic domain of the large nicking subunit (55), suggesting two possible routes of evolution of BspD6I: (i) duplication of the catalytic domain as the small subunit; (ii) deletion of the DNA-binding domain of one of the subunits of an ancestral homodimeric BspD6I REase; followed by additional selective pressure that optimized the amino acid sequence of the small subunit for dimerization and second-strand cleavage activity. Resolution of more atomic structures of this class of heterodimeric Type IIS REases will tell if this apparent domain duplication phenomenon is typical.

Class IV: one catalytic site per protein chain; act as heterodimers; neither of the monomers is active in the absence of its partner

Type IIA BbvCI REase (CCTCAGC −5/−2) has an interesting variation of the heterodimeric molecular architecture. Unlike Type IIS BtsI and BsrDI, where the subunits differ in size substantially (38 kDa versus 18 kDa in BtsI, 56 kDa versus 25 kDa in BsrDI), the two subunits of BbvCI are comparable in size (31 and 32 kDa, respectively). Neither isolated subunit of BbvCI can nick DNA (Figure 2). However, when rendered catalytically inactive by mutagenesis, the $R1^-R2^+$ heterodimers nick the top strand of the target sequence, while $R1^+R2^-$ heterodimers nick the bottom strand only (56,57). This indicates that each of the R1 and R2 subunits contain a catalytic site that cleaves the bottom and top strand of the target sequence, respectively. It also implies that each subunit has its own DNA-binding domain that recognizes half of the recognition site. Cleavage within the recognition sequence suggests that the amino acid residues responsible for DNA recognition and cleavage are not separated into different domains—much like the canonical Type IIP REases. The REase Bpu10I (CCTNAGC −5/−2) that recognizes a highly similar target sequence and adopts a similar molecular architecture (58) is amenable to similar manipulations for the generation of strand-specific nicking variants. BseYI (CCCAGC −5/−2) contains two subunits with 58.3 and 37.1 kDa, respectively, and recognize a target sequence similar to BbvCI's. It is likely that BseYI also belongs to this class of Type IIA REases (R. Morgan and X. Y. Xu, unpublished data).

Class V: half of a catalytic site (phospholipase D-type) per protein chain; act as homodimers on each strand sequentially

In Mva1269I, nicking of the bottom strand is a prerequisite for cleavage of the top strand. This nick-dependent cleavage of the second strand has also been observed in BfiI (ACTGGG 5/4) (49,50). BfiI and its isoschizomer BmrI, contains the phospholipase D (PLD) catalytic motif (37). By attacking the scissile phosphodiester bond directly with a nucleophilic histidine, the PLD-type nucleases do not require metal ions for cleavage. The catalytic site consists of two copies of the HXK motif, where the histidine from one motif acts as a nucleophile that attacks the scissile phosphodiester bond directly, and the histidine from the other HXK motif acts as a proton donor to

protonate the 3′ leaving group (59). In BfiI/BmrI, only one HXK motif is present per protein chain. Therefore, dimerization is needed to generate a complete catalytic site at the dimerization interface (Figure 2). The cleavage of the two strands of ds-DNA is strictly sequential: the bottom strand is cleaved before the top strand (50,60). The possibility of oligomerization and hairpin formation for second-strand cleavage has been ruled out, suggesting that ds cleavage is achieved by protein and/or DNA conformation rearrangement. As in other PLD nucleases, the cleavage reaction of BfiI proceeds through a covalent phosphohistidine intermediate where the nucleophilic histidine (His105) from one subunit attacks and is covalently linked to the 5′-phosphate of the cleaved strand, while the histidine from another subunit protonates the hydroxyl of the 3′-leaving group. The covalent phosphohistidine is then attacked by a nucleophilic water molecule activated by His105 of the other subunit to complete the hydrolysis (49). Whether the histidine from both subunits takes on a different role by chance or each of them serves a specific role was unknown until recently. By creating heterodimeric BfiI variants of WT/H105A and their N-terminal truncations, it was shown that the His105 of the subunit not bound to the recognition sequence acts as a nucleophile to attack both strands, whereas His105 of the subunit bound to the recognition sequence protonates the leaving 3′-group and activates a water molecule to hydrolyze the phosphohistidine intermediate for both strands (61), implying a change of orientation of the protein between the two cleavage events. Inhibition of this re-orientation mechanism will confirm this hypothesis and probably facilitate the creation of bottom-strand nicking BfiI variants.

The sequential cleavage of the two strands has allowed for verification of the presence of substrate-assisted catalysis (62). Using bottom-strand nicked substrate with or without the 5′-phosphate at the nicked site, Siksnys and co-workers showed that at $pH < 8.5$, the cleavage of the top strand is highly reduced in the presence of the 5′-phosphate (50). The absence of 5′-phosphate resulted in a overall low cleavage of the top strand. Therefore, it was concluded that the deprotonation of the 5′-phosphate of the bottom strand is required for the cleavage of the top strand, consistent to the substrate-assisted catalysis model (62).

Homing endonucleases

Homing is the process by which mobile microbial intervening genetic sequences—groups I or II introns or inteins—are duplicated into host genes that lack such a sequence (63–67). This process is induced by a site-specific endonuclease encoded by an ORF harbored within the intervening sequence (68). The endonuclease specifically recognizes a target sequence corresponding to the intron insertion site and usually generates a double-strand break that is repaired by the cellular DNA-repair machinery, leading to the duplication of the intron and the endonuclease ORF into the target site. However, in the case of both certain naturally occurring and engineered HEases, the generation of a single-strand break appears to be sufficient for homing and/or targeted homologous recombination.

HEase genes are widespread and found within introns and inteins in all biological super-kingdoms; each is associated with a unique host genome (ranging from phage, to cyanobacteria and eubacteria, to single-cell eukaryotes) (69). In order to promote precise intron transfer and avoid deleterious cleavage of their host genomes, HEases are highly specific to their homing site by recognizing longer DNA sequences (15–37 bp). However, they exhibit sufficient site recognition flexibility to retain genetic mobility in the face of target site variation across diverging host strains. Unlike REases that make side-chain-base contacts to every base within a recognition site, HEases use a strategy in which variable numbers of contacts are made to individual base pairs across a long-target site to achieve overall high specificity but variable recognition fidelity across the site (70). Target site polymorphism that are tolerated by HEases are strongly correlated with the polymorphism of the sequence of the host target site (71). In an extreme case, the nicking HEase I-BasI from *Bacillus* phage *Bastille* can bind to and nick the intronless site of the DNA-polymerase gene of both *Bastille* and *SPO1* strain, even though the sequence of the two strains are very different (27).

HEases can be classified into four major classes based on their structural folds: ββα-metal (includes HNH and His–Cys HEases), GIY-YIG, LAGLIDADG and PD-(D/E)XK. ββα-metal and GIY-YIG HEases are found in both free-standing and intron-associated HEases in phage and in many cases generate single-strand nicks in their DNA target sites (69). These two folds have also been found and studied in Type IIP REases (Figure 3A) (72–75) (Mak *et al.*, in press). Although having a different structure fold, the catalytic residues of the ββα-metal and GIY-YIG endonucleases are spatially conserved and these enzymes appear to share a very similar single-metal hydrolytic mechanism where a tyrosine or histidine residue acts as an activated base to generate a nucleophilic hydroxylate ion from a water molecule (Figure 3B).

The ββα-metal, GIY-YIG and LAGLIDADG catalytic motifs, in the absence of structural embellishments or accessory domains, are relatively non-specific endonucleases, often put to use by phage and bacteria as part of an arsenal of degradative enzymes that can assist in invasion or cellular defense (76,77). In the case of HEases, the nuclease domains are generally tethered to additional DNA-binding regions (78,79), and may also contain additional inserted sequence motifs that facilitate target specificity (80). The presence of the PD-(D/E)XK motif commonly used by REases in HEases is only become apparent recently with the resolution of the crystal structure of I-Ssp6803I in the presence of its target DNA (81). The structure shows that I-Ssp6803I adapts a tetrameric configuration for binding to the long-homing site but only uses two of the subunits for cleavage of the target homing site.

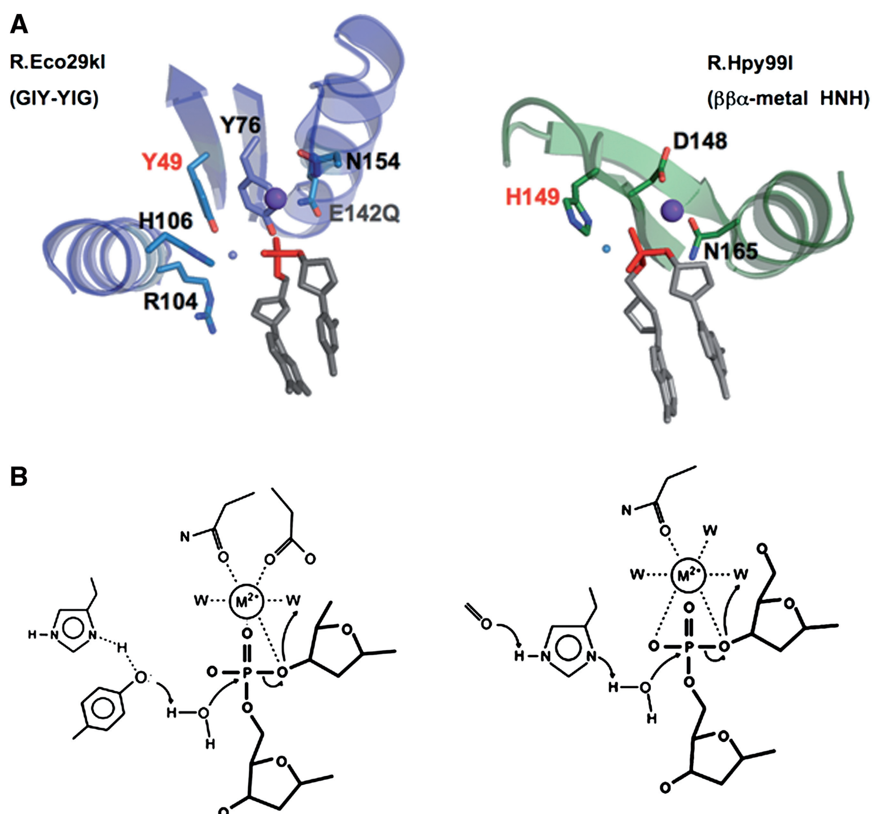

Figure 3. Comparison of two endonuclease motifs that are heavily represented by monomeric endonuclease domains with significant nicking activity. (**A**) On the left is the DNA-bound catalytic site of the GIY-YIG endonuclease R.Eco29kI; on the right is the DNA-bound catalytic site of the HNH endonuclease R.Hpy99I. The proposed mechanism and overall catalytic architecture is largely the same (with the exception of the substitution of a tyrosine for a histidine for general base), but are grafted onto two completely separate protein fold topologies. The large purple spheres in the structures are bound magnesium ions; the small light blue spheres are the nucleophilic water molecules. (**B**) Postulated mechanisms of phosphodiester bond hydrolysis by the GIY-YIG (left) and HNH (right) nuclease active sites.

Naturally occurring nicking HNH HEases

The HNH enzymes I-HmuI and I-BasI (isoschizomers encoded within *Bacillus* phage SPO1 and *Bastille*, respectively) have been studied using a combination of mutagenic, biochemical and structural analyses (27,79,82–85). These proteins bind to the same DNA sequences and nick the coding strand at the identical position, 3 nt downstream of the eventual intron insertion site: AACGCTCAGCAATTCCCACGT (none/−18) (84,85).

Although I-HmuI and I-BsaI are isoschizomers, the relative specificities of I-HmuI and I-BasI are slightly different (27): I-HmuI nicks both intron-minus and intron-plus target sequences from its own genome (*SPO1* phage) with a preference for the intron-minus substrate. In contrast, I-BasI is able to nick the intron-minus target from both host genomes, although the actual homing sites are very difference in sequence (27). In both cases, it is assumed that the eventual resolution of the enzyme-induced nick, which leads to intron homing, may require DNA replication in order to convert the initial nick into a double-strand break. However, as discussed below, recent studies with engineered LAGLIDADG HEases may indicate that a single-strand break can be recombinogenic under certain conditions.

Sequential cleavage of GIY-YIG HEases

The GIY-YIG HEase such as I-TevI possesses one catalytic site per protein chain and largely appears to generate ds breaks. This requires the enzyme to form a transient dimer to create two adjacent active sites (as in Type IIS FokI) or reposition a single-active site on each strand during a single-binding event with its cognate target site (86,87) (similar to Type IIS BfiI). Both I-TevI and I-BmoI have been reported to undergo a sequential cleavage mechanism where the bottom strand is nicked first, bending the DNA at the nicked site to facilitate the nicking of the top strand (88,89). By using a bioinformatics and mutagenesis approach to identify highly co-evolved residues, a residue that is involved in the second-cleavage event has been identified in I-BmoI—mutation of Ile70 to Asn leads to the accumulation of a nicked intermediate, which is eventually converted to ds cleavage product over time (90). More research is needed to understand the role of this residue and the mechanism of sequential cleavage of GIY-YIG HEases in general.

LAGLIDADG HEases and nicking variants

LAGLIDADG endonucleases contain two similar core folds of mixed α/β topology, with their namesake

sequence motifs forming two α-helices that are packed together at the domain or subunit interface, where they each contribute a catalytic residue to an active site (69). Two distinct subfamilies of LAGLIDADG endonucleases have been identified. Those that contain a single LAGLIDADG motif per protein chain (such as I-CreI) form homodimers that recognize palindromic and pseudo-palindromic target sites. In contrast, enzymes that contain two motifs per protein chain form asymmetric monomers that recognize correspondingly asymmetric homing sites, with each domain cleaving each of the two strands (Figure 4A). Both subclasses of these endonucleases display a canonical two-metal ion mechanism for phosphodiester hydrolysis, with a metal-bound water molecule in each active site appropriately positioned for an in-line hydrolytic attack on the scissile phosphate group (Figure 4B) (91). As well, the two active sites found in the LAGLIDADG enzymes are very closely juxtaposed, and in many cases appear to physically share one bound catalytic metal ion (out of three total) (91–93).

Members of the monomeric LAGLIDADG enzyme subfamily often display a significant asymmetry in their rates of cleavage of individual DNA strands within their target sites. The algal endonuclease I-CpaII preferentially nicks the bottom strand of its target site at very low magnesium (94). The yeast HEase I-SceI has a higher binding affinity for the 3′-DNA half-site, leading to the accumulation of nicked intermediates during the cleavage reaction (95). The archaeal endonuclease I-DmoI and the fungal enzyme I-AniI have both been shown to preferentially cleave the coding strand of their target sites (96,97).

The natural tendency of monomeric LAGLIDADG enzymes to cleave one DNA strand more efficiently suggests that cleavage of the two strands is sequential. A wealth of crystallographic and biochemical studies has revealed information about the cleavage mechanism and active-site architecture of LAGLIDADG enzymes. In stark contrast to other endonuclease families, a well-ordered network of solvent molecules, that surround the nucleophilic water moiety, is distributed in a large pocket surrounding the DNA scissile phosphate group. These ordered solvent molecules are positioned and coordinated by several basic residues that line the periphery of the active site. These residues, which appear to be involved in solvent-mediated interactions and proton transfer, are highly diverse chemically and structurally. Mutation of the peripheral residues Gln47 and Lys98 of the homodimeric I-CreI LAGLIDADG endonuclease causes significant reductions in catalytic efficiency with little effect on either overall substrate-binding affinity or the structure of the enzyme-DNA complex (98). Similar mutations of the monomeric I-AniI result in a variety of catalytic effects, ranging from complete inactivation, to accentuation of their natural tendency to cleave one DNA strand more rapidly than its counterpart, to the formation of highly active nicking enzymes (Figure 4C) (99). Mutation of Lys227 to Met results in high nicking activity, but no significant ds cleavage activity within the experiment conditions (Figure 4C) (99).

The crystal structures of I-SceI with uncleaved and bottom-nicked substrates suggest a sequential cleavage mechanism that involves the recruitment of a fourth metal ion for second-strand cleavage (100). Lys223 (corresponds to Lys227 of I-AniI) in the C-terminal catalytic site coordinates the nucleophilic water molecule for the attack of the scissile phosphate of the bottom strand (101). Lys122 is involved in the water network that coordinates the fourth metal ion and the second nucleophilic water molecule for the cleavage of the top strand (93,100). Mutations of Lys122 and Lys223 to Ile result in changes of the rate constant of the individual nicking steps such that the K122I mutation suppresses top-strand cleavage whereas the K223I mutation suppresses bottom-strand cleavage, resulting in a bottom-strand nicking and a top-strand nicking variant, respectively (Figure 5) (101). Both mutants, however, generate ds break products over extended reaction time.

Interactions at the interface of the two catalytic domains also seem to play a role in the ds cleavage activity of LAGLIDADG HEases. Mutation of the domain interface residues of I-DmoI results in a preference to nick than making ds breaks (102). Given the loosely conserved nature of the peripheral residues in LAGLIDADG active sites and the use of a solvent network for the coordination and deprotonation of the nucleophile rather than a single explicit protein side chain, it seems clear that LAGLIDADG active sites display variable amounts of mechanistic redundancy in their ability to carry out acid/base catalysis. Nonetheless, the overall mutagenic strategy to shift the activity of LAGLIDADG enzymes from DNA cleavage to DNA nicking by mutation of the peripheral and domain interface residues may be generalizable across a wide range of diverse protein scaffolds.

SELECTION METHODS FOR ISOLATION OF NICKING ACTIVITIES

Enzymes that possess two readily identifiable catalytic sites, such as LAGLIDADG HEases, Mva1269I and BtsCI and their group members, are easier subjects for engineering nicking activity. When little is known about the identity of the functional residues, a good selection system is perhaps the most important part of the engineering and selection of nicking activity (and any desired enzymatic properties). Nicked plasmid has a very different electrophoretic mobility than the closed circular ones, and *E. coli* appears to tolerate low-level expression of some NEases in the absence of the cognate methyltransferase (MTase). Nicking variants of SapI have been selected by making use of these properties (53). A SapI site and a BbvCI site are added to the expression plasmid vector such that sequential nicking by top-strand nicking variants of SapI and Nb.BbvCI will result in a 10-nt 3′-recessive end for ligating to an adaptor for amplification in PCR (Figure 6). Similarly, sequential nicking by bottom-strand nicking variants of SapI and Nt.BbvCI will result in a 4-base 5′-recessive end available for ligation to an adaptor for amplification by

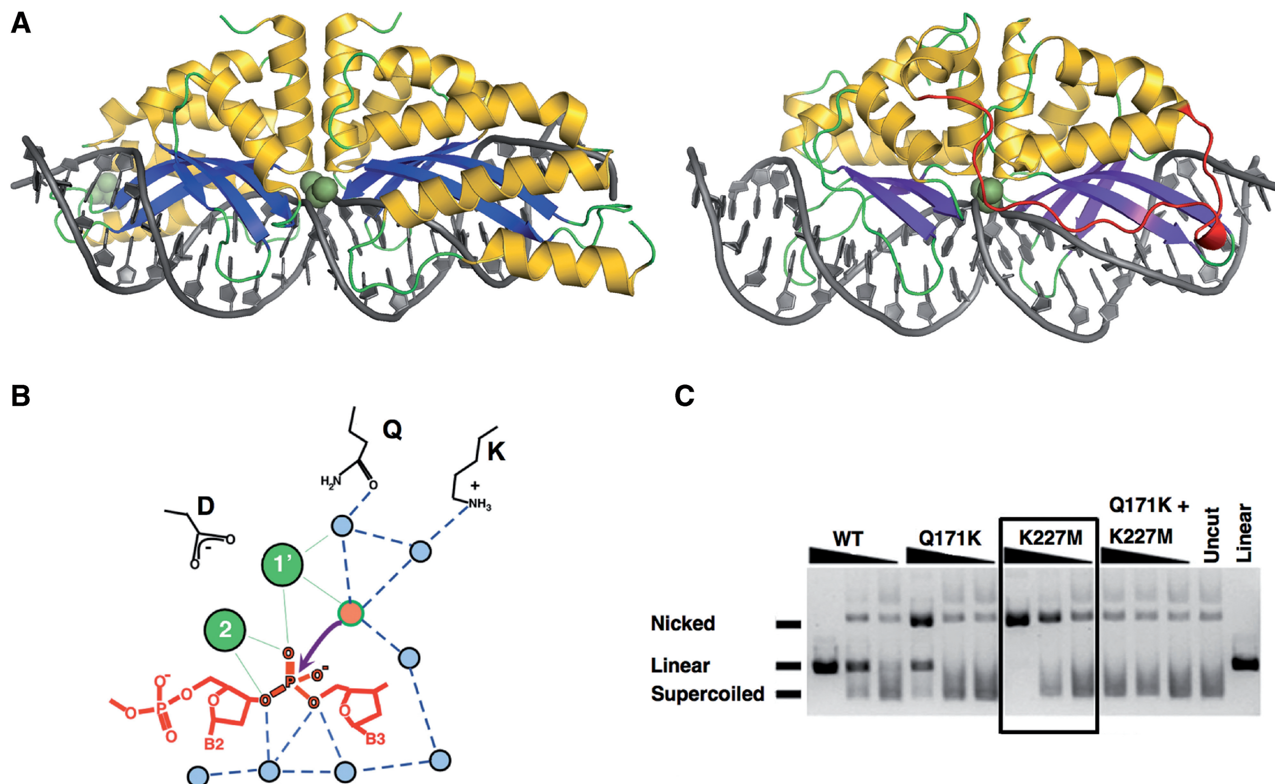

Figure 4. LAGLIDADG endonucleases and engineered nicking I-AniI variant. (**A**) Members of the LAGLIDADG endonucleases exist both as homodimers that recognize, bind and cleave palindromic and near-palindromic DNA-target sequences (such as I-CreI, left side) and monomeric, pseudo-symmetric monomers that recognize and cleave asymmetric DNA-target sequences (such as I-AniI, right side). These latter monomeric endonucleases often display considerable asymmetry in the relative rates of top versus bottom strand cleavage. (**B**) A canonical LAGLIDADG catalytic site positions bound metal ions (green spheres labeled 1′ and 2) within direct contact distance to the oxygen atoms of the scissile phosphate, and then surrounds the nucleophilic water (red sphere) with a shell of ordered water molecules that are positioned, in part, through contacts with side chains of peripheral residues. While the identity of those side chains are only moderately conserved, a glutamine residue and/or a lysine residue are usually found to be involved in these contacts. (**C**) Mutation of these residues in I-AniI generate a variety of behaviors ranging from partial nicking activity (Q171K), to complete conversion to strong nicking activity (K227M; boxed lanes) to a complete loss of endonuclease activity (Q171M/K227M double mutation). Shown on the gel for comparison is activity on the same plasmid substrate by wild-type (WT) I-AniI (which generates a ds break).

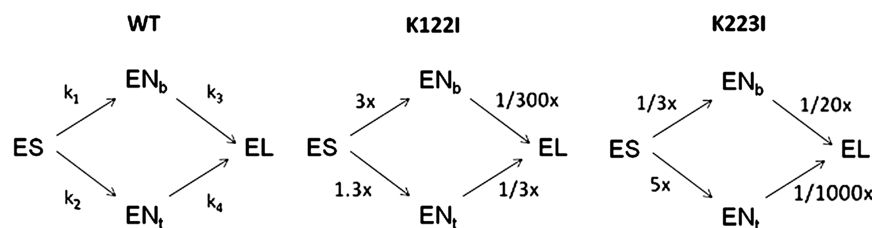

Figure 5. Rate constants of sequential cleavage of I-SceI WT, K122I and K223I mutant. The enzyme-substrate complex (ES) is converted to the enzyme-bottom-nicked intermediate (EN_b) at a rate constant of k_1 and then to the enzyme-linear product (EL) at a rate constant of k_3. The enzyme-substrate can also be converted to the enzyme-top-nicked intermediate (EN_t) at a rate constant of k_2 and then to the enzyme-linear product (EL) at a rate constant of k_4. The rate constants for mutants K122I and K223I are represented as folds compared to those of the WT enzyme.

PCR. A randomly mutated library of the *sapIR* gene was subjected to transformation of the *E. coli* host cells without the protection of the cognate MTases. The expression plasmid DNA was nicked by the nicking SapI variants *in vivo*. The nicked plasmid DNA was then isolated, gel-purified and subjected to Nt.BbvCI or Nb.BbvCI cleavage. After ligation with the appropriate adaptor, the mutated ORF was amplified and re-cloned into host cells with the cognate MTases. Nicking activity of the SapI mutants were then assayed. Through this selection process, residues that are important for the cleavage of each strand of the ds-target sequence were identified, and top- and bottom-strand nicking-SapI variants were created (53). This selection method should be generally applicable for engineering nicking activity of other restriction/HEases.

APPLICATION OF NEases

By cleaving a specific strand of dsDNA and leaving the opposite strand intact, NEases open the door to applications that cannot be achieved by ds breaking REases. In conjunction with strand-displacing or nick-translating DNA polymerases, which extend the 3′-recessive end of the nicked site and regenerate the nicking site itself, NEases can amplify specific segments of sample DNA without the need of thermocycling as in EXPAR and related amplification schemes. The same methodology can also be used to introduce fluorescently labeled nucleotides into long/chromosomal DNA *in vitro* for visualization. The implementation of these sample and/or signal amplification schemes can lead to simple but sensitive and specific methods for the detection of target organisms in the field. Finally, NEases can create nicked DNA for *in vivo* gene targeting by inducing homologous recombination. Gaps can be introduced into long DNA molecules for the study of DNA repair and recombination, as well as DNA translocases activity.

The applications described below were designed based on the sequence and strand specificity of the NEases being used. Nonetheless, these applications are general and other NEases can be used as long as the corresponding nicking site is properly located within the DNA.

Amplification of DNA signal

NEases generate nicked DNA with 3′-recessive ends available for extension by strand-displacing DNA polymerases. The nicking site can then be regenerated for the next round of nicking extension without the need of thermocycling. The amplified DNA signal can then be quantified by a variety of detection methods: capillary electrophoresis (103); reverse-phase liquid chromatography (104,105); real-time detection of the amplified DNA using DNA-binding fluorescent dyes and real-time thermocycler (106); or annealing the amplified DNA to complementary sequences conjugated to colloidal gold

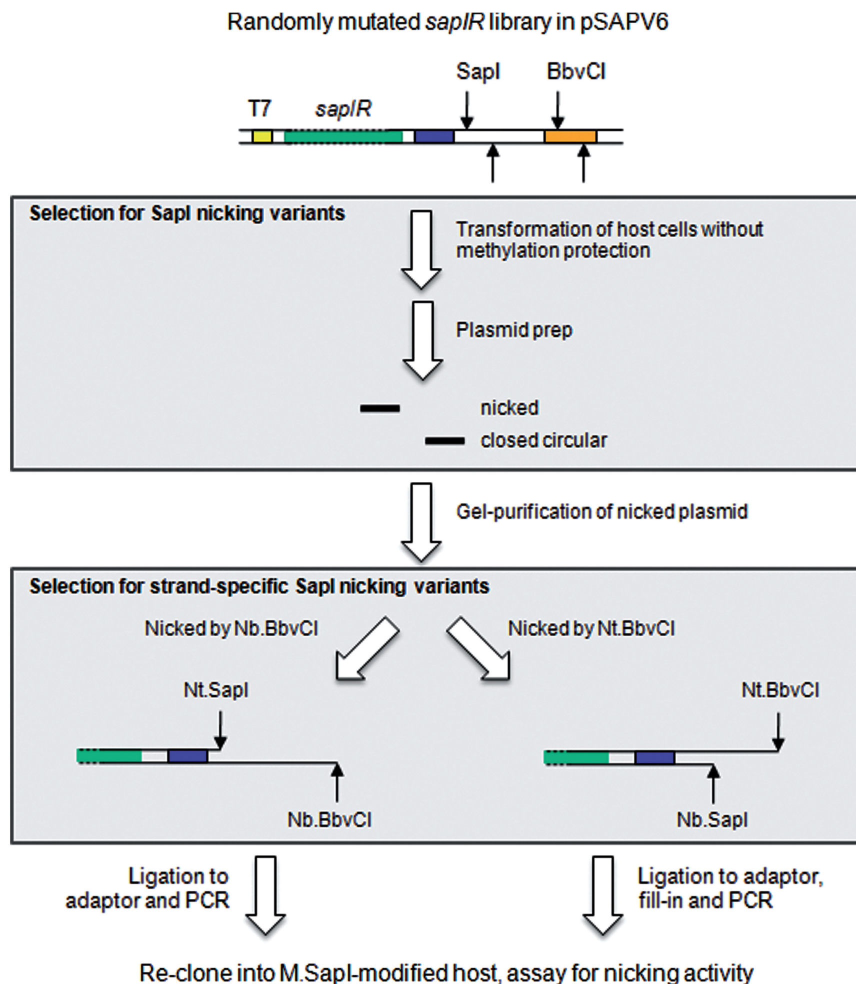

Figure 6. Selection for strand-specific SapI nicking variants. The T7 expression vector for *sapIR* (pSAPV6) was modified such that nicking by Nt.SapI/Nb.BbvCI or Nb.SapI/Nt.BbvCI resulted in overhangs that can be ligated to adaptors for PCR amplification. A randomly mutated pSAPV6 *sapIR* library was isolated from an *E. coli* host without the protection from the cognate MTases so that the plasmid vectors that expressed nicking variants of SapI were nicked *in vivo*. The nicked plasmid vector was gel-purified and further nicked by Nt.BbvCI or Nb.BbvCI, followed by ligation to the appropriate ss adaptor (and fill-in for the Nb.SapI/Nt.BbvCI reaction). After PCR amplification, the mutated *sapIR* gene was re-cloned into host cells pre-modified by SapI MTases. Nicking activity was then assayed from the crude extract of induced cultures. T7 promoter is in yellow color, the *sapIR* gene in green, the SapI recognition site in blue and BbvCI recognition site in orange. The cleavage sites of SapI and BbvCI and their nicking variants are indicated by arrows.

particles (107). Several schemes have been developed to amplify specific DNA segments.

In Exponential Amplification Reaction (EXPAR), a short segment of the target analyte DNA (trigger) is amplified using Nt.BstNBI and the strand-displacing Bst DNA polymerase in a linear or exponential mode. An amplification template that incorporates two copies of the sequence complementary to the trigger separated by a BstNBI site circumvents the need of a nicking site in the trigger sequence (Figure 7A). The exponential mode has been shown to achieve a >100-fold amplification in under 5 min (105–108). EXPAR has also been applied to the construction of DNA-based logic gates and circuits (X. L. Wang, personal communication).

NESA is designed to amplify ss DNA-target sequences. A fluorescently labeled ss-oligonucleotide probe is annealed to the heat-denatured genomic DNA target or other ss DNA sources through prior multiple displacement amplification (MDA) (103,109). Sequence- and strand-specific nicking was carried out by Nt.AlwNI at 58°C such that the nicked probe dissociates from the template and is subsequently detected by capillary electrophoresis. NESA has been shown to efficiently detect single-copy target from 10 fmol of DNA and distinguish DNA isolated from similar *Bacillus* species (103). An extended version of NESA incorporates rolling circle amplification of the target DNA and amplification of signal by the use of a molecular beacon (Figure 7B) (110). The combined amplification has been shown to bring about a detection limit in the femtomolar range (110). NESA has also been adapted to detect *Streptococcus pyogenes* (111). A signal amplification method (without target sequence amplification) using Nt.BspD6I and a molecular beacon has also been reported (112). A NEases mediated DNA Amplification scheme amplifies random ss-DNA fragments through the continuous nicking and extension of nicked DNA by the frequent NEase Nt.CviPII and Bst DNA–DNA polymerase (23). The ss-DNA generated can be used as the template for NESA.

The primer generation-rolling circle amplification (PG-RCA) scheme makes use of a circular ss DNA probe that contains the complementary sequence of the target DNA and a Nb.BsmI site (Figure 7C). Replication and nicking of the target sequence by Vent (exo^-) DNA polymerase and Nb.BsmI, respectively, has been shown to amplify the target DNA sequence from ~60 molecules of *Listeria monocytogenes* genomic DNA (104).

A detection scheme for human herpes virus by extension at the 3′-ends of vicinal nicks on the virus DNA by Nb.BsmI and Nb.BsrDI has been reported. In this scheme, the NEases generate gaps of 15–25 nt on the target virus DNA. After annealing by gap-specific oligonucleotides with a 3′-flap and immobilization, RCA is used to amplify the signal (113). Such a detection scheme has been shown to be sensitive and specific, and is expected to be applicable to other viruses such as EBV and HCMV.

In addition to DNA, sequence-specific detection of RNA has also been reported. In this detection scheme, a hairpin DNA probe is designed to open up when annealed to the target RNA sequence. The opening of the probe exposes a sequence complementary to a quenched fluorescent probe with a 5′-fluorophore and a 3′-quencher. Annealing with the quenched probe generates a ds-BbvCI site such that Nb.BbvCI nicks the strand containing quenched fluorescent probe and releases the fluorophore from being quenched (114). This RNA detection scheme has been shown to be quantitative and not requiring purification of the RNA, and should be applicable to the detection of ss DNA.

Chromosomal DNA bar coding (nanocoding)

Sequence-specific nicks generated by NEases can be fluorescently labeled by incorporating fluorescent nucleotides using a nick-translating DNA polymerase. The labeled DNA is then stretched through a nanoslit (115,116), on a modified glass surface (117) or on an optical mapping surface (54) (Figure 8). Single DNA molecule BbvCI maps of BAC (115), λ DNA and human adenovirus type 2 DNA (117), BsmI maps of human rhinovirus 14 DNA and human rhinovirus 89 (117) and BspQI maps of T7 phage DNA (54) have been produced. This NEase-based nanocoding methodology can become a quick and highly sensitive method to study SNP within nicking sites. For example, the nicking variants of BtsCI (a neoschizomer of FokI) should be useful in nanocoding the FokI polymorphism of the vitamin D receptor in cancer risk assessment (118,119).

Gene targeting based on homologous recombination

The use of natural and designer zinc-finger nucleases (ZFNs) for targeted *in vivo* deletion and insertion through generation of ds breaks has been well studied and exhibits big promise of practical use in gene targeting (120–122). Recently, strand-specific nicks generated by NEases have also been shown to induce homologous recombination and hence gene insertion (99,123,124). Adeno-associated virus (AAV) Rep78 and Rep68 proteins that introduce a sequence-specific nick in locus AAVS1 on human chromosome 19 have been shown to induce insertion of a functional 4.1-kb sequence into HeLa cells through homologous recombination (124). The nicking variant of I-AniI has been shown to stimulate homologous recombination in human T293 cells at an efficiency approximately one-fourth that of the wild-type ds-breaking I-AniI enzyme, both when measuring transient recombination in *cis*, and in assays of recombination of a chromosomal target gene in *trans* (99). A more recent study using an adenoviral delivery system showed that the I-AniI nickase-induced recombination events do not appear to require subsequent conversion to ds breaks (M. J. Metzger, manuscript in preparation). These nicking enzyme-induced recombination events do not result in significant non-homologous end-joining (NHEJ) events and appear to greatly reduce overall cell toxicity when the protein is expressed.

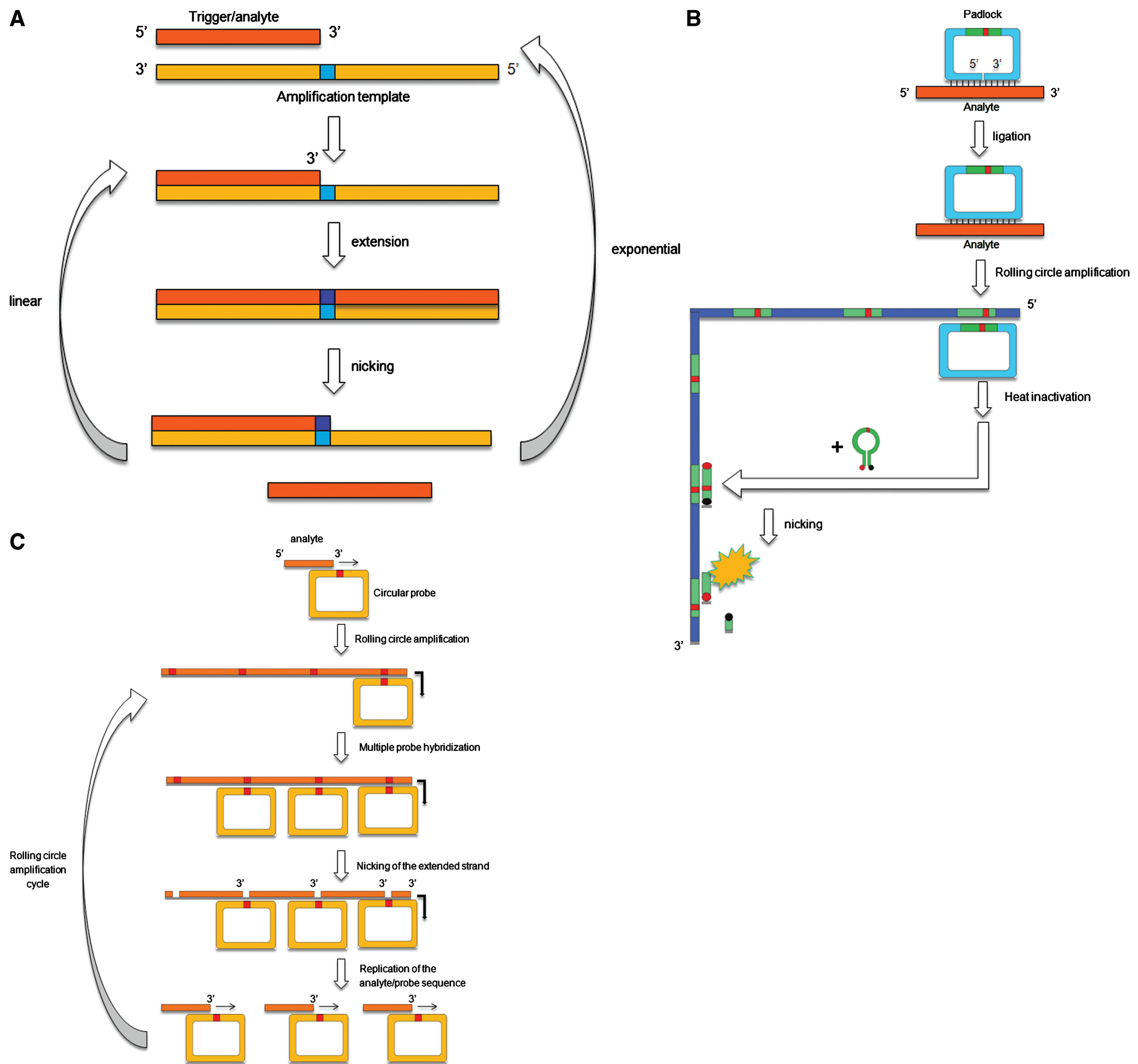

Figure 7. NEase-dependent amplification. (**A**) EXPAR. In EXPAR, the target analyte DNA (trigger; dark orange) anneals to a synthetic amplification template that contains two copies of the sequence complementary to the target sequence (light orange) separated by a sequence complementary to the BstNBI site (light blue). At an elevated temperature (60°C) Bst DNA polymerase large fragment extends at the 3′-end of the trigger DNA and creates a ds-BtsNBI site (dark blue/light blue) followed by an extra copy of the trigger sequence. The thermostable Nt.BstNBI NEase then nicks the top strand of the BstNBI site. By the combined action of the elevated temperature and the strand-displacing activity of Bst DNA polymerase, the newly synthesized trigger sequence is detached from the amplification template and allows the next cycle of extension to proceed in the linear amplification scheme. In the exponential amplification protocol, the extended-cleaved triggers are removed from the amplification template so that the newly synthesized triggers can start a new cycle of extension-nicking. (**B**) In the extended NEase signal amplification (NESA) scheme that involves multiple temperatures and steps, the initial target-analyte sequence (dark orange) was first amplified by annealing to a padlock DNA designed to close at the target sequence (light blue), followed by ligation of the closed padlock DNA and rolling-circle amplification. The resulting concantemers consist of a tandem repeat of the sequence complementary to a molecular beacon with a BstNBI site (red). The molecular beacon is then allowed to anneal to the concantemers, creating a ds-BstNBI site where Nt.BstNBI nicks the molecular beacon and emits a fluorescent signal. (**C**) Primer generation-rolling circle amplification (PG-RCA). Similar to EXPAR, the amplification template (light orange) consists of a nicking site (red) sandwiched between two copies of the sequence complementary to the target analyte (dark orange). In this particular example, a BsmI site is used. In PG-RCA, the amplification template is made circular for rolling circle amplification. After annealing of the analyte DNA, Vent (exo^{-}) DNA polymerase extends the 3′-end and creates a concantemer of BsmI sites sandwiched between the analyte sequences. Multiple annealing of the circular probe to the concatemers creates ds-strand BsmI sites. Nb.BsmI then nicks the concantemer strands and generates multiple copies of the analyte for the next round of rolling circle amplification.

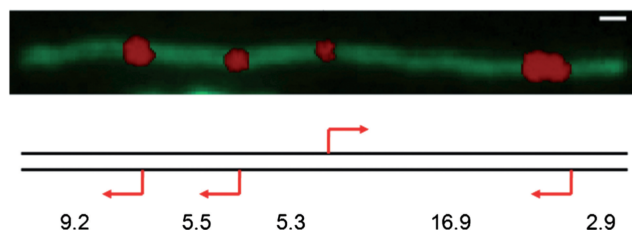

Figure 8. NEases dependent nanocoding. (Upper panel) A T7 bacteriophage DNA molecule was nicked by Nt.BspQI. Alexa Fluor 647-dUTP was incorporated by the nick-translating *E. coli* DNA polymerase I. The DNA is stained by YOYO-1 (green) and the nick-translated sites (red) are revealed by FRET signal from the Alexa fluorophore. The white bar indicates 1 μm. (Lower panel) Scale for the T7 bacteriophage DNA in kilobites. The red arrows indicate the orientation of the BspQI sites and the direction of nick translation.

Generation of nicked or gapped substrate for the study of DNA recombination and translocation

Recombination initiated by nicks or gaps as opposed to ds breaks has not been well-studied largely due to a paucity of reagents that reliably nick DNA in a site-specific fashion. The availability of nicking variants of Type IIS and Type IIA REases and HEases makes it possible to generate nicked or gapped DNA for mechanistic analyses of DNA repair and recombination. By designing two nicking sites close enough on the same strand, nicking of the sites creates an intervening sequence that melts away. Nicked and gapped DNA has been used in the study of DNA mismatch repair mechanism in a mammalian model system (125,126).

Gapped DNA has also become imperative in the study of the helicase domain of EcoP124I. The ability to generate long-DNA molecules with a ss-DNA gap has facilitated the discrimination of the DNA translocation activity from DNA unwinding activity of the Type I REase (127). The availability of top- and bottom-nicking variants of BbvCI makes it possible to generate identical DNA molecules with a gap on either the 5′–3′ or 3′–5′ strand for the determination of strand dependence of gap bypass of EcoP124I (127). These assays are believed to be applicable to other helicases and translocases.

Other applications

The generation of extendable 3′-recessive ends with the potential of creating long-single-stranded downstream sequence on the opposite strand is a great asset of NEases. By making use of this property, a new sample preparation scheme has been reported for pyrosequencing that allows quantitative genotyping of SNP and gene expression analysis (128). Also, the 3′-end of the nicked strand directs *E. coli* exonuclease III (Exo III) to degrade the nicked strand specifically but leaving the un-nicked strand intact. For example, pairs of top- and bottom-strand nicking enzymes that recognize the same target site, such as Nt.BspQI/Nb.BspQI and Nt.BtsCI/Nb.BtsCI can generate either Watson or Crick strand from the same plasmid construct in conjunction with Exo III (54). A heteroduplex reporter plasmid for mismatch repair has been created by this means (129). Gapped DNA generated by multiple nicks on the same DNA strand allows incorporation of modified nucleotides anywhere within a long-DNA molecules (130). By putting a nicking site and a restriction site in tandem, long-sticky ends can be created for gene cloning without the need of DNA ligase (131).

As a neoschizomer to FokI, sequential cleavage of Nt.BtsCI or Nb.BtsCI with wild-type FokI can generate 11-nt 5′ or 9-nt 3′ overhangs, respectively. These overhangs can be ligated to oligonucleotide duplexes that carry single-strand regions complementary to the overhangs for cloning and labeling (51). The 11-nt 3′-recessive end can also be filled in with fluorescently labeled nucleotides efficiently for the generation of fluorescent probes (51).

The single-strand nicking activity, combined with the annealing-denaturing with a complementary sequence, has been exploited to create a molecular motor. A nanomotor device that transports a single-stranded cargo DNA along a ds-DNA track using Nb.BbvCI has been reported. In this 'burnt-bridge' strategy, a double-strand DNA track is lined with equally spaced single-stranded stators with sequence complementary to the cargo DNA. Annealing of the cargo DNA to the stators creates a ds-BbvCI site (132). Nb.BbvCI nicks the stator strand and releases the cargo DNA for annealing to the next stator along the track. The transport is halted by a single-base mismatch at the BbvCI site of the last stator such that the cargo-stator is not nicked.

FUTURE CHALLENGES AND OPPORTUNITIES

So much has been learned about the cleavage mechanisms of Types IIS and IIA REases and LAGLIDADG HEases, yet there are still many unknowns. Class III enzymes, for example, have a small subunit that exhibits cleavage activity only in the presence of the large subunit. It remains unclear how this switch happens. Similarly, it is still unclear why the presence of both subunits of the nicking variants of the Class IV BbvCI (a Type IIA REase) is required for nicking to occur, even though only one of them is catalytically active. More research is needed to understand the intricate modulation of cleavage activity of these enzymes.

Just like the ds-breaking parent enzymes, some engineered NEases exhibit star activity. It will be interesting to see if the mutations that bring about low star activity for the parent REases can also lower the star activity of the corresponding NEases.

Nanocoding in its current version can only achieve 0.5-kb resolution between two labeled nicked sites. Improvement of fluorescence imaging techniques can achieve higher resolution (<100 bp) for more precise determination of genetic mutations and multiple nicking site profiling.

The fusion of the methyl-binding domain (MBD) of human MeCP2 (133–135) with the DNA-cleavage domain of BmrI (Class V) or FokI (Class I) had been shown to cleavage 5mCpG sequences preferentially over unmodified DNA (136). This study suggests that it is

possible to generate NEases specific to 5mCpG by tethering the MBD of 5mCpG-binding proteins such as hMeCP2 and MBD2b (137–139) to an active DNA-nicking domain, such as those from Class III enzymes (i.e. Nb.BtsI, Nb.BsrDI). The availability of such 5mCpG-specific NEases will be of tremendous value in the study of epigenetics as reagents for signal amplification and nanocoding. The nicking domain of Class III enzymes may also be tethered to DNA-recognition domains, such as zinc fingers and transcription factors, to create custom-specific NEases that can make nicks close enough *in vivo* to create ds breaks for gene targeting.

CONCLUDING REMARKS

Canonical Type IIP REases combine their DNA-binding and DNA-cleavage residues into one catalytic domain. Obligatory homodimerization results in cleavage of both DNA strands of ds-DNA at the equivalent positions of palindromic sequences. Type II REases that recognize asymmetric sequences, Types IIS and IIA REases, in particular, use a variety of strategies to recognize non-palindromic target sites and cleave at variable distances, as revealed by intensive research in the last decade. These include the use of two catalytic sites to cleave each of the strands of ds-DNA and cleavage of each of the strands sequentially using the same catalytic site. The study of the cleavage mechanism of the former case has resulted in useful NEases. LAGLIDADG HEases, such as I-SceI and I-AniI, also use two catalytic sites for cleavage of the each of two strands of ds-DNA target. Studies of their cleavage mechanisms have also resulted in nicking variants useful for gene targeting among others.

NEases open a new door of application not achievable by REases that make ds breaks. The nicks created by NEases can be directly fluorescently labeled by nick translation, allowing for nanocoding of chromosomal and virus DNA. Gaps created by multiple nicks can be used to incorporate modified nucleotides into long DNA molecules. Combined with strand-displacing DNA polymerases, such as Bst DNA polymerase and ϕ-29 DNA polymerase, the nick generates a 3′-extendable end for the amplification of target DNA in a sequence-specific or non-specific manner. Lastly, NEases can also generate sequence-specific nicks and gaps for the study of DNA repair and recombination and DNA translocation.

NEases have only been commercially available in the last few years and the selection of sequence specificity is still limited. However, with the information already available on the cleavage mechanisms of various classes of Type IIS and Type IIA REases and LAGLIDADG HEases, more engineering effort will result in a wider selection of NEases specificity for a broader application in molecular medicine and biotechnology.

ACKNOWLEDGEMENTS

We thank Richard Roberts, Bill Jack and Frederick Gimble for commenting on the article and Geoff Wilson for discussions, Don Comb and Jim Ellard for support and encouragement.

FUNDING

National Institutes of Health (grants R01GM49857 and NIH RL1 CA833133 to Stoddard laboratory). Funding for open access charge: New England Biolabs, Inc.

Conflict of interest statement. None declared.

REFERENCES

1. Li,H., Stampfer,M.J., Hollis,J.B., Mucci,L.A., Gaziano,J.M., Hunter,D., Giovannucci,E.L. and Ma,J. (2007) A prospective study of plasma vitamin D metabolites, vitamin D receptor polymorphisms, and prostate cancer. *PLoS Med.*, **4**, e103.
2. Chang,H.W., Yang,C.H., Chang,P.L., Cheng,Y.H. and Chuang,L.Y. (2006) SNP-RFLPing: restriction enzyme mining for SNPs in genomes. *BMC Genomics*, **7**, 30.
3. Zilberman,D. and Henikoff,S. (2007) Genome-wide analysis of DNA methylation patterns. *Development*, **134**, 3959–3965.
4. Nielsen,K.L. (2008) DeepSAGE: higher sensitivity and multiplexing of samples using a simpler experimental protocol. *Methods Mol. Biol.*, **387**, 81–94.
5. Porter,D., Yao,J. and Polyak,K. (2006) SAGE and related approaches for cancer target identification. *Drug Discov. Today*, **11**, 110–118.
6. Sorber,K., Chiu,C., Webster,D., Dimon,M., Ruby,J.G., Hekele,A. and DeRisi,J.L. (2008) The long march: a sample preparation technique that enhances contig length and coverage by high-throughput short-read sequencing. *PLoS One*, **3**, e3495.
7. Kim,S., Kim,Y., Liang,T., Sinsheimer,J.S. and Chow,S.A. (2006) A high-throughput method for cloning and sequencing human immunodeficiency virus type 1 integration sites. *J. Virol.*, **80**, 11313–11321.
8. Luria,S.E. and Human,M.L. (1952) A nonhereditary, host-induced variation of bacterial viruses. *J. Bacteriol.*, **64**, 557–569.
9. Bertani,G. and Weigle,J.J. (1953) Host controlled variation in bacterial viruses. *J. Bacteriol.*, **65**, 113–121.
10. Sharp,P.A., Sugden,B. and Sambrook,J. (1973) Detection of two restriction endonuclease activities in Haemophilus parainfluenzae using analytical agarose–ethidium bromide electrophoresis. *Biochemistry*, **12**, 3055–3063.
11. Garfin,D.E. and Goodman,H.M. (1974) Nucleotide sequences at the cleavage sites of two restriction endonucleases from Hemophilus parainfluenzae. *Biochem. Biophys. Res. Commun.*, **59**, 108–116.
12. Smith,H.O. and Wilcox,K.W. (1970) A restriction enzyme from Hemophilus influenzae. I. Purification and general properties. *J. Mol. Biol.*, **51**, 379–391.
13. Goldstein,L., Thomas,M. and Davis,R.W. (1975) EcoRI endonuclease cleavage map of bacteriophage P4-DNA. *Virology*, **66**, 420–427.
14. von Gabain,A., Hayward,G.S. and Bujard,H. (1976) Physical mapping of the HindIII, EcoRI, Sal and Sma restriction endonuclease cleavage fragments from bacteriophage T5 DNA. *Mol. Gen Genet.*, **143**, 279–290.
15. Roberts,R.J. (2005) How restriction enzymes became the workhorses of molecular biology. *Proc. Natl Acad. Sci. USA*, **102**, 5905–5908.
16. Roberts,R.J., Vincze,T., Posfai,J. and Macelis,D. (2010) REBASE–a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.*, **38**, D234–D236.
17. Roberts,R.J., Belfort,M., Bestor,T., Bhagwat,A.S., Bickle,T.A., Bitinaite,J., Blumenthal,R.M., Degtyarev,S., Dryden,D.T., Dybvig,K. *et al.* (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.*, **31**, 1805–1812.

18. Vanamee,E.S. and Aggarwal,A.K. (2004) In Messerschmidt,A., Bode,W. and Cygler,M. (eds), *Handbook of Metalloproteins*, Vol. 3. John Wiley & Sons, Ltd., Chichester, pp. 742–756.
19. Galburt,E.A. and Stoddard,B.L. (2002) Catalytic mechanisms of restriction and homing endonucleases. *Biochemistry*, **41**, 13851–13860.
20. Grigorescu,A., Horvath,M., Wilkosz,P.A., Chandrasekhar,K. and Rosenberg,J.M. (2004) In Pingoud,A. (ed.), *Restriction Endonucleases*, Vol. 14. Springer, New York, pp. 137–177.
21. Cowan,J.A. (2004) In Pingoud,A. (ed.), *Restriction Endonucleases*, Vol. 14. Springer, New York, pp. 339–360.
22. Horton,J.R., Blumenthal,R.M. and Cheng,X. (2004) In Pingoud,A. (ed.), *Restriction Endonucleases*. Springer, New York, pp. 316–392.
23. Chan,S.H., Zhu,Z., Van Etten,J.L. and Xu,S.Y. (2004) Cloning of CviPII nicking and modification system from chlorella virus NYs-1 and application of Nt.CviPII in random DNA amplification. *Nucleic Acids Res.*, **32**, 6187–6199.
24. Chan,S.H., Zhu,Z., Dunigan,D.D., Van Etten,J.L. and Xu,S.Y. (2006) Cloning of Nt.CviQII nicking endonuclease and its cognate methyltransferase: M.CviQII methylates AG sequences. *Protein Expr. Purif.*, **49**, 138–150.
25. Landthaler,M. and Shub,D.A. (2003) The nicking homing endonuclease I-BasI is encoded by a group I intron in the DNA polymerase gene of the Bacillus thuringiensis phage Bastille. *Nucleic Acids Res.*, **31**, 3071–3077.
26. Landthaler,M., Lau,N.C. and Shub,D.A. (2004) Group I intron homing in Bacillus phages SPO1 and SP82: a gene conversion event initiated by a nicking homing endonuclease. *J. Bacteriol.*, **186**, 4307–4314.
27. Landthaler,M., Shen,B.W., Stoddard,B.L. and Shub,D.A. (2006) I-BasI and I-HmuI: two phage intron-encoded endonucleases with homologous DNA recognition sequences but distinct DNA specificities. *J. Mol. Biol.*, **358**, 1137–1151.
28. Xu,S.Y., Zhu,Z., Zhang,P., Chan,S.H., Samuelson,J.C., Xiao,J., Ingalls,D. and Wilson,G.G. (2007) Discovery of natural nicking endonucleases Nb.BsrDI and Nb.BtsI and engineering of top-strand nicking variants from BsrDI and BtsI. *Nucleic Acids Res.*, **35**, 4608–4618.
29. Morgan,R.D., Calvet,C., Demeter,M., Agra,R. and Kong,H. (2000) Characterization of the specific DNA nicking activity of restriction endonuclease N.BstNBI. *Biol. Chem*, **381**, 1123–1125.
30. Higgins,L.S., Besnier,C. and Kong,H. (2001) The nicking endonuclease N.BstNBI is closely related to type IIs restriction endonucleases MlyI and PleI. *Nucleic Acids Res.*, **29**, 2492–2501.
31. Perevyazova,T.A., Rogulin,E.A., Zheleznaya,L.A. and Matvienko,N.I. (2003) Cloning and sequencing of the gene of site-specific nickase N.BspD6I. *Biochemistry*, **68**, 984–987.
32. Yunusova,A.K., Rogulin,E.A., Artyukh,R.I., Zheleznaya,L.A. and Matvienko,N.I. (2006) Nickase and a protein encoded by an open reading frame downstream from the nickase BspD6I gene form a restriction endonuclease complex. *Biochemistry*, **71**, 815–820.
33. Szybalski,W., Kim,S.C., Hasan,N. and Podhajska,A.J. (1991) Class-IIS restriction enzymes–a review. *Gene*, **100**, 13–26.
34. Wah,D.A., Hirsch,J.A., Dorner,L.F., Schildkraut,I. and Aggarwal,A.K. (1997) Structure of the multimodular endonuclease FokI bound to DNA. *Nature*, **388**, 97–100.
35. Wah,D.A., Bitinaite,J., Schildkraut,I. and Aggarwal,A.K. (1998) Structure of FokI has implications for DNA cleavage. *Proc. Natl Acad. Sci. USA*, **95**, 10564–10569.
36. Zaremba,M., Urbanke,C., Halford,S.E. and Siksnys,V. (2004) Generation of the BfiI restriction endonuclease from the fusion of a DNA recognition domain to a non-specific nuclease from the phospholipase D superfamily. *J. Mol. Biol.*, **336**, 81–92.
37. Grazulis,S., Manakova,E., Roessle,M., Bochtler,M., Tamulaitiene,G., Huber,R. and Siksnys,V. (2005) Structure of the metal-independent restriction enzyme BfiI reveals fusion of a specific DNA-binding domain with a nonspecific nuclease. *Proc. Natl Acad. Sci. USA*, **102**, 15797–15802.
38. Kriukiene,E. (2006) Domain organization and metal ion requirement of the Type IIS restriction endonuclease MnlI. *FEBS Lett.*, **580**, 6115–6122.
39. Bitinaite,J., Wah,D.A., Aggarwal,A.K. and Schildkraut,I. (1998) FokI dimerization is required for DNA cleavage. *Proc. Natl Acad. Sci. USA*, **95**, 10570–10575.
40. Vanamee,E.S., Santagata,S. and Aggarwal,A.K. (2001) FokI requires two specific DNA sites for cleavage. *J. Mol. Biol.*, **309**, 69–78.
41. Skowron,P., Kaczorowski,T., Tucholski,J. and Podhajska,A.J. (1993) Atypical DNA-binding properties of class-IIS restriction endonucleases: evidence for recognition of the cognate sequence by a FokI monomer. *Gene*, **125**, 1–10.
42. Bath,A.J., Milsom,S.E., Gormley,N.A. and Halford,S.E. (2002) Many type IIs restriction endonucleases interact with two recognition sites before cleaving DNA. *J. Biol. Chem.*, **277**, 4024–4033.
43. Sanders,K.L., Catto,L.E., Bellamy,S.R. and Halford,S.E. (2009) Targeting individual subunits of the FokI restriction endonuclease to specific DNA strands. *Nucleic Acids Res.*, **37**, 2105–2115.
44. Besnier,C.E. and Kong,H. (2001) Converting MlyI endonuclease into a nicking enzyme by changing its oligomerization state. *EMBO Rep.*, **2**, 782–786.
45. Xu,Y., Lunnen,K.D. and Kong,H. (2001) Engineering a nicking endonuclease N.AlwI by domain swapping. *Proc. Natl Acad. Sci. USA*, **98**, 12990–12995.
46. Armalyte,E., Bujnicki,J.M., Giedriene,J., Gasiunas,G., Kosinski,J. and Lubys,A. (2005) Mva1269I: a monomeric type IIS restriction endonuclease from Micrococcus varians with two EcoRI- and FokI-like catalytic domains. *J. Biol. Chem.*, **280**, 41584–41594.
47. Guan,C. and Kumar,S. (2005) A single catalytic domain of the junction-resolving enzyme T7 endonuclease I is a non-specific nicking endonuclease. *Nucleic Acids Res.*, **33**, 6225–6234.
48. Guan,C., Kumar,S., Kucera,R. and Ewel,A. (2004) Changing the enzymatic activity of T7 endonuclease by mutations at the beta-bridge site: alteration of substrate specificity profile and metal ion requirements by mutation distant from the catalytic domain. *Biochemistry*, **43**, 4313–4322.
49. Sasnauskas,G., Connolly,B.A., Halford,S.E. and Siksnys,V. (2007) Site-specific DNA transesterification catalyzed by a restriction enzyme. *Proc. Natl Acad. Sci. USA*, **104**, 2115–2120.
50. Sasnauskas,G., Halford,S.E. and Siksnys,V. (2003) How the BfiI restriction enzyme uses one active site to cut two DNA strands. *Proc. Natl Acad. Sci. USA*, **100**, 6410–6415.
51. Too,P.H., Zhu,Z., Chan,S.H. and Xu,S.Y. (2010) Engineering Nt.BtsCI and Nb.BtsCI nicking enzymes and applications in generating long overhangs. *Nucleic Acids Res.*, **38**, 1294–1303.
52. Zhu,Z., Samuelson,J.C., Zhou,J., Dore,A. and Xu,S.Y. (2004) Engineering strand-specific DNA nicking enzymes from the type IIS restriction endonucleases BsaI, BsmBI, and BsmAI. *J. Mol. Biol.*, **337**, 573–583.
53. Samuelson,J.C., Zhu,Z. and Xu,S.Y. (2004) The isolation of strand-specific nicking endonucleases from a randomized SapI expression library. *Nucleic Acids Res.*, **32**, 3661–3671.
54. Zhang,P., Too,P.H., Samuelson,J.C., Chan,S.H., Vincze,T., Doucette,S., Backstrom,S., Potamousis,K.D., Schramm,T.M., Forrest,D. *et al.* Engineering BspQI nicking enzymes and application of N.BspQI in DNA labeling and production of single-strand DNA. *Protein Expr. Purif.*, **69**, 226–234.
55. Kachalova,G.S., Rogulin,E.A., Yunusova,A.K., Artyukh,R.I., Perevyazova,T.A., Matvienko,N.I., Zheleznaya,L.A. and Bartunik,H.D. (2008) Structural analysis of the heterodimeric type IIS restriction endonuclease R.BspD6I acting as a complex between a monomeric site-specific nickase and a catalytic subunit. *J. Mol. Biol.*, **384**, 489–502.
56. Bellamy,S.R., Milsom,S.E., Scott,D.J., Daniels,L.E., Wilson,G.G. and Halford,S.E. (2005) Cleavage of individual DNA strands by the different subunits of the heterodimeric restriction endonuclease BbvCI. *J. Mol. Biol.*, **348**, 641–653.
57. Heiter,D.F., Lunnen,K.D. and Wilson,G.G. (2005) Site-specific DNA-nicking mutants of the heterodimeric restriction endonuclease R.BbvCI. *J. Mol. Biol.*, **348**, 631–640.
58. Stankevicius,K., Lubys,A., Timinskas,A., Vaitkevicius,D. and Janulaitis,A. (1998) Cloning and analysis of the four genes coding for Bpu10I restriction-modification enzymes. *Nucleic Acids Res.*, **26**, 1084–1091.

59. Stuckey,J.A. and Dixon,J.E. (1999) Crystal structure of a phospholipase D family member. *Nat. Struct. Biol.*, **6**, 278–284.
60. Lagunavicius,A., Sasnauskas,G., Halford,S.E. and Siksnys,V. (2003) The metal-independent type IIs restriction enzyme BfiI is a dimer that binds two DNA sites but has only one catalytic centre. *J. Mol. Biol.*, **326**, 1051–1064.
61. Sasnauskas,G., Zakrys,L., Zaremba,M., Cosstick,R., Gaynor,J.W., Halford,S.E. and Siksnys,V. (2010) A novel mechanism for the scission of double-stranded DNA: BfiI cuts both 3′-5′ and 5′-3′ strands by rotating a single active site. *Nucleic Acids Res.*, **38**, 2399–2410.
62. Jeltsch,A., Alves,J., Wolfes,H., Maass,G. and Pingoud,A. (1993) Substrate-assisted catalysis in the cleavage of DNA by the EcoRI and EcoRV restriction enzymes. *Proc. Natl Acad. Sci. USA*, **90**, 8499–8503.
63. Dujon,B. (1989) Group I introns as mobile genetic elements: facts and mechanistic speculations—a review. *Gene*, **82**, 91–114.
64. Lambowitz,A.M. and Belfort,M. (1993) Introns as mobile genetic elements. *Annu. Rev. Biochem.*, **62**, 587–622.
65. Belfort,M. and Perlman,P.S. (1995) Mechanisms of intron mobility. *J. Biol. Chem.*, **270**, 30237–30240.
66. Belfort,M. and Roberts,R.J. (1997) Homing endonucleases - keeping the house in order. *Nucleic Acids Res.*, **25**, 3379–3388.
67. Chevalier,B.S. and Stoddard,B.L. (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res.*, **29**, 3757–3774.
68. Jacquier,A. and Dujon,B. (1985) An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell*, **41**, 383–394.
69. Stoddard,B.L. (2005) Homing endonuclease structure and function. *Q. Rev. Biophys.*, **38**, 49–95.
70. Chevalier,B., Turmel,M., Lemieux,C., Monnat,R.J. and Stoddard,B.L. (2003) Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI. *J. Mol. Biol.*, **329**, 253–269.
71. Scalley-Kim,M., McConnell-Smith,A. and Stoddard,B.L. (2007) Coevolution of homing endonuclease specificity and its host target sequence. *J. Mol. Biol.*, **372**, 1305–1319.
72. Sokolowska,M., Czapinska,H. and Bochtler,M. (2009) Crystal structure of the beta beta alpha-Me type II restriction endonuclease Hpy99I with target DNA. *Nucleic Acids Res.*, **37**, 3799–3810.
73. Saravanan,M., Bujnicki,J.M., Cymerman,I.A., Rao,D.N. and Nagaraja,V. (2004) Type II restriction endonuclease R.KpnI is a member of the HNH nuclease superfamily. *Nucleic Acids Res.*, **32**, 6129–6135.
74. Kaminska,K.H., Kawai,M., Boniecki,M., Kobayashi,I. and Bujnicki,J.M. (2008) Type II restriction endonuclease R.Hpy188I belongs to the GIY-YIG nuclease superfamily, but exhibits an unusual active site. *BMC Struct. Biol.*, **8**, 48.
75. Ibryashkina,E.M., Zakharova,M.V., Baskunov,V.B., Bogdanova,E.S., Nagornykh,M.O., Den'mukhamedov,M.M., Melnik,B.S., Kolinski,A., Gront,D., Feder,M. *et al.* (2007) Type II restriction endonuclease R.Eco29kI is a member of the GIY-YIG nuclease superfamily. *BMC Struct. Biol.*, **7**, 48.
76. Carlson,K. and Wiberg,J.S. (1983) In vivo cleavage of cytosine-containing bacteriophage T4 DNA to genetically distinct, discretely sized fragments. *J. Virol.*, **48**, 18–30.
77. Kleanthous,C., Kuhlmann,U.C., Pommer,A.J., Ferguson,N., Radford,S.E., Moore,G.R., James,R. and Hemmings,A.M. (1999) Structural and mechanistic basis of immunity toward endonuclease colicins. *Nat. Struct. Biol.*, **6**, 243–252.
78. VanRoey,P., Waddling,C.A., Fox,K.M., Belfort,M. and Derbyshire,V. (2001) Intertwined structure of the DNA-binding domain of intron endonuclease I-TevI with its substrate. *EMBO J.*, **20**, 3631–3637.
79. Shen,B.W., Landthaler,M., Shub,D.A. and Stoddard,B.L. (2004) DNA binding and cleavage by the HNH homing endonuclease I-HmuI. *J. Mol. Biol.*, **342**, 43–56.
80. Mehta,P., Katta,K. and Krishnaswamy,S. (2004) HNH family subclassification leads to identification of commonality in the His-Me endonuclease superfamily. *Protein Sci.*, **13**, 295–300.
81. Zhao,L., Bonocora,R.P., Shub,D.A. and Stoddard,B.L. (2007) The restriction fold turns to the dark side: a bacterial homing endonuclease with a PD-(D/E)-XK motif. *EMBO J.*, **26**, 2432–2442.
82. Goodrich-Blair,H., Scarlato,V., Gott,J.M., Xu,M.Q. and Shub,D.A. (1990) A self-splicing group I intron in the DNA polymerase gene of Bacillus subtilis bacteriophage SP01. *Cell*, **63**, 417–424.
83. Goodrich-Blair,H. and Shub,D.A. (1996) Beyond homing: competition between intron endonucleases confers a selective advantage on flanking genetic markers. *Cell*, **84**, 211–221.
84. Landthaler,M.L. and Shub,D.A. (2003) The nicking homing endonuclease I-BasI is encoded by a group I intron in the DNA polymerase gene of the Bacillus thuringiensis phage Bastille. *Nucleic Acids Res.*, **31**, 3071–3077.
85. Landthaler,M., Lau,N.C. and Shub,D.A. (2004) Group I intron homing in Bacillus phages SP01 and SP82: a gene conversion event initiated by a nicking homing endonuclease. *J. Bacteriol.*, **186**, 4307–4314.
86. Mueller,J., Smith,D., Bryk,M. and Belfort,M. (1995) Intron-encoded endonuclease I-TevI binds as a monomer to effect sequential cleavage via conformational changes in the td homing site. *EMBO J.*, **14**, 5724–5735.
87. VanRoey,P., Meehan,L., Kowalski,J.C., Belfort,M. and Derbyshire,V. (2002) Catalytic domain structure and hypothesis for function of GIY-YIG intron endonuclease I-TevI. *Nat. Struct. Biol.*, **9**, 806–811.
88. Mueller,J.E., Smith,D., Bryk,M. and Belfort,M. (1995) Intron-encoded endonuclease I-TevI binds as a monomer to effect sequential cleavage via conformational changes in the td homing site. *EMBO J.*, **14**, 5724–5735.
89. Carter,J.M., Friedrich,N.C., Kleinstiver,B. and Edgell,D.R. (2007) Strand-specific contacts and divalent metal ion regulate double-strand break formation by the GIY-YIG homing endonuclease I-BmoI. *J. Mol. Biol.*, **374**, 306–321.
90. Kleinstiver,B.P., Fernandes,A.D., Gloor,G.B. and Edgell,D.R. (2010) A unified genetic, computational and experimental framework identifies functionally relevant residues of the homing endonuclease I-BmoI. *Nucleic Acids Res.*, **38**, 2411–2427.
91. Chevalier,B., Sussman,D., Otis,C., Noel,A.J., Turmel,M., Lemieux,C., Stephens,K., Monnat,R.J. Jr and Stoddard,B.L. (2004) Metal-dependent DNA cleavage mechanism of the I-CreI LAGLIDADG homing endonuclease. *Biochemistry*, **43**, 14015–14026.
92. Chevalier,B.S., Monnat,R.J. Jr and Stoddard,B.L. (2001) The homing endonuclease I-CreI uses three metals, one of which is shared between the two active sites. *Nat. Struct. Biol.*, **8**, 312–316.
93. Moure,C.M., Gimble,F.S. and Quiocho,F.A. (2003) The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity. *J. Mol. Biol.*, **334**, 685–695.
94. Turmel,M., Mercier,J.P., Cote,V., Otis,C. and Lemieux,C. (1995) The site-specific DNA endonuclease encoded by a group I intron in the Chlamydomonas pallidostigmatica chloroplast small subunit rRNA gene introduces a single-strand break at low concentrations of Mg2+. *Nucleic Acids Res.*, **23**, 2519–2525.
95. Perrin,A., Buckle,M. and Dujon,B. (1993) Asymmetrical recognition and activity of the I-SceI endonuclease on its site and on intron-exon junctions. *EMBO J.*, **12**, 2939–2947.
96. Geese,W.J., Kwon,Y.K., Wen,X. and Waring,R.B. (2003) In vitro analysis of the relationship between endonuclease and maturase activities in the bi-functional group I intron-encoded protein, I-AniI. *Eur. J. Biochem.*, **270**, 1543–1554.
97. Silva,G.H. and Belfort,M. (2004) Analysis of the LAGLIDADG interface of the monomeric homing endonuclease I-DmoI. *Nucleic Acids Res.*, **32**, 3156–3168.
98. Chevalier,B., Sussman,D., Otis,C., Noel,A.J., Turmel,M., Lemieux,C., Stephens,K.R.J., Monnat,J. and Stoddard,B.L. (2004) Analysis of the symmetric metal-dependent I-CreI homing endonuclease mechanism. *Biochemistry*, **43**, 4015–4026.
99. McConnell Smith,A., Takeuchi,R., Pellenz,S., Davis,L., Maizels,N., Monnat,R.J. Jr. and Stoddard,B.L. (2009) Generation of a nicking enzyme that stimulates site-specific gene

conversion from the I-AniI LAGLIDADG homing endonuclease. *Proc. Natl Acad. Sci. USA*, **106**, 5099–5104.
100. Moure,C.M., Gimble,F.S. and Quiocho,F.A. (2008) Crystal structures of I-SceI complexed to nicked DNA substrates: snapshots of intermediates along the DNA cleavage reaction pathway. *Nucleic Acids Res.*, **36**, 3287–3296.
101. Niu,Y., Tenney,K., Li,H. and Gimble,F.S. (2008) Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA-nicking activity. *J. Mol. Biol.*, **382**, 188–202.
102. Silva,G.H. and Belfort,M. (2004) Analysis of the LAGLIDADG interface of the monomeric homing endonuclease I-DmoI. *Nucleic Acids Res.*, **32**, 3156–3168.
103. Kiesling,T., Cox,K., Davidson,E.A., Dretchen,K., Grater,G., Hibbard,S., Lasken,R.S., Leshin,J., Skowronski,E. and Danielsen,M. (2007) Sequence specific detection of DNA using nicking endonuclease signal amplification (NESA). *Nucleic Acids Res.*, **35**, e117.
104. Murakami,T., Sumaoka,J. and Komiyama,M. (2009) Sensitive isothermal detection of nucleic-acid sequence by primer generation-rolling circle amplification. *Nucleic Acids Res.*, **37**, e19.
105. Van Ness,J., Van Ness,L.K. and Galas,D.J. (2003) Isothermal reactions for the amplification of oligonucleotides. *Proc. Natl Acad. Sci. USA*, **100**, 4504–4509.
106. Tan,E., Erwin,B., Dames,S., Ferguson,T., Buechel,M., Irvine,B., Voelkerding,K. and Niemz,A. (2008) Specific versus nonspecific isothermal DNA amplification through thermophilic polymerase and nicking enzyme activities. *Biochemistry*, **47**, 9987–9999.
107. Tan,E., Erwin,B., Dames,S., Voelkerding,K. and Niemz,A. (2007) Isothermal DNA amplification with gold nanosphere-based visual colorimetric readout for herpes simplex virus detection. *Clin. Chem.*, **53**, 2017–2020.
108. Tan,E., Wong,J., Nguyen,D., Zhang,Y., Erwin,B., Van Ness,L.K., Baker,S.M., Galas,D.J. and Niemz,A. (2005) Isothermal DNA amplification coupled with DNA nanosphere-based colorimetric detection. *Anal. Chem.*, **77**, 7984–7992.
109. Dean,F.B., Hosono,S., Fang,L., Wu,X., Faruqi,A.F., Bray-Ward,P., Sun,Z., Zong,Q., Du,Y., Du,J. *et al.* (2002) Comprehensive human genome amplification using multiple displacement amplification. *Proc. Natl Acad. Sci. USA*, **99**, 5261–5266.
110. Li,J.J., Chu,Y., Lee,B.Y. and Xie,X.S. (2008) Enzymatic signal amplification of molecular beacons for sensitive DNA detection. *Nucleic Acids Res.*, **36**, e36.
111. Dawson,E.D., Taylor,A.W., Smagala,J.A. and Rowlen,K.L. (2009) Molecular detection of Streptococcus pyogenes and Streptococcus dysgalactiae subsp. equisimilis. *Mol. Biotechnol.*, **42**, 117–127.
112. Zheleznaya,L.A., Kopein,D.S., Rogulin,E.A., Gubanov,S.I. and Matvienko,N.I. (2006) Significant enhancement of fluorescence on hybridization of a molecular beacon to a target DNA in the presence of a site-specific DNA nickase. *Anal. Biochem.*, **348**, 123–126.
113. Kuhn,H. and Frank-Kamenetskii,M.D. (2008) Labeling of unique sequences in double-stranded DNA at sites of vicinal nicks generated by nicking endonucleases. *Nucleic Acids Res.*, **36**, e40.
114. Hosoda,K., Matsuura,T., Kita,H., Ichihashi,N., Tsukada,K., Urabe,I. and Yomo,T. (2008) A novel sequence-specific RNA quantification method using nicking endonuclease, dual-labeled fluorescent DNA probe, and conformation-interchangeable oligo-DNA. *RNA*, **14**, 584–592.
115. Jo,K., Dhingra,D.M., Odijk,T., de Pablo,J.J., Graham,M.D., Runnheim,R., Forrest,D. and Schwartz,D.C. (2007) A single-molecule barcoding system using nanoslits for DNA analysis. *Proc. Natl Acad. Sci. USA*, **104**, 2673–2678.
116. Jo,K., Schramm,T.M. and Schwartz,D.C. (2009) A single-molecule barcoding system using nanoslits for DNA analysis: nanocoding. *Methods Mol. Biol.*, **544**, 29–42.
117. Xiao,M., Phong,A., Ha,C., Chan,T.F., Cai,D., Leung,L., Wan,E., Kistler,A.L., DeRisi,J.L., Selvin,P.R. *et al.* (2007) Rapid DNA mapping by fluorescent single molecule detection. *Nucleic Acids Res.*, **35**, e16.
118. Raimondi,S., Johansson,H., Maisonneuve,P. and Gandini,S. (2009) Review and meta-analysis on vitamin D receptor polymorphisms and cancer risk. *Carcinogenesis*, **30**, 1170–1180.
119. Mocellin,S. and Nitti,D. (2008) Vitamin D receptor polymorphisms and the risk of cutaneous melanoma: a systematic review and meta-analysis. *Cancer*, **113**, 2398–2407.
120. Le Provost,F., Lillico,S., Passet,B., Young,R., Whitelaw,B. and Vilotte,J.L. (2010) Zinc finger nuclease technology heralds a new era in mammalian transgenesis. *Trends Biotechnol.*, **28**, 134–141.
121. Katada,H. and Komiyama,M. (2009) Artificial restriction DNA cutters as new tools for gene manipulation. *Chembiochem*, **10**, 1279–1288.
122. Miller,J.C., Holmes,M.C., Wang,J., Guschin,D.Y., Lee,Y.L., Rupniewski,I., Beausejour,C.M., Waite,A.J., Wang,N.S., Kim,K.A. *et al.* (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.*, **25**, 778–785.
123. Lee,G.S., Neiditch,M.B., Salus,S.S. and Roth,D.B. (2004) RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell*, **117**, 171–184.
124. van Nierop,G.P., de Vries,A.A., Holkers,M., Vrijsen,K.R. and Goncalves,M.A. (2009) Stimulation of homology-directed gene targeting at an endogenous human locus by a nicking endonuclease. *Nucleic Acids Res.*, **37**, 5725–5736.
125. Wang,H. and Hays,J.B. (2001) Simple and rapid preparation of gapped plasmid DNA for incorporation of oligomers containing specific DNA lesions. *Mol. Biotechnol.*, **19**, 133–140.
126. Wang,H. and Hays,J.B. (2002) Mismatch repair in human nuclear extracts. Quantitative analyses of excision of nicked circular mismatched DNA substrates, constructed by a new technique employing synthetic oligonucleotides. *J. Biol. Chem.*, **277**, 26136–26142.
127. Stanley,L.K., Seidel,R., van der Scheer,C., Dekker,N.H., Szczelkun,M.D. and Dekker,C. (2006) When a helicase is not a helicase: dsDNA tracking by the motor protein EcoR124I. *EMBO J.*, **25**, 2230–2239.
128. Song,Q., Wu,H., Feng,F., Zhou,G., Kajiyama,T. and Kambara,H. (2010) Pyrosequencing on nicked dsDNA generated by nicking endonucleases. *Anal. Chem.*, **82**, 2074–2081.
129. Zhou,B., Huang,C., Yang,J., Lu,J., Dong,Q. and Sun,L.Z. (2009) Preparation of heteroduplex enhanced green fluorescent protein plasmid for in vivo mismatch repair activity assay. *Anal. Biochem.*, **388**, 167–169.
130. Ang,W.H., Brown,W.W. and Lippard,S.J. (2009) Preparation of mammalian expression vectors incorporating site-specifically platinated-DNA lesions. *Bioconjug. Chem.*, **20**, 1058–1063.
131. Bitinaite,J., Rubino,M., Varma,K.H., Schildkraut,I., Vaisvila,R. and Vaiskunaite,R. (2007) USER friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Res.*, **35**, 1992–2002.
132. Bath,J., Green,S.J. and Turberfield,A.J. (2005) A free-running DNA motor powered by a nicking enzyme. *Angew. Chem. Int. Ed. Engl.*, **44**, 4358–4361.
133. Ho,K.L., McNae,I.W., Schmiedeberg,L., Klose,R.J., Bird,A.P. and Walkinshaw,M.D. (2008) MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell*, **29**, 525–531.
134. Meehan,R.R., Lewis,J.D. and Bird,A.P. (1992) Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res.*, **20**, 5085–5092.
135. Nan,X., Campoy,F.J. and Bird,A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*, **88**, 471–481.
136. Fomenkov,A., Too,P.H., Chan,S.H., Vaisvila,R., Cantin,B.A., Mazzola,L., Tam,V. and Xu,S.Y. (2008) Targeting DNA 5mCpG sites with chimeric endonucleases. *Anal. Biochem.*, **381**, 135–141.

137. Rauch,T. and Pfeifer,G.P. (2005) Methylated-CpG island recovery assay: a new technique for the rapid detection of methylated-CpG islands in cancer. *Lab. Invest.*, **85**, 1172–1180.
138. Fodermayr,M., Proll,J., Zach,O., Wechselberger,C. and Lutz,D. (2009) In vitro detection of methylated DNA via recombinant protein MBD2b. *Mol. Biol. Rep.*, **36**, 1859–1862.
139. Boeke,J., Ammerpohl,O., Kegel,S., Moehren,U. and Renkawitz,R. (2000) The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A. *J. Biol. Chem.*, **275**, 34963–34967.