NAR Breakthrough Article

SURVEY AND SUMMARY

Highlights of the DNA cutters: a short history of the restriction enzymes

Wil A. M. Loenen[1],*, David T. F. Dryden[2],*, Elisabeth A. Raleigh[3],*, Geoffrey G. Wilson[3],* and Noreen E. Murray[†]

[1]Leiden University Medical Center, Leiden, the Netherlands, [2]EaStChemSchool of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, Scotland, UK and [3]New England Biolabs, Inc., 240 County Road, Ipswich, MA 01938, USA

ABSTRACT

In the early 1950's, 'host-controlled variation in bacterial viruses' was reported as a non-hereditary phenomenon: one cycle of viral growth on certain bacterial hosts affected the ability of progeny virus to grow on other hosts by either restricting or enlarging their host range. Unlike mutation, this change was reversible, and one cycle of growth in the previous host returned the virus to its original form. These simple observations heralded the discovery of the endonuclease and methyltransferase activities of what are now termed Type I, II, III and IV DNA restriction-modification systems. The Type II restriction enzymes (e.g. EcoRI) gave rise to recombinant DNA technology that has transformed molecular biology and medicine. This review traces the discovery of restriction enzymes and their continuing impact on molecular biology and medicine.

INTRODUCTION

Restriction endonucleases (REases) such as EcoRI are familiar to virtually everyone who has worked with DNA. Currently, >19 000 putative REases are listed on REBASE (http://rebase.neb.com) (1). REases are classified into four main types, Type I, II, III and IV, with subdivisions for convenience; almost all require a divalent metal cofactor such as Mg^{2+} for activity (Table 1 and Figure 1).

Type II REases represent the largest group of characterized enzymes owing to their usefulness as tools for recombinant DNA technology, and they have been studied extensively. Over 300 Type II REases, with >200 different sequence-specificities, are commercially available. Far fewer Type I, III and IV enzymes have been characterized, but putative examples are being identified daily through bioinformatic analysis of sequenced genomes (Table 1).

Here we present a non-specialists perspective on important events in the discovery and understanding of REases. Studies of these enzymes have generated a wealth of information regarding DNA–protein interactions and catalysis, protein family relationships, control of restriction activity and plasticity of protein domains, as well as providing essential tools for molecular biology research. Discussion of the equally fascinating DNA-methyltransferase (MTase) enzymes that almost always accompany REases *in vivo* is beyond the scope of this review, but we note that base flipping, first discovered in the HhaI MTase (2), is not confined to these enzymes alone, but appears to be a common phenomenon that is also used by certain REases (3) and in other nucleic acid-binding enzymes (4–7).

Most research interest has focused on Type I and II enzymes for historical and practical reasons, so this history is weighted to their treatment. The molecular, genetic and enzymological properties of these have been extensively reviewed [see e.g. (8–12)], and separate reviews of the Type I, III and IV systems appear elsewhere in this journal.

*To whom correspondence should be addressed. Tel: +31 85 878 0248; Email: w.a.m.loenen@lumc.nl
Correspondence may also be addressed to David T. F. Dryden. Tel: +44 131 650 4735; Fax: +44 131 650 6453; Email: david.dryden@ed.ac.uk
Correspondence may also be addressed to Elisabeth A. Raleigh. Tel: +1 978 380 7238; Fax: +1 978 921 1350; Email: raleigh@neb.com
Correspondence may also be addressed to Geoffrey G. Wilson. Tel: +1 978 380 7370; Fax: +1 978 921 1350; Email: wilson@NEB.com
[†]Noreen Murray passed away during the early stages of preparing this review, which is fondly dedicated to her memory.

The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

Table 1. Characterization and organization of the genes and subunits of the four Types of restriction enzymes

Type	Type I	Type II	Type III	Type IV
Features	Oligomeric REase and MTase complex Require ATP hydrolysis for restriction Cleave variably, often far from recognition site 'DEAD-box' translocating REase bipartite DNA recognition domain	Separate REase and MTase or combined REase~MTase fusion Cleave within or at fixed positions close to recognition site Many different subtypes	Combined REase + MTase complex ATP required for restriction Cleave at fixed position outside recognition site 'DEAD-box' REase	Methylation-dependent REase Cleave at variable distance from recognition site Cleave m6A, m5C, hm5C and/or other modified DNA Many different types
Example	e.g. EcoKI	e.g. EcoRI	e.g. EcoP1I	No 'typical' example
Genes	*hsdR, hsdM, hsdS*	e.g. *ecorIR, ecorIM*	e.g. *ecoP1IM, ecoP1IR*	e.g. *mcrA, mcrBC, mrr*
Subunits	~135, ~62 and ~52 kDa	~31 and ~38 kDa for EcoRI	~106 and ~75 kDa for EcoP1I	Unrelated proteins
Proteins	REase: 2R + 2M + S MTase: 2M + S (±2R)	Orthodox REase: 2R Orthodox MTase: M	REase: 1 or 2 R + 2M MTase: 2M (±2R)	Varies
REBASE	104 enzymes, 47 genes cloned, 34 genes sequenced, 5140 putatives	3938 enzymes, 633 genes cloned, 597 sequenced, 9632 putatives	21 enzymes, 19 genes cloned & sequenced, 1889 putatives	18 enzymes & genes cloned, 15 sequenced, 4822 putatives

Type I and II are currently divided in 5 and 11 different subclasses, respectively. Few enzymes have been well-characterized, but based on the current avalanche of sequence information many putative genes belonging to all Types and subtypes are being identified and listed on the restriction enzyme website (http://rebase.neb.com). The modification-dependent Type IV enzymes are highly diverse and only a few have been characterized in any detail. In each case, an example is given of one of the best-characterized enzymes within the different Types I, II and III. Note that Type II enzymes range from simple (shown here for EcoRI) to more complex systems (see Table 2 for the diversity of Type II subtypes).
REBASE count is as of 16 September 2013 (http://rebase.neb.com/cgi-bin/statlist).

Functional roles of R-M Systems
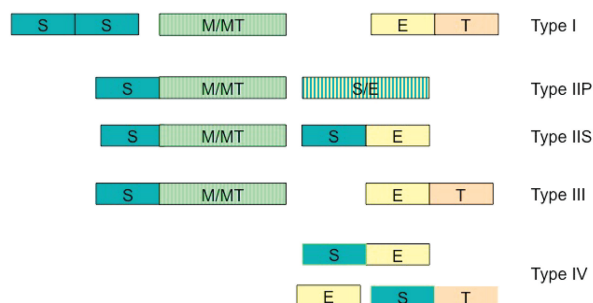
Figure 1. Schematic representation of the functional roles filled in different ways in R-M systems. The functions served include the following: S, DNA sequence specificity; MT, methyltransferase catalytic activity; M, SAM binding; E, endonuclease; T, translocation. Boxes outline functions that are filled by distinct protein domains. Different colours indicate different functions, while different boxes represent distinct domains. Domains in a single polypeptide are abutted, and those in separate polypeptides spaced apart. The order of domains in a polypeptide may vary—e.g. not all Type IIS enzymes have the cleavage domain at the C-terminus. In some cases, functions are integrated with each other, e.g. S and E functions of Type IIP (striped box); in other cases, separate domains carry them out, e.g. Type IIS. See Table 2 for the complexity and diversity of Type II subtypes. The large families of Type I and Type II systems are currently subdivided in 5 and 11 different groups, respectively. The Type I families are distinguished by homology; the Type II groups are distinguished by catalytic properties rather than sequence homology. Type IV enzymes were initially identified as hydroxymethylation-dependent restriction enzymes and currently comprise a highly diverse family. Two examples are shown with and without a translocation domain.

THE FIRST HIGHLIGHTS

Discovery of 'host-controlled variation'

Many important scientific developments in the first half of the 20th century laid the groundwork for to the discovery of restriction and modification (R-M). These included the discoveries of radiation and to the ability to incorporate isotopes in living cells; the molecular building blocks of DNA, RNA and protein; 'filterable agents' (viruses); the isolation of *Escherichia coli* and other bacteria, and of their viruses [called (bacterio)phage] and plasmids. Enabling technical advances included development of electron microscopy, ultracentrifugation, chromatography, electrophoresis and radiographic crystallography. Key was the emerging field of microbial genetics, which flourished owing to the discovery of lysogeny, conjugation, transduction, recombination and mutation.

Preliminary descriptions of the phenomenon of R-M were published by Luria and Human (1952) (13), Anderson and Felix (1952) (14) and Bertani and Weigle (1953) (15). These reports of 'host-controlled variation in bacterial viruses' were reviewed by Luria (1953) (16). Host-controlled variation referred to the observation that the efficiency with which phage infected new bacterial hosts depended on the host on which they previously grew. Phage that propagated efficiently on one bacterial strain could lose that ability if grown for even a single cycle on a different strain. The loss was not due to mutation, and one cycle of growth on the previous

strain returned the virus to its original state once more. R-M systems of all types initially were investigated in this same way, by comparing the 'efficiency of plating' (eop; number of plaques on the test host divided by the number of plaques on a permissive host) on alternate bacterial hosts (17–21). Eop values would range from ∼10^{-1} to 10^{-5}, thus indicating that R-M systems were effective barriers to the uptake of DNA; see (16,22–26) for early reviews.

DNA modification

A decade after these initial reports, Werner Arber and Daisy Dussoix, using phage lambda as experimental system, showed that it was the phage DNA that carried the host-range imprint (17). Different specificities could be imprinted concomitantly both by the bacterium itself (by what were later recognized to be the Type I EcoKI and EcoBI systems) and by phage P1 in its latent prophage state (the Type III EcoP1I system). Gunther Stent suggested that DNA methylation might be the basis for the modification imprint, thus prompting Arber to show that methionine was required in the growth medium to produce the imprint on the DNA (27). This important finding coincided with the discovery of RNA-methyltransferase and MTase activities in bacteria that catalysed the formation of m5C and m6A (28). Arber's interest in the biochemical mechanisms of R-M was driven in part by insight that R-M enzymes would prove useful for analysing DNA molecules and DNA–protein interactions. He concluded a landmark 1965 review of host-controlled modification with the following words: 'Looking toward future developments…it is to be hoped that the enzymes involved in production and control of host specificity will be isolated and characterized. Such studies, paralleled with investigations of the genes controlling R-M and of their expression, should eventually permit an explanation of the high degree of strain specificity, for example "by a mechanism of recognition of certain base sequences". If this last idea should be correct one may further speculate that a restriction enzyme might "provide a tool for the sequence-specific cleavage of DNA"' (22) (our double quotes).

Sequence-specific DNA-cleavage

As chance would have it, the R-M systems studied by Arber, Type I and Type III (25), do not provide simple enzymes for the sequence-specific cleavage of DNA (see further below). However, REases with the desired sequence-specific cleavage were soon isolated, and these set the stage for the advances in gene analysis and manipulation, collectively called 'recombinant DNA technology', that quickly followed. The first of these new enzymes, HindII, was discovered in Hamilton ('Ham') Smith's laboratory at Johns Hopkins Medical School in 1970 (29). This was subsequently termed a Type II REase as its properties were distinct from the Type I REases (25). Purified from *Haemophilus influenzae* serotype d, HindII (originally called endonuclease R) was found to act as a homodimer and to cleave DNA at the symmetric (though degenerate) sequence GTY'RAC (Y = C or T; R = A or

G;' indicates the cut site) (29,30). Subsequently, what was thought to be pure HindII was found to be a mixture of HindII and a second REase made by the same bacterium, HindIII. HindIII cleaved DNA at a different symmetric sequence, A'AGCTT (31,32); [see (33) for a thought-provoking discussion]. The existence of HindIII came to light during experiments to characterize the MTase activities of *H. influenzae*. These experiments showed that the HindII and HindIII MTases acted at the same DNA sequences as those cleaved by the REases. They modified these sequences rather than cleaving them, producing GTYRm6AC and m6AAGCTT, respectively (34–36).

The universe of enzymes in this Type II category expanded rapidly. As Smith's work proceeded on the east coast of the USA, REases with similar behaviour but different specificity were discovered in the laboratory of Herb Boyer at the University of California, San Francisco, on the west coast. Here, PhD student Robert Yoshimori (37) benefited from the experience of Daisy Dussoix, who had moved from Werner Arber's lab to UCSF. Yoshimori investigated restriction systems present on plasmids in clinical *E. coli* isolates, and purified what became known as EcoRI and EcoRII (37,38). The EcoRI REase was found to cleave G'AATT C (39,40) and the corresponding M.EcoRI MTase to modify the inner adenines in this sequence, producing GAm6ATTC (41). The EcoRII REase was found to cleave 'CCWGG (W = A or T), and the M.EcoRII MTase to modify the inner cytosines, producing Cm5CWGG (42,43).

Staggered cuts and the advent of genetic engineering

In contrast to the Type I and Type III enzymes studied in the 1960's, EcoRI and HindIII cleave DNA within their recognition sites and, most importantly, produce staggered cuts. Since the recognition sites are symmetric, this means that every fragment is flanked by the same single-stranded extension, allowing any fragment to anneal (via the extensions) to any other fragment, thus setting the stage for recombining DNA fragments and 'cloning'. These findings were presented at the 1972 EMBO Workshop on Restriction, organized by Werner Arber (see Supplement S1 for details of the program and attendees). Figure 2 shows a photograph of participants at this Workshop, recalled by Noreen Murray as the most exciting meeting in the history of the REases, with discussions on the impact of this vital new information on 'sticky ends' and the implications for novel DNA manipulation. The recently described DNA ligase (44) would allow the joining of DNA fragments with the same sticky ends. EcoRI and HindIII spurred the development of recombinant DNA work through the availability of both purified enzymes and of replicatable carriers known as vectors. Both phage lambda (45) and various plasmids (46,47) were developed into vectors into which DNA fragments generated by EcoRI and HindIII could be ligated.

Fittingly, in 1978, Werner Arber was awarded the Nobel Prize together with Dan Nathans and Ham Smith

EMBO Workshop on Restriction 1972

Figure 2. Photograph of the participants at the EMBO Workshop on restriction in Leuenberg (Basel), Switzerland, 26–30 September 1972, organized by Werner Arber, who took the picture (Archive Noreen Murray). Supplement S1 contains a list with names of the attendees and the program of this meeting, and puts names to faces as far as the attendees could be identified (from the archives of Noreen Murray and Werner Arber).

in recognition for their pioneering work on R-M (www. nobelprize.org).

Emerging genetic and enzymatic complexity

While the 1972 review by Matt Meselson *et al.* (26) mentions only the recognition sequence of HindII, the pace soon quickened. The discovery of new restriction enzymes skyrocketed, as laborious *in vivo* phage-plating assays were replaced by rapid *in vitro* DNA-cleavage assays of cell extracts. Elucidation of differences in recognition and cutting led to the classification of additional distinct classes, or types, of restriction enzymes (25,48), which with extensions and subdivisions has stood the test of time: Type I (exemplified by EcoKI, EcoBI, EcoR124, the 'classical' enzymes); Type II (EcoRI, HindIII, EcoRV, the 'orthodox' enzymes); and Type III (EcoP1I and EcoP15I), Table 1 and Figure 1. Type IV (modification-dependent REases Mcr and Mrr) was added later (49). Sequencing and biochemistry have since led to subdivisions within the Type I (see below) and Type II systems (Table 2) [see (49,50) and http:// rebase.neb.com for nomenclature and details].

The recombinant DNA scare

In their 1975 review (51), Nathans and Smith discuss methods for DNA cleavage and separation of the resulting fragments on gels, as well as the use of REases in other applications, e.g. the physical mapping of chromosomes, taking Simian Virus SV40 (SV40) as an example.

The debate on the safety of recombinant DNA technology started soon after the 1972 EMBO Workshop and reports on the transfer of eukaryotic DNA into *E. coli* [documented by (52)]. The debate was extremely heated, but by 1990 many of the fears had abated as the anticipated dangers did not materialize and the advantages of DNA cloning, and the ability to produce large quantities of pharmaceutically important proteins such as insulin, hormones and vaccines became clear.

FURTHER HIGHLIGHTS IN THE STUDY OF TYPE I R-M SYSTEMS

Type I families are defined by complementation and display sequence conservation

Type I REases were originally identified in *E. coli* and other enteric organisms as barriers to DNA entry. They turned out to be oligomeric proteins encoded by the three *host specificity determinant* (*hsd*) genes: a restriction (R), modification (M) and recognition (S for specificity) gene, respectively (Table 1 and Figure 1). Before the development of DNA sequencing, genetic complementation tests defined the *hsdR*, *hsdM* and *hsdS* genes (53,54). DNA hybridization studies, and probing with antibodies directed at EcoKI, established that EcoKI and EcoBI were more closely related to each other than to EcoAI, the Type I system in *E. coli* 15T$^-$ [reviewed in (8)]. This approach based on biological interaction led to the division of these systems into families: the Type IA (EcoKI, EcoBI,

Table 2. Nomenclature of Type II restriction enzymes

Subtype	Features of restriction enzymes[a]	Examples
Type IIP	Palindromic recognition sequence; recognized by both homodimeric and monomeric enzymes; cleavage occurs symmetrically, usually within the recognition sequence	Prototypes EcoRI & EcoRV
Type IIA	Asymmetric recognition sequence	FokI
Type IIB	Cleavage on both sides of the recognition sequence	BcgI
Type IIC	Single, combination R-M polypeptide	HaeIV
Type IIE	Two sequences required for cleavage, one serving as allosteric effector	EcoRII, Sau3AI
Type IIF	Two sequences required for cleavage, concerted reaction by homotetramer	SfiI
Type IIG	Requires AdoMet cofactor for both R-M	Eco57I
Type IIH	Separate M and S subunits; MTase organization similar to Type I systems	BcgI
Type IIM	Require methylated recognition sequence; Type IIP or Type IIA	DpnI
Type IIS	Asymmetric recognition sequence; cleavage at fixed positions usually outside recognition sequence	FokI
Type IIT	Heterodimeric restriction enzyme.	Bpu10I, BslI
Putatives	All subtypes	
Control	Control proteins of Type II restriction enzymes	C.BamHI, C.PvuII

The characteristics of the orthodox Type IIP enzymes originally distinguished this group of enzymes from the Type I and III R-M systems. Type IIP is the largest group, owing to its valuable role in molecular science and its commercial value, but the current classification and growing number of R-M systems (putatively) identified, makes it clear that Type II enzymes are highly diverse and the boundaries with the other types are beginning to blur; see also Figure 3 and text for details.
[a]These classifications reflect enzyme properties and activities, and not their evolutionary relationships. The classifications are not exclusive, and one enzyme can often belong several classes. Thus BcgI, for example, is Type IIA, B, C, G and H (see text for details).

EcoDI and *Salmonella typhimurium* StySPI); Type IB (EcoAI, EcoEI and *Citrobacter freundii* CfrAI); Type IC (EcoR124, EcoDXXI, EcoprrI) (8); and later, Type ID (StyBLI and *Klebsiella pneumoniae* KpnAI) (9,55,56) and Type IE (KpnBI) (57); see reviews for further details (8–10,58,59). Other organisms will have their own families, for example, *Staphylococcus aureus* has at least two families [(60) and unpublished DTFD results].

Preparation, cofactor requirements and structures

Landmark studies on purified enzymes in the wake of the 1962 Arber and Dussoix articles (17,61) date to 1968. Stu Linn and Werner Arber in Switzerland and Matt Meselson and Bob Yuan in the USA, respectively, purified EcoBI and EcoKI. They used restriction of phages fd and lambda as their assay for detecting the enzymes during purification, a laborious process (62,63). This was no simple matter; Bob Yuan recalls that 'the fall flew by in deep frustration' until he and Matt discovered that the enzyme needed S-adenosylmethionine (SAM) for activity in addition to Mg^{2+} and ATP (See Supplement S2 for his personal story). The same cofactor requirement was also found for EcoBI (62,64), reviewed in (25,26). Twenty-five years later, we have come to appreciate that SAM, like ATP, is a widely used cofactor in many metabolic reactions (65,66).

A long-awaited breakthrough did not happen until much later: The structures of the subunits and assembled Type I R-M enzymes. Two structures of S subunits appeared in 2005 and culminated in 2012 with the structure of two complete R-M enzymes containing two R subunits, two M subunits and one S subunit (67–69).

Type I enzymes cut away from the target site

In 1972, Horiuchi and Zinder showed that the DNA recognition site of EcoBI is not the cleavage site (70).

They cut ^{3}H-labelled double-strand RF DNA of phage f1 (a relative of phage M13) with EcoBI, denatured and renatured the DNA and then treated with EcoBI a second time. This resulted in a heterogeneous distribution of small DNA fragments on alkaline sucrose gradients leading to the conclusion that EcoBI cuts at a variable distance from its target site. This feature is now known to be common to all Type I restriction enzymes. It was later shown by Studier that EcoKI would preferentially cleave DNA approximately half-way between consecutive target sites (71). This feature is also common to all Type I restriction enzymes although the distribution of cleavage locations can be broad.

DNA translocation to reach the cutting sites uses molecular motors

Translocation was first observed in electron microscope (EM) studies that showed DNA looping by EcoBI and EcoKI. These were interpreted as reaction intermediates, formed by the enzymes translocating along the DNA while remaining attached to their recognition sites (72,73). In the case of EcoKI, studies with full-length phage lambda DNA and relaxed or supercoiled circular DNA showed that EcoKI translocates the DNA past itself, concomitant with a large conformational change of the enzyme, creating large bidirectional loops clearly visible in the EM. Recent studies confirm that on DNA binding, the enzyme strongly contracts from an open to a compact form (58,69,74). In contrast, EcoBI appeared to form loops in only one direction. Later studies with EcoBI did show supercoiled structures like EcoKI; however, translocation was still unidirectional, and without any apparent strand selectivity in the cleavage reaction (75,76). The translocation process would explain the cleavage observed half-way between consecutive target sites on a DNA molecule as two translocating enzyme

molecules would collide roughly half-way between target sites.

The R-subunit of Type I enzymes belongs to the SNF2 helicase/translocase superfamily of proteins. These appear to be the result of an ancient fusion between nuclease and ATP-dependent RecA-like (AAA+ or 'motor') domains, a linkage found in many enzymes involved in DNA repair, replication, recombination and chromosome remodelling (77–88). As such, Type I enzymes could prove useful for understanding the action of SNF2 enzymes in higher organisms, including the coordinated steps of DNA scanning, recognition, binding and alteration of the helical structure, that allow other domains or subunits to move and touch the DNA. All of these steps are required to prevent indiscriminate nuclease activity (69).

The key to the functionality of the Type I REase and other SNF2 proteins is their enormous flexibility, allowing large conformational changes. First noticed for EcoKI by Yuan et al (73), and more recently for RecB and EcoR124 (69,89), large protein motions may be a general feature of SNF2 proteins. In line with such large-scale domain movement, mutational analysis of EcoR124 showed long-range effects, e.g. nuclease mutants affect the distant helicase domain leading to a reduced translocation and ATP usage rate, a decrease in the off rate, slower restart and turnover. In other words, the nuclease and motor domain together are 'more than the sum of their parts' (89,90).

Plasticity of type I DNA sequence recognition: hybrid specificities and phase variation

Type I enzymes recognize bipartite DNA sequences [e.g. AAC(N6)GTGC for EcoKI]. The S subunit has a duplicated organization: two ~150 aa variable regions alternate with smaller conserved regions, which are highly similar within each of the five families. Each variable region recognizes one part of the bipartite target sequence.

A key event in understanding the significance and mechanism of variation of sequence specificity was the discovery of a brand-new specificity resulting from a genetic cross (91,92). As a result of crossing-over in the conserved central region between the two variable regions, hybrid specificities were found. This change in specificity was found to occur *in vivo* and *in vitro* and was first noted for *Salmonella* species (91–94). An extensive treatment of this topic is found in the accompanying Type I review.

A variety of genetic processes promote variation in S subunits proteins of Type I and Type III enzymes. In *Lactococcus lactis*, entering plasmids may bring *hsdS* genes with them (95). The segmented organization of *hsdS*, with two DNA recognition domains, lends itself to variation by DNA rearrangement. Site-specific recombination leads to expression of S proteins with alternative recognition domains in *Mycoplasma pulmonis*, thus generating combinatorial variations of recognition sequence (96). Such plasticity of restriction specificity is also inferred in *Bacteroides fragilis* (97) and other species.

## FURTHER HIGHLIGHTS IN THE STUDY OF TYPE II R-M SYSTEMS

Subdivisions of type II enzymes

Type II REases are defined rather broadly as enzymes that cleave DNA at a fixed position with respect to their recognition sequence, and produce distinct DNA-fragment banding patterns during gel electrophoresis. These REases are extremely varied and occur in many structural forms. The Type II classification was used originally to define the simplest kind of REases, exemplified by HindII and EcoRI, that recognize symmetric DNA sequences and require Mg^{2+} ions for cleavage activity (25). Enzymes of this sort generally act as homodimers and cleave DNA within their recognition sequences. *In vivo*, they function in conjunction with a separate modification MTase that acts independently as a monomer (Table 1). The first distinction made among Type II enzymes concerned REases such as HphI and FokI that recognize asymmetric sequences and cleave a short distance away, to one side. These were designated Type IIS (98).

As the number of REases producing distinct fragments grew, it became clear that many unrelated proteins were included in the category (99). Rather than dividing these into further Types based on their phylogenies, it was agreed that 'Type II' should become a utilitarian classification that reflected enzymatic behaviour rather than evolutionary relatedness, and for convenience, a number of Type II groups, corresponding to particular enzymatic behaviours, were defined (49) (Table 2). Each of these groups, A, B, C, E, F, G, H, M, P, S and T, should be thought of, not as an exclusive subdivision, but rather as an icon that signifies some specific property. Enzymes may exhibit more than one salient property and thus belong to more than one group. HindIII and EcoRI remain simple; they are members of just the one, Type IIP, group ('P' for *P*alindromic). BcgI, in contrast, is complicated since it recognizes an asymmetric DNA sequence (= Type IIA); cleaves on both sides of that sequence (= Type IIB); and comprises a fused endonuclease-methyltransferase subunit (= Type IIC) plus a Type I-like DNA-specificity subunit (= Type IIH). BcgI, thus, is a member of multiple groups (100–102). DpnI (Gm6A'TC) is a Type IIM REase, which cleaves its recognition sequence only when the sequence is methylated (103). DpnI is also a member of the Type IIP group since its recognition sequence is palindromic sequence and cleavage is internal and symmetric (Figure 3 and http://rebase.neb.com).

Type IIP ('orthodox') REases such as EcoRI and HindIII were crucial to the development of recombinant DNA technology. Certain 'unorthodox' enzymes have also been widely used. Sau3AI ('GATC) is a monomeric Type IIE REase that dimerizes on the DNA, inducing DNA loops. Two recognition sites must be bound for activity; one is cleaved while the other acts as allosteric effector (104). EcoRII is somewhat similar, and many REases are now known to cleave only as dimers of dimers bound to two separate sites.

Type II restriction enzymes grouped by cleavage properties
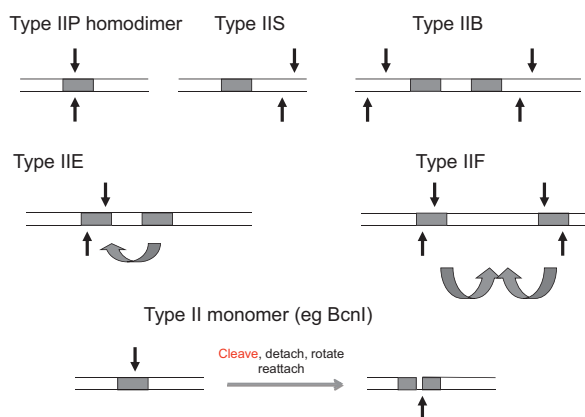
Figure 3. Type II restriction enzymes grouped by cleavage properties. 'Orthodox' IIP enzymes (e.g. EcoRI, EcoRV) cut at the recognition site. Type IIS cut away from the site (e.g. FokI, BfiI). Type IIB require two recognition sites and cut on the outside (e.g. BplI). Type IIE require two recognition sites, and one of the two sites acts as allosteric effector (e.g. EcoRII). Type IIF require two sites and cut at both sites as a tetramer after bringing the two regions together by looping the DNA (e.g. SfiI). Enzymes such as BcnI act as a monomer, in contrast to most Type II REases that act as dimers. See Table 2 and text for further details.

Predicting enzyme families: sequences, structures and bioinformatics

Early amino acid sequences of Type II enzymes (e.g. EcoRI, EcoRV, PvuII and BamHI) showed them to be almost completely unrelated (105–109). When crystal structures appeared (110–114), commonalities began to emerge. The motif PD-(D/E)XK was identified as a common feature (115,116). This motif also appeared in other nucleases, e.g. lambda exonuclease (117) and the Tn7 transposase protein TnsA (118). This motif is the catalytic core of a Mg^{2+}-dependent nuclease.

REBASE, the Restriction Enzyme Database set up by Rich Roberts to keep track of RM specificities and indicate how to acquire the enzymes, made possible the next phase of understanding. First on paper (119,120), then via the nascent Internet by File Transfer Protocol and finally on the World Wide Web (1,121) this resource makes available a focused organized data set allowing computational analysis of sequences and structures as well as access to individual topics of interest [e.g. (99,122–129)].

Recently, Sau3A (104) and several other REases proved able to cut DNA/RNA hybrids (130). The rarity of this property (6 of 223 surveyed) suggests that any biological roles for this ability will be specialized, but the property could be used to study the ubiquitous small RNA molecules that regulate expression in all domains of life (131).

FURTHER HIGHLIGHTS IN THE STUDY OF TYPE III R-M SYSTEMS

Type III enzymes have properties that are intermediate between Types I and II (Table 1 and Figure 1). In general, Type III enzymes recognize asymmetric sequences, cleave 25–27 nucleotides away from their recognition site and use ATP and SAM as cofactors, although they do not have an absolute requirement for the latter. Particularly interesting topics include control of the phage-borne R.EcoP1I REase activity following infection and how newly replicated DNA can be protected when only one strand of the recognition sequence is protected by methylation (132–139).

An early result showed that two copies of the target site were required for DNA cleavage but that these sites had to be in a head-to-head orientation (135,140). A head-to-tail orientation prevented cleavage. How this communication between the two target sites was achieved when ATP hydrolysis was insufficient for DNA translocation like the Type I enzymes (59) has provoked much discussion (141). It appears that DNA looping may have a role in bringing the sites together (142,143), but recent single-molecule analyses (144,145) show strong evidence for enzyme diffusion along the DNA triggered by an ATP-dependent conformational change as a novel mechanism for bringing two copies of the enzyme together to give cleavage, see also (83,146). The long-awaited atomic structure of a Type III R-M enzyme should resolve many of the complexities of these enzymes [AK Aggarwal, personal communication (147,148)].

FURTHER HIGHLIGHTS IN THE STUDY OF TYPE IV RESTRICTION SYSTEMS

Modification-dependent restriction was first observed with populations of phage T4 that contained hydroxymethylcytosine (hm5C)-substituted DNA (13), reviewed in (149,150). This original discovery relied on the fortuitous use of *Shigella dysenteria* SH as permissive host: it lacks both of the *E. coli* K-12 hm5C-targeted endonucleases and also the donor for the protective modification, glucosylation. This allowed glucoseless phage to be propagated in *Shigella*, while picking apart the *E. coli* K-12 set of restricting and modifying genes.

Key advances in the early years lay in determining the nature of the modifications in T-even phage DNA and the genes that enable them. hm5C is incorporated into the DNA during synthesis, and then glucose residues are added in different configurations. The host provides the glucosyl donor (151,152), while the phage provides the glucosyltransferase enzymes (153–155). With these genetic tools in hand, the host genes mediating the phage restriction activity were identified (156). These were named *rglA* and *rglB* (*r*estricts *gl*ucoseless phage) because they mediate restriction of hmC-containing phage that lack the further glucose modification.

In the 1980s, the focus switched to other modifications, particularly m5C, with efforts to clone Type II MTases and eukaryotic DNA into *E. coli* (157–159). The m5C-specific functions *mcrA* and *mcrB* were mapped (160) and were shown to be identical to the *rglA* and *rglB* genes (161). A third modification–dependent enzyme was found to recognize m6A as well as m5C (162). Using the genetic tools described above, glucose-specific activity was identified (163,164). Most recently, a newly described

DNA modification (165) has provided new targets (166) for Type IV enzymes: phosphorothioate linkages in the phosphodiester backbone.

The utility of all these discoveries was, at first, the ability to avoid them (167–169): these restriction systems were found to underlie difficulties encountered in the introduction of foreign MTases into *E. coli* (157,158,170). On the positive side, use of Type IV restriction *in vivo* also allowed enrichment of clone libraries for active eukaryotic genomic sequence, since much transcriptionally silenced DNA is heavily methylated [e.g. (171)].

Type IV enzymes have aroused considerable interest in recent years following the rediscovery of hm5C in the DNA of higher eukaryotes (172–175). This finding could portend the discovery of further, as yet unknown or neglected, DNA modifications. The ability of Type IV enzymes to distinguish between C, m5C, hm5C and other molecular variations of cytosine implicates these enzymes as useful tools for studies of epigenetic phenomena; the commercially available enzyme McrBC has been used for the study of such modification patterns (176,177).

Much history may remain to be written. The accompanying review focuses on structural and enzymatic properties of the systems that are known, and sketches some of the evolutionary pressures faced by restriction systems as they compete with each other and with invading replicons.

CONTROL OF RESTRICTION

Double-stranded cleavage of cellular DNA is extremely deleterious to the host cell, even when it can be repaired. Early in the study of restriction systems, the ease of moving systems among strains with differing systems by conjugation or transduction was noted. This suggested that regulation must be present to enable exchange of activities. More recently, the sporadic distribution of R-M systems in genomes of closely related strains strongly suggests that acquisition of a new system is a relatively frequent event in nature as well. Thus, coordination of expression or activity of the R-M activities is a key research topic. Transcriptional or translational control of Type I systems has not been documented, despite efforts to find it (8,178). However post-translational control is exerted at several levels and is described in the accompanying review on Type I R-M systems. The control of Type II R-M systems recapitulates the mechanisms for other regulatory systems and is described here.

Transcriptional control of Type II enzyme expression

In contrast to Type I enzymes, transcriptional control has been found for Type II enzymes. Most of the Type II systems that have been examined have the problem of integrating control of the modification and restriction activities separately, since they are embodied in separate proteins. Once again, the introduction of these genes into a naïve host is of special interest.

Control of restriction of Type II enzymes: the case of EcoRI

Expression of the MTase gene and methylation of the host DNA before synthesis of the REase is the obvious solution and the so-called 'Hungarian trick' was the basis for the cloning of many of the first restriction enzymes (179). The lab of Ichizo Kobayashi investigated the regulation of the EcoRI gene, *ecoRIR* (180–182). This gene is upstream of the modification gene, *ecoRIM*. The M gene has its own promoters embedded in *ecoRIR* and no transcription terminator between the genes, so *ecoRIM* can be transcribed with and without *ecoRIR*. Using primer extension to locate the start sites and gene fusions to assess expression, two adjacent promoters for *ecoRIM* as well as two reverse promoters were found within *ecoRIR*. These convergent promoters negatively affect each other [as in lambda (183)]. Transcription from the reverse promoter is terminated by the forward promoters and generates a small antisense RNA. The presence of the antisense RNA gene *in trans* reduced lethality mediated by cleavage of under-methylated chromosomes after loss of the EcoRI plasmid (post-segregational killing) (182,184).

Dual transcription control by C proteins

The Blumenthal laboratory provided the first evidence for temporal control in the plasmid-based PvuII system of *Proteus vulgaris* (108,185). A similar open reading frame with similar function was also found contemporaneously in the BamHI system (186,187). In the PvuII system, the MTase is expressed without delay from an independent promoter and protects the host DNA. The REase gene is in an operon with that for an autogenous activator/repressor protein, C.PvuII. Low basal expression from the *pvuIIC* promoter leads to accumulation of the activator, thereby boosting transcription of the C and REase genes (108,185) (Figure 4).

The C protein binds to palindromic DNA sequences (C boxes) defining two sites upstream of its gene: O_L, associated with activation, and O_R, associated with repression. The C protein activates expression of its own gene as well as that of the REase (188). The regulation is similar to gene control in phage lambda: differential binding affinities for the promoters in turn depend on differential DNA sequence and dual symmetry recognition. C proteins belong to the helix-turn-helix family of transcriptional regulators that include the cI and cro repressor proteins of lambdoid phages.

In the wake of PvuII and BamHI, other R-M systems were discovered that were controlled by C proteins, including BglII (189), Eco72I (190), EcoRV (191), Esp396I (192) and SmaI (193). Currently, Rebase lists 19 documented C proteins, as well as 432 putatives based on sequence data (16 September 2013, http://rebase.neb.com). The organization of the genes in the system and regulatory details differ from system to system (108,185,194). There is no published evidence addressing the question of whether R-M systems as a whole evolve in concert with the C proteins. An interesting system would be one homologous to a C-regulated system but without the C gene.

The PvuII and Esp1396I operons
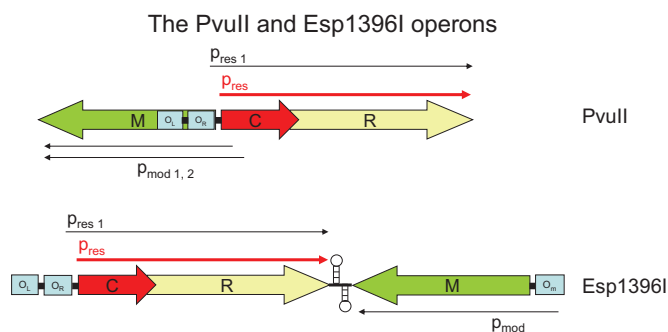
Figure 4. Intricate control of restriction in the operons of the Type II R-M systems of PvuII and Esp1396I by controlling C proteins. A small C gene upstream of, and partially overlapping with, R is coexpressed from p_{res1}, located within the M gene, at low level with R after entry of the self-transmissible PvuII plasmid into a new host, while M is expressed at normal levels from its own two promoters p_{mod1} and p_{mod2} located within the C gene. A similar C protein operates in Esp1396I, but in this case the genes are convergently transcribed with transcription terminator structures in between, and M is expressed from a promoter under negative control of operator O_R, when engaged by C protein in a manner similar to that of the PvuII system. Briefly, the C protein binds to two palindromic sequences (C boxes) defining operator sites O_R and O_L upstream of the C and R genes. After initial low-level expression of C.PvuII protein from the weak promoter p_{res1}, positive feedback by high-affinity binding of a C protein dimer to the distal O_L site later stimulates expression from the second promoter p_{res}, resulting in a leaderless transcript and more C and R protein. The proximal site O_R is a much weaker binding site, but C protein bound at O_L enhances the affinity of O_R for C protein, and at high levels of C protein, the protein-O_R complex downregulates expression of C and R. In this way, C protein is both an activator and negative regulator of its own transcription. In addition, it is a negative regulator of M, which makes sense as overmethylation of DNA may also be harmful to the cell (see text for further details). C.Esp1396I controls O_R, O_L and O_M in a similar manner as described above. In this way, C proteins keep both R-M under control, and have been tentatively identified in >300 R-M systems (Table 2).

Structure of C proteins

The first structures of C-proteins appeared in 2005: C.AhdI from Geoff Kneale's laboratory (195), and C.BclI from a consortium of workers (196). The structures were solved without bound DNA, and while they confirmed the close relationship between C-proteins and helix-turn-helix DNA-binding proteins in general, they did not reveal details of the interactions between C-proteins and their C-box binding sites in DNA (195,197–204). That came 4 years later with the crystal structure of C.Esp1396I bound to DNA (205). This structure, coupled with experimental investigations, revealed the mechanics of the genetic switch and the nature of the sequence-specific and non-specific interactions with the promoters controlling the C/R and M genes (205–208). C.Esp1396I bound as a tetramer, with two dimers bound adjacently on the 35-bp operator sequence $O_L + O_R$ (206). This cooperative binding of dimers to the DNA operator controls the switch from activation to repression of the C and R genes.

Biological consequences of transcriptional regulation

The existence of C proteins explains why it was difficult to introduce some R-M genes in *E. coli*. For instance, the BamHI system of *Bacillus amyloliquefaciens* could only

be maintained in *E. coli* when the REase and MTase were present on one plasmid with an additional copy of the MTase on a second plasmid (209). Further analysis suggested that in *Bacillus subtilis*, a host more closely related to the original expression of R-M was even more stringently regulated (109). C.BamHI enhanced activity of the REase 100-fold in *E. coli*, but at least 1000-fold in *B. subtilis*. In *E. coli*, the C protein repressed expression of the MTase 15-fold. The *B. subtilis* vegetative RNA polymerase is known be more stringent in its promoter sequence requirements than that of *E. coli* (210), possibly accounting for the difference in behaviour in the two species.

Crosstalk among the C genes of similar specificity can allow exclusion with R-M systems of different sequence specificities because of the premature activation of the R gene. The *pvuIIC* and *bamHIC* genes define one incompatibility group of exclusion, whereas *ecoRVC* defines another (211). Entry of a second R-M system thus becomes lethal, a phenomenon called 'apoptotic mutual exclusion' (211).

THE IMPACT OF RESTRICTION ENZYMES

The technical ingenuity applied to the use of restriction enzymes warrants a separate detailed Survey and Summary or indeed an entire book. For instance, their use led to the production of insulin from recombinant bacteria and yeast by Genentech, thus greatly increasing the supply for diabetics and the production of a recombinant vaccine for Hepatitis B by Biogen to treat the hundreds of millions of people at risk of infection by this virus. More recently they have been redesigned to create artifical nucleases, the Zinc-finger nucleases and the TAL-effector nucleases, which have potential for gene targeting and gene therapy (212). Here, we limit ourselves to a few other examples with significant scientific or public impact.

Genetic engineering

Type II enzymes yielded many practical benefits, as *E. coli* K12, its genes and its vectors became the workhorses of molecular biology in the 1970s for cloning, generation of libraries, DNA sequencing, detection and overproduction of enzymes, hormones, etc [e.g. (45,213–224)]. The applications of Type II enzymes continued to expand, especially after the arrival of synthetic DNA, *in vitro* packaging of DNA in phage particles and improved bacterial hosts and vectors for overexpression and stabilization of proteins [see e.g. (225–232)].

A historical perspective on the above topics is beyond the scope of this review. However, a couple of vignettes illustrate how use of REases enabled the research community to leverage a store of understanding to create tools for new advances.

The *lacZ* gene, which had EcoRI sites suitable for early vectors, and its encoded enzyme, beta-galactosidase, had a long history of investigation. Its utility in the creation of cloning vectors relied on identification of domains within the encoded protein, namely a large catalytic domain and

a small multimerization domain. This was discovered by the Muller-Hill group in 1974 (233,234). The 25 N-terminal residues of the small domain can be replaced by peptides of any size and origin without destroying the ability of the multimerization domain to interact with the catalytic domain (233,234). As a result, vectors with short stretches of DNA carrying multiple restriction sites could be created (235–237). Cloning into these sites interrupted the translation of the small domain, destroying its ability to interact with the separately expressed large one. This made possible rapid screening of bacterial colonies on an agar plate for those lacking the activity of LacZ using a colour assay.

In addition, vectors carrying the intact gene but with multiple cloning sites allowed EcoRI-based DNA constructs for transcriptional and translational fusions to the *lacZ* gene (238–250). The majority (90%) of such LacZ-fusion proteins are stable, allowing purification of chimeric antigens, as well as detection of positive clones with colour assays (238,251). Mutagenesis studies in the laboratory of Jeffrey Miller used the *lacZ* gene in phage f1, allowing the rapid detection of spontaneous or induced base substitutions and frameshifts (252–254). This resulted in e.g. LacZ-transgenic mice for studies on DNA damage in different organs and tissues in mammalian cells (255,256).

DNA fingerprinting

Restriction enzymes are tools for monitoring Restriction Fragment Length Polymorphisms, allowing the location of mutations, generation of human linkage maps, identification of disease genes (such as sickle cell trait or Huntington disease), and last, but not least, the DNA fingerprinting technique developed by Alec Jeffreys (257–267). DNA fingerprinting (268) allows the solution of paternity cases, the identification of criminals and their victims and the exoneration of the falsely accused. The use of REases in this system enabled the creation of suitable procedures for such identification, although PCR has largely displaced REases in this application.

REases have also proven useful for identifying pathogenic bacterial strains, most recently of *S. aureus sp* with antibiotic-resistance and virulence factors mediated by mobile genetic elements, e.g. the methicillin-resistant *S. aureus* (MRSA) bacteria (269). Such strains pose a great threat to humans and animals (270).

FINAL THOUGHTS

In 1977, Werner Arber proposed that REases might have additional functions in the cell (271), and this is an idea to keep in mind given that much of the study of restriction enzymes has been aimed at creating tools rather than a basic study of their behaviour in their natural hosts.

For example, the actions of translocating enzymes such as the Type I and IV enzymes at a replication fork or other variant structure are one such possibility (272,273). This activity may seem of arcane interest, but a broader understanding especially of the translocating enzymes could further understanding of genome stabilization activities in all domains of life. Applications to genome manipulation or medicine could emerge. Action at aberrant structures is a major topic of interest in medicine (274).

Lastly, it is interesting to speculate on the condition of molecular biology and all of its associated sciences at the present day if the simple experiment of spreading bacteria and phage on agar plates to follow the restriction-modification phenomenon (13–15) had not been pursued. It is clear that a multi-billion dollar biotechnology industry would not have been spawned, medical diagnostics and the treatment of many diseases would have been severely retarded, genomics and genome sequencing projects would have been difficult if not impossible and their support of bioinformatics and evolutionary studies would also not have been possible, thus greatly diminishing our current appreciation of the spectacular diversity of life on earth.

REFERENCES

1. 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.
2. Roberts,R.J. and Cheng,X. (1998) Base flipping. *Annu. Rev. Biochem.*, **67**, 181–198.
3. Horton,J.R., Liebert,K., Bekes,M., Jeltsch,A. and Cheng,X. (2006) Structure and substrate recognition of the Escherichia coli DNA adenine methyltransferase. *J. Mol. Biol.*, **358**, 559–570.
4. Cheng,X. and Blumenthal,R.M. (2002) Cytosines do it, thymines do it, even pseudouridines do it—base flipping by an enzyme that acts on RNA. *Structure.*, **10**, 127–129.
5. Yang,W. (2011) Surviving the sun: repair and bypass of DNA UV lesions. *Protein Sci.*, **20**, 1781–1789.
6. Brooks,S.C., Adhikary,S., Rubinson,E.H. and Eichman,B.F. (2013) Recent advances in the structural mechanisms of DNA glycosylases. *Biochim. Biophys. Acta*, **1834**, 247–271.
7. Tubbs,J.L., Pegg,A.E. and Tainer,J.A. (2007) DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair (Amst)*, **6**, 1100–1115.
8. Murray,N.E. (2000) Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol. Mol. Biol. Rev.*, **64**, 412–434.
9. Murray,N.E. (2002) 2001 Fred Griffith review lecture. Immigration control of DNA in bacteria: self versus non-self. *Microbiology*, **148**, 3–20.

10. Loenen,W.A. (2003) Tracking EcoKI and DNA fifty years on: a golden story full of surprises. *Nucleic Acids Res.*, **31**, 7059–7069.
11. Pingoud,A. and Jeltsch,A. (2001) Structure and function of type II restriction endonucleases. *Nucleic Acids Res.*, **29**, 3705–3727.
12. Pingoud,A., Fuxreiter,M., Pingoud,V. and Wende,W. (2005) Type II restriction endonucleases: structure and mechanism. *Cell Mol. Life Sci.*, **62**, 685–707.
13. Luria,S.E. and Human,M.L. (1952) A nonhereditary, host-induced variation of bacterial viruses. *J. Bacteriol.*, **64**, 557–569.
14. Anderson,E.S. and Felix,A. (1952) Variation in Vi-phage II of Salmonella typhi. *Nature*, **170**, 492–494.
15. Bertani,G. and Weigle,J.J. (1953) Host controlled variation in bacterial viruses. *J. Bacteriol.*, **65**, 113–121.
16. Luria,S.E. (1953) Host-induced modifications of viruses. *Cold Spring Harb. Symp. Quant. Biol.*, **18**, 237–244.
17. Arber,W. and Dussoix,D. (1962) Host specificity of DNA produced by *Escherichia coli*. I. Host controlled modification of bacteriophage lambda. *J. Mol. Biol.*, **5**, 18–36.
18. Arber,W., Hattman,S. and Dussoix,D. (1963) On the host-controlled modification of bacteriophage lambda. *Virology*, **21**, 30–35.
19. Glover,W., Schell,J., Symonds,N. and Stacey,K.A. (1963) The control of host-induced modification by phage P1. *Genet. Res.*, **4**, 480–482.
20. Lederberg,S. (1966) Genetics of host-controlled restriction and modification of deoxyribonucleic acid in *Escherichia coli*. *J. Bacteriol.*, **91**, 1029–1036.
21. Franklin,N.C. and Dove,W.F. (1969) Genetic evidence for restriction targets in the DNA of phages lambda and phi 80. *Genet. Res.*, **14**, 151–157.
22. Arber,W. (1965) Host-controlled modification of bacteriophage. *Annu. Rev. Microbiol.*, **19**, 365–378.
23. Arber,W. and Linn,S. (1969) DNA modification and restriction. *Annu. Rev. Biochem.*, **38**, 467–500.
24. Arber,W. (1971) Host-Controlled Variation. In: Hershey,A.D. (ed.), *Bacteriophage Lambda*. Cold Spring Harbor Laboratory, New York, pp. 83–96.
25. Boyer,H.W. (1971) DNA restriction and modification mechanisms in bacteria. *Annu. Rev. Microbiol.*, **25**, 153–176.
26. Meselson,M., Yuan,R. and Heywood,J. (1972) Restriction and modification of DNA. *Annu. Rev. Biochem.*, **41**, 447–466.
27. Arber,W. (1965) Host specificit of DNA produced by *Escherichia coli* V The role of methionine in the production of host specificity. *J. Mol. Biol.*, **11**, 247–256.
28. Gold,M., Hurwitz,J. and Anders,M. (1963) The enzymatic methylation of RNA and DNA, II. On the species specificity of the methylation enzymes. *Proc. Natl Acad. Sci. USA*, **50**, 164–169.
29. Kelly,T.J. Jr and Smith,H.O. (1970) A restriction enzyme from *Hemophilus influenzae*. II. *J. Mol. Biol.*, **51**, 393–409.
30. 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.
31. Landy,A., Ruedisueli,E., Robinson,L., Foeller,C. and Ross,W. (1974) Digestion of deoxyribonucleic acids from bacteriophage T7, lambda, and phi 80h with site-specific nucleases from *Hemophilus influenzae* strain Rc and strain Rd. *Biochemistry*, **13**, 2134–2142.
32. Old,R., Murray,K. and Boizes,G. (1975) Recognition sequence of restriction endonuclease III from *Hemophilus influenzae*. *J. Mol. Biol.*, **92**, 331–339.
33. Halford,S.E. (2009) An end to 40 years of mistakes in DNA-protein association kinetics? *Biochem. Soc. Trans.*, **37**, 343–348.
34. Roy,P.H. and Smith,H.O. (1973) DNA methylases of *Hemophilus influenzae* Rd. I. Purification and properties. *J. Mol. Biol.*, **81**, 427–444.
35. Roy,P.H. and Smith,H.O. (1973) DNA methylases of *Hemophilus influenzae* Rd. II. Partial recognition site base sequences. *J. Mol. Biol.*, **81**, 445–459.
36. Roszczyk,E. and Goodgal,S. (1975) Methylase activities from Haemophilus influenzae that protect *Haemophilus parainfluenzae* transforming deoxyribonucleic acid from inactivation by Haemophilus influenzae endonuclease R. *J. Bacteriol.*, **123**, 287–293.
37. Yoshimori,R. (1971) A genetic and biochemical analysis of the restriction and modification of DNA by resistance transfer factors. 146 p. Ph.D Thesis University of California, San Francisco. Thesis/Dissertation.
38. Yoshimori,R., Roulland-Dussoix,D. and Boyer,H.W. (1972) R factor-controlled restriction and modification of deoxyribonucleic acid: restriction mutants. *J. Bacteriol.*, **112**, 1275–1279.
39. Hedgpeth,J., Goodman,H.M. and Boyer,H.W. (1972) DNA nucleotide sequence restricted by the RI endonuclease. *Proc. Natl Acad. Sci. USA*, **69**, 3448–3452.
40. Mertz,J.E. and Davis,R.W. (1972) Cleavage of DNA by R 1 restriction endonuclease generates cohesive ends. *Proc. Natl Acad. Sci. USA*, **69**, 3370–3374.
41. Dugaiczyk,A., Hedgpeth,J., Boyer,H.W. and Goodman,H.M. (1974) Physical identity of the SV40 deoxyribonucleic acid sequence recognized by the Eco RI restriction endonuclease and modification methylase. *Biochemistry*, **13**, 503–512.
42. Boyer,H.W., Chow,L.T., Dugaiczyk,A., Hedgpeth,J. and Goodman,H.M. (1973) DNA substrate site for the EcoRII restriction endonuclease and modification methylase. *Nat. New Biol.*, **244**, 40–43.
43. Bigger,C.H., Murray,K. and Murray,N.E. (1973) Recognition sequence of a restriction enzyme. *Nat. New Biol.*, **244**, 7–10.
44. Lehman,I.R. (1974) DNA ligase: structure, mechanism, and function. *Science*, **186**, 790–797.
45. Murray,N.E. and Murray,K. (1974) Manipulation of restriction targets in phage lambda to form receptor chromosomes for DNA fragments. *Nature*, **251**, 476–481.
46. Cohen,S.N., Chang,A.C., Boyer,H.W. and Helling,R.B. (1973) Construction of biologically functional bacterial plasmids *in vitro*. *Proc. Natl Acad. Sci. USA*, **70**, 3240–3244.
47. Hershfield,V., Boyer,H.W., Yanofsky,C., Lovett,M.A. and Helinski,D.R. (1974) Plasmid ColEl as a molecular vehicle for cloning and amplification of DNA. *Proc. Natl Acad. Sci. USA*, **71**, 3455–3459.
48. Smith,H.O. and Nathans,D. (1973) Letter: a suggested nomenclature for bacterial host modification and restriction systems and their enzymes. *J. Mol. Biol.*, **81**, 419–423.
49. Roberts,R.J., Belfort,M., Bestor,T., Bhagwat,A.S., Bickle,T.A., Bitinaite,J., Blumenthal,R.M., Degtyarev,S.K., 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.
50. Tock,M.R. and Dryden,D.T. (2005) The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.*, **8**, 466–472.
51. Nathans,D. and Smith,H.O. (1975) Restriction endonucleases in the analysis and restructuring of DNA molecules. *Annu. Rev. Biochem.*, **44**, 273–293.
52. Watson,J.D. and Tooze,J. (1981) *The DNA Story: Documentary History of Gene Cloning*. W.H. Freeman and Company, San Francisco.
53. Boyer,H.W. and Roulland-Dussoix,D. (1969) A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.*, **41**, 459–472.
54. Hubacek,J. and Glover,S.W. (1970) Complementation analysis of temperature-sensitive host specificity mutations in *Escherichia coli*. *J. Mol. Biol.*, **50**, 111–127.
55. Titheradge,A.J., King,J., Ryu,J. and Murray,N.E. (2001) Families of restriction enzymes: an analysis prompted by molecular and genetic data for type ID restriction and modification systems. *Nucleic Acids Res.*, **29**, 4195–4205.
56. Kasarjian,J.K., Hidaka,M., Horiuchi,T., Iida,M. and Ryu,J. (2004) The recognition and modification sites for the bacterial type I restriction systems KpnAI, StySEAI, StySENI and StySGI. *Nucleic Acids Res.*, **32**, e82.
57. Chin,V., Valinluck,V., Magaki,S. and Ryu,J. (2004) KpnBI is the prototype of a new family (IE) of bacterial type I restriction-modification system. *Nucleic Acids Res.*, **32**, e138.
58. Dryden,D.T., Murray,N.E. and Rao,D.N. (2001) Nucleoside triphosphate-dependent restriction enzymes. *Nucleic Acids Res.*, **29**, 3728–3741.
59. Bourniquel,A.A. and Bickle,T.A. (2002) Complex restriction enzymes: NTP-driven molecular motors. *Biochimie*, **84**, 1047–1059.

60. Roberts,G.A., Houston,P.J., White,J.H., Chen,K., Stephanou,A.S., Cooper,L.P., Dryden,D.T. and Lindsay,J.A. (2013) Impact of target site distribution for Type I restriction enzymes on the evolution of methicillin-resistant *Staphylococcus aureus* (MRSA) populations. *Nucleic Acids Res.*, **41**, 7472–7484.

61. Dussoix,D. and Arber,W. (1962) Host specificity of DNA produced by *Escherichia coli*. II. Control over acceptance of DNA from infecting phage lambda. *J. Mol. Biol.*, **5**, 37–49.

62. Linn,S. and Arber,W. (1968) Host specificity of DNA produced by *Escherichia coli*, X. *In vitro* restriction of phage fd replicative form. *Proc. Natl Acad. Sci. USA*, **59**, 1300–1306.

63. Meselson,M. and Yuan,R. (1968) DNA restriction enzyme from *E. coli. Nature*, **217**, 1110–1114.

64. Roulland-Dussoix,D. and Boyer,H.W. (1969) The *Escherichia coli* B restriction endonuclease. *Biochim. Biophys. Acta*, **195**, 219–229.

65. Loenen,W.A. (2006) S-adenosylmethionine: jack of all trades and master of everything? *Biochem. Soc. Trans.*, **34**, 330–333.

66. Loenen,W.A.M. (2010) S-adenosylmethionine: simple agent of methylation and secret to aging and metabolism? In: Tollefsbol,T.O. (ed.), *Epigenetics of Aging* Springer, pp. 107–131.

67. Kim,J.S., DeGiovanni,A., Jancarik,J., Adams,P.D., Yokota,H., Kim,R. and Kim,S.H. (2005) Crystal structure of DNA sequence specificity subunit of a type I restriction-modification enzyme and its functional implications. *Proc. Natl Acad. Sci. USA*, **102**, 3248–3253.

68. Calisto,B.M., Pich,O.Q., Pinol,J., Fita,I., Querol,E. and Carpena,X. (2005) Crystal structure of a putative type I restriction-modification S subunit from Mycoplasma genitalium. *J. Mol. Biol.*, **351**, 749–762.

69. Kennaway,C.K., Taylor,J.E., Song,C.F., Potrzebowski,W., Nicholson,W., White,J.H., Swiderska,A., Obarska-Kosinska,A., Callow,P., Cooper,L.P. *et al.* (2012) Structure and operation of the DNA-translocating type I DNA restriction enzymes. *Genes Dev.*, **26**, 92–104.

70. Horiuchi,K. and Zinder,N.D. (1972) Cleavage of bacteriophage fl DNA by the restriction enzyme of *Escherichia coli* B. *Proc. Natl Acad. Sci. USA*, **69**, 3220–3224.

71. Studier,F.W. and Bandyopadhyay,P.K. (1988) Model for how type I restriction enzymes select cleavage sites in DNA. *Proc. Natl Acad. Sci. USA*, **85**, 4677–4681.

72. Rosamond,J., Endlich,B. and Linn,S. (1979) Electron microscopic studies of the mechanism of action of the restriction endonuclease of *Escherichia coli* B. *J. Mol. Biol.*, **129**, 619–635.

73. Yuan,R., Hamilton,D.L. and Burckhardt,J. (1980) DNA translocation by the restriction enzyme from *E. coli* K. *Cell*, **20**, 237–244.

74. Kennaway,C.K., Obarska-Kosinska,A., White,J.H., Tuszynska,I., Cooper,L.P., Bujnicki,J.M., Trinick,J. and Dryden,D.T. (2009) The structure of M.EcoKI Type I DNA methyltransferase with a DNA mimic antirestriction protein. *Nucleic Acids Res.*, **37**, 762–770.

75. Endlich,B. and Linn,S. (1985) The DNA restriction endonuclease of *Escherichia coli* B. II. Further studies of the structure of DNA intermediates and products. *J. Biol. Chem.*, **260**, 5729–5738.

76. Endlich,B. and Linn,S. (1985) The DNA restriction endonuclease of *Escherichia coli* B. I. Studies of the DNA translocation and the ATPase activities. *J. Biol. Chem.*, **260**, 5720–5728.

77. Davies,G.P., Martin,I., Sturrock,S.S., Cronshaw,A., Murray,N.E. and Dryden,D.T. (1999) On the structure and operation of type I DNA restriction enzymes. *J. Mol. Biol.*, **290**, 565–579.

78. Fairman-Williams,M.E., Guenther,U.P. and Jankowsky,E. (2010) SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.*, **20**, 313–324.

79. Gorbalenya,A.E. and Koonin,E.V. (1991) Endonuclease (R) subunits of type-I and type-III restriction-modification enzymes contain a helicase-like domain. *FEBS Lett.*, **291**, 277–281.

80. Gorbalenya,A.E., Koonin,E.V., Donchenko,A.P. and Blinov,V.M. (1988) A conserved NTP–motif in putative helicases. *Nature*, **333**, 22.

81. Hall,M.C. and Matson,S.W. (1999) Helicase motifs: the engine that powers DNA unwinding. *Mol. Microbiol.*, **34**, 867–877.

82. Singleton,M.R., Dillingham,M.S. and Wigley,D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.*, **76**, 23–50.

83. Szczelkun,M.D., Friedhoff,P. and Seidel,R. (2010) Maintaining a sense of direction during long-range communication on DNA. *Biochem. Soc. Trans.*, **38**, 404–409.

84. Tuteja,N. and Tuteja,R. (2004) Prokaryotic and eukaryotic DNA helicases. Essential molecular motor proteins for cellular machinery. *Eur. J. Biochem.*, **271**, 1835–1848.

85. Umate,P., Tuteja,N. and Tuteja,R. (2011) Genome-wide comprehensive analysis of human helicases. *Commun. Integr. Biol.*, **4**, 118–137.

86. Ramanathan,A. and Agarwal,P.K. (2011) Evolutionarily conserved linkage between enzyme fold, flexibility, and catalysis. *PLoS. Biol.*, **9**, e1001193.

87. Mahdi,A.A., Briggs,G.S., Sharples,G.J., Wen,Q. and Lloyd,R.G. (2003) A model for dsDNA translocation revealed by a structural motif common to RecG and Mfd proteins. *EMBO J.*, **22**, 724–734.

88. Rudolph,C.J., Upton,A.L., Briggs,G.S. and Lloyd,R.G. (2010) Is RecG a general guardian of the bacterial genome? *DNA Repair (Amst)*, **9**, 210–223.

89. Sisakova,E., Stanley,L.K., Weiserova,M. and Szczelkun,M.D. (2008) A RecB-family nuclease motif in the Type I restriction endonuclease EcoR124I. *Nucleic Acids Res.*, **36**, 3939–3949.

90. Sisakova,E., Weiserova,M., Dekker,C., Seidel,R. and Szczelkun,M.D. (2008) The interrelationship of helicase and nuclease domains during DNA translocation by the molecular motor EcoR124I. *J. Mol. Biol.*, **384**, 1273–1286.

91. Bullas,L.R., Colson,C. and Van,P.A. (1976) DNA restriction and modification systems in Salmonella. SQ, a new system derived by recombination between the SB system of Salmonella typhimurium and the SP system of Salmonella potsdam. *J. Gen. Microbiol.*, **95**, 166–172.

92. Fuller-Pace,F.V., Bullas,L.R., Delius,H. and Murray,N.E. (1984) Genetic recombination can generate altered restriction specificity. *Proc. Natl Acad. Sci. USA*, **81**, 6095–6099.

93. Nagaraja,V., Shepherd,J.C. and Bickle,T.A. (1985) A hybrid recognition sequence in a recombinant restriction enzyme and the evolution of DNA sequence specificity. *Nature*, **316**, 371–372.

94. Fuller-Pace,F.V. and Murray,N.E. (1986) Two DNA recognition domains of the specificity polypeptides of a family of type I restriction enzymes. *Proc. Natl Acad. Sci. USA*, **83**, 9368–9372.

95. Schouler,C., Gautier,M., Ehrlich,S.D. and Chopin,M.C. (1998) Combinational variation of restriction modification specificities in *Lactococcus lactis. Mol. Microbiol.*, **28**, 169–178.

96. Dybvig,K., Sitaraman,R. and French,C.T. (1998) A family of phase-variable restriction enzymes with differing specificities generated by high-frequency gene rearrangements. *Proc. Natl Acad. Sci. USA*, **95**, 13923–13928.

97. Cerdeno-Tarraga,A.M., Patrick,S., Crossman,L.C., Blakely,G., Abratt,V., Lennard,N., Poxton,I., Duerden,B., Harris,B., Quail,M.A. *et al.* (2005) Extensive DNA inversions in the *B. fragilis* genome control variable gene expression. *Science*, **307**, 1463–1465.

98. Szybalski,W., Kim,S.C., Hasan,N. and Podhajska,A.J. (1991) Class-IIS restriction enzymes—a review. *Gene*, **100**, 13–26.

99. Bujnicki,J.M. (2004) Molecular Phylogenetics of Restriction Enzymes. In: Pingoud,A. (ed.), *Restriction Enzymes*, Vol. 14. Springer, Berlin; New York, pp. 63–93.

100. Kong,H. and Smith,C.L. (1997) Substrate DNA and cofactor regulate the activities of a multi-functional restriction-modification enzyme, BcgI. *Nucleic Acids Res.*, **25**, 3687–3692.

101. Smith,R.M., Marshall,J.J., Jacklin,A.J., Retter,S.E., Halford,S.E. and Sobott,F. (2013) Organization of the BcgI restriction-modification protein for the cleavage of eight phosphodiester bonds in DNA. *Nucleic Acids Res.*, **41**, 391–404.

102. Smith,R.M., Jacklin,A.J., Marshall,J.J., Sobott,F. and Halford,S.E. (2013) Organization of the BcgI restriction-modification protein for the transfer of one methyl group to DNA. *Nucleic Acids Res.*, **41**, 405–417.

103. Lacks,S. and Greenberg,B. (1975) A deoxyribonuclease of *Diplococcus pneumoniae* specific for methylated DNA. *J. Biol. Chem.*, **250**, 4060–4066.

104. Friedhoff,P., Lurz,R., Luder,G. and Pingoud,A. (2001) Sau3AI, a monomeric type II restriction endonuclease that dimerizes on the DNA and thereby induces DNA loops. *J. Biol. Chem.*, **276**, 23581–23588.

105. Greene,P.J., Gupta,M., Boyer,H.W., Brown,W.E. and Rosenberg,J.M. (1981) Sequence analysis of the DNA encoding the Eco RI endonuclease and methylase. *J. Biol. Chem.*, **256**, 2143–2153.

106. Newman,A.K., Rubin,R.A., Kim,S.H. and Modrich,P. (1981) DNA sequences of structural genes for Eco RI DNA restriction and modification enzymes. *J. Biol. Chem.*, **256**, 2131–2139.

107. Bougueleret,L., Schwarzstein,M., Tsugita,A. and Zabeau,M. (1984) Characterization of the genes coding for the Eco RV restriction and modification system of *Escherichia coli. Nucleic Acids Res.*, **12**, 3659–3676.

108. Tao,T. and Blumenthal,R.M. (1992) Sequence and characterization of pvuIIR, the PvuII endonuclease gene, and of pvuIIC, its regulatory gene. *J. Bacteriol.*, **174**, 3395–3398.

109. Brooks,J.E., Nathan,P.D., Landry,D., Sznyter,L.A., Waite-Rees,P., Ives,C.L., Moran,L.S., Slatko,B.E. and Benner,J.S. (1991) Characterization of the cloned BamHI restriction modification system: its nucleotide sequence, properties of the methylase, and expression in heterologous hosts. *Nucleic Acids Res.*, **19**, 841–850.

110. McClarin,J.A., Frederick,C.A., Wang,B.C., Greene,P., Boyer,H.W., Grable,J. and Rosenberg,J.M. (1986) Structure of the DNA-Eco RI endonuclease recognition complex at 3 A resolution. *Science*, **234**, 1526–1541.

111. Kim,Y.C., Grable,J.C., Love,R., Greene,P.J. and Rosenberg,J.M. (1990) Refinement of Eco RI endonuclease crystal structure: a revised protein chain tracing. *Science*, **249**, 1307–1309.

112. Winkler,F.K., Banner,D.W., Oefner,C., Tsernoglou,D., Brown,R.S., Heathman,S.P., Bryan,R.K., Martin,P.D., Petratos,K. and Wilson,K.S. (1993) The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.*, **12**, 1781–1795.

113. Cheng,X., Balendiran,K., Schildkraut,I. and Anderson,J.E. (1994) Structure of PvuII endonuclease with cognate DNA. *EMBO J.*, **13**, 3927–3935.

114. Newman,M., Strzelecka,T., Dorner,L.F., Schildkraut,I. and Aggarwal,A.K. (1994) Structure of restriction endonuclease BamHI and its relationship to EcoRI. *Nature*, **368**, 660–664.

115. Anderson,J.E. (1993) Restriction endonucleases and modification methylases. *Curr. Opin. Struct. Biol.*, **3**, 24–30.

116. Aggarwal,A.K. (1995) Structure and function of restriction endonucleases. *Curr. Opin. Struct. Biol.*, **5**, 11–19.

117. Kovall,R.A. and Matthews,B.W. (1998) Structural, functional, and evolutionary relationships between lambda-exonuclease and the type II restriction endonucleases. *Proc. Natl Acad. Sci. USA*, **95**, 7893–7897.

118. Hickman,A.B., Li,Y., Mathew,S.V., May,E.W., Craig,N.L. and Dyda,F. (2000) Unexpected structural diversity in DNA recombination: the restriction endonuclease connection. *Mol. Cell*, **5**, 1025–1034.

119. Roberts,R.J. (1976) Restriction endonucleases. *CRC Crit. Rev. Biochem.*, **4**, 123–164.

120. Roberts,R.J. (1985) Restriction and modification enzymes and their recognition sequences. *Nucleic Acids Res.*, **13(Suppl)**, r165–r200.

121. Roberts,R.J. and Macelis,D. (1993) REBASE—restriction enzymes and methylases. *Nucleic Acids Res.*, **21**, 3125–3137.

122. 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.

123. Orlowski,J. and Bujnicki,J.M. (2008) Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses. *Nucleic Acids Res.*, **36**, 3552–3569.

124. Bujnicki,J.M. and Rychlewski,L. (2001) Grouping together highly diverged PD-(D/E)XK nucleases and identification of novel superfamily members using structure-guided alignment of sequence profiles. *J. Mol. Microbiol. Biotechnol.*, **3**, 69–72.

125. Sokolowska,M., Czapinska,H. and Bochtler,M. (2011) Hpy188I-DNA pre- and post-cleavage complexes—snapshots of the GIY-YIG nuclease mediated catalysis. *Nucleic Acids Res.*, **39**, 1554–1564.

126. Cymerman,I.A., Obarska,A., Skowronek,K.J., Lubys,A. and Bujnicki,J.M. (2006) Identification of a new subfamily of HNH nucleases and experimental characterization of a representative member, HphI restriction endonuclease. *Proteins*, **65**, 867–876.

127. Chan,S.H., Opitz,L., Higgins,L., O'loane,D. and Xu,S.Y. (2010) Cofactor requirement of HpyAV restriction endonuclease. *PLoS One*, **5**, e9071.

128. Mak,A.N., Lambert,A.R. and Stoddard,B.L. (2010) Folding, DNA recognition, and function of GIY-YIG endonucleases: crystal structures of R.Eco29kI. *Structure*, **18**, 1321–1331.

129. 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.

130. Murray,I.A., Stickel,S.K. and Roberts,R.J. (2010) Sequence-specific cleavage of RNA by Type II restriction enzymes. *Nucleic Acids Res.*, **38**, 8257–8268.

131. Mochizuki,A., Yahara,K., Kobayashi,I. and Iwasa,Y. (2006) Genetic addiction: selfish gene's strategy for symbiosis in the genome. *Genetics*, **172**, 1309–1323.

132. Hadi,S.M., Bachi,B., Iida,S. and Bickle,T.A. (1983) DNA restriction—modification enzymes of phage P1 and plasmid p15B. Subunit functions and structural homologies. *J. Mol. Biol.*, **165**, 19–34.

133. Redaschi,N. and Bickle,T.A. (1996) Posttranscriptional regulation of EcoP1I and EcoP15I restriction activity. *J. Mol. Biol.*, **257**, 790–803.

134. Kruger,D.H., Kupper,D., Meisel,A., Reuter,M. and Schroeder,C. (1995) The significance of distance and orientation of restriction endonuclease recognition sites in viral DNA genomes. *FEMS Microbiol. Rev.*, **17**, 177–184.

135. Meisel,A., Bickle,T.A., Kruger,D.H. and Schroeder,C. (1992) Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. *Nature*, **355**, 467–469.

136. Humbelin,M., Suri,B., Rao,D.N., Hornby,D.P., Eberle,H., Pripfl,T., Kenel,S. and Bickle,T.A. (1988) Type III DNA restriction and modification systems EcoP1 and EcoP15. Nucleotide sequence of the EcoP1 operon, the EcoP15 mod gene and some EcoP1 mod mutants. *J. Mol. Biol.*, **200**, 23–29.

137. Bachi,B., Reiser,J. and Pirrotta,V. (1979) Methylation and cleavage sequences of the EcoP1 restriction-modification enzyme. *J. Mol. Biol.*, **128**, 143–163.

138. Piekarowicz,A. and Brzezinski,R. (1980) Cleavage and methylation of DNA by the restriction endonuclease HinfIII isolated from Haemophilus influenzae Rf. *J. Mol. Biol.*, **144**, 415–429.

139. Kruger,D.H., Schroeder,C., Reuter,M., Bogdarina,I.G., Buryanov,Y.I. and Bickle,T.A. (1985) DNA methylation of bacterial viruses T3 and T7 by different DNA methylases in *Escherichia coli* K12 cells. *Eur. J. Biochem.*, **150**, 323–330.

140. Meisel,A., Mackeldanz,P., Bickle,T.A., Kruger,D.H. and Schroeder,C. (1995) Type III restriction endonucleases translocate DNA in a reaction driven by recognition site-specific ATP hydrolysis. *EMBO J.*, **14**, 2958–2966.

141. Dryden,D.T., Edwardson,J.M. and Henderson,R.M. (2011) DNA translocation by type III restriction enzymes: a comparison of current models of their operation derived from ensemble and single-molecule measurements. *Nucleic Acids Res.*, **39**, 4525–4531.

142. Crampton,N., Roes,S., Dryden,D.T., Rao,D.N., Edwardson,J.M. and Henderson,R.M. (2007) DNA looping and translocation provide an optimal cleavage mechanism for the type III restriction enzymes. *EMBO J.*, **26**, 3815–3825.

143. Crampton,N., Yokokawa,M., Dryden,D.T., Edwardson,J.M., Rao,D.N., Takeyasu,K., Yoshimura,S.H. and Henderson,R.M. (2007) Fast-scan atomic force microscopy reveals that the type III restriction enzyme EcoP15I is capable of DNA translocation and looping. *Proc. Natl Acad. Sci. USA*, **104**, 12755–12760.

144. Ramanathan,S.P., van,A.K., Sears,A., Peakman,L.J., Diffin,F.M., Szczelkun,M.D. and Seidel,R. (2009) Type III restriction enzymes communicate in 1D without looping between their target sites. *Proc. Natl Acad. Sci. USA*, **106**, 1748–1753.

145. Schwarz,F.W., van,A.K., Toth,J., Seidel,R. and Szczelkun,M.D. (2011) DNA cleavage site selection by Type III restriction

enzymes provides evidence for head-on protein collisions following 1D bidirectional motion. *Nucleic Acids Res.*, **39**, 8024–8051.

146. Szczelkun,M.D. (2011) Translocation, switching and gating: potential roles for ATP in long-range communication on DNA by Type III restriction endonucleases. *Biochem. Soc. Trans.*, **39**, 589–594.

147. Gupta,Y.K., Yang,L., Chan,S.H., Samuelson,J.C., Xu,S.Y. and Aggarwal,A.K. (2012) Structural insights into the assembly and shape of Type III restriction-modification (R-M) EcoP15I complex by small-angle X-ray scattering. *J. Mol. Biol.*, **420**, 261–268.

148. Wyszomirski,K.H., Curth,U., Alves,J., Mackeldanz,P., Moncke-Buchner,E., Schutkowski,M., Kruger,D.H. and Reuter,M. (2012) Type III restriction endonuclease EcoP15I is a heterotrimeric complex containing one Res subunit with several DNA-binding regions and ATPase activity. *Nucleic Acids Res.*, **40**, 3610–3622.

149. Revel,H.R. and uria,S.E. (1970) DNA-glucosylation in T-even phage: genetic determination and role in phagehost interaction. *Annu. Rev. Genet.*, **4**, 177–192.

150. Revel,H.R. (1983) DNA modification: glucosylation. In: Mathews,C.K., Kutter,E.M., Mosig,G. and Berget,P. (eds), *Bacteriophage T4*. American Society of Microbiology, Washington DC, pp. 156–165.

151. Hattman,S. and Fukasawa,T. (1963) Host-induced modification of T-even phages due to defective glucosylation of their DNA. *Proc. Natl Acad. Sci. USA*, **50**, 297–300.

152. Shedlovsky,A. and Brenner,S. (1963) A chemical basis for the host-induced modification of T-even bacteriophages. *Proc. Natl Acad. Sci. USA*, **50**, 300–305.

153. Revel,H.R., Hattman,S. and Luria,S.E. (1965) Mutants of bacteriophages T2 and T6 defective in alpha-glucosyl transferase. *Biochem. Biophys. Res. Commun.*, **18**, 545–550.

154. Georgopoulos,C.P. (1967) Isolation and preliminary characterization of T4 mutants with nonglucosylated DNA. *Biochem. Biophys. Res. Commun.*, **28**, 179–184.

155. Georgopoulos,C.P. and Revel,H.R. (1971) Studies with glucosyl transferase mutants of the T-even bacteriophages. *Virology*, **44**, 271–285.

156. Revel,H.R. (1967) Restriction of nonglucosylated T-even bacteriophage: properties of permissive mutants of *Escherichia coli* B and K12. *Virology*, **31**, 688–701.

157. Noyer-Weidner,M., Diaz,R. and Reiners,L. (1986) Cytosine-specific DNA modification interferes with plasmid establishment in *Escherichia coli* K12: involvement of rglB. *Mol. Gen. Genet.*, **205**, 469–475.

158. Raleigh,E.A. and Wilson,G. (1986) *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. *Proc. Natl Acad. Sci. USA*, **83**, 9070–9074.

159. Blumenthal,R.M. (1987) The PvuII restriction-modification system: cloning, characterization, and use in revealing an *E. coli* barrier to certain methylases or methylated DNA. In: Chirikjian,J.G. (ed.), *Gene Amplification and Analysis*, Vol. 5. Elsevier, New York, pp. 227–245.

160. Ravi,R.S., Sozhamannan,S. and Dharmalingam,K. (1985) Transposon mutagenesis and genetic mapping of the rglA and rglB loci of *Escherichia coli*. *Mol. Gen. Genet.*, **198**, 390–392.

161. Raleigh,E.A., Trimarchi,R. and Revel,H. (1989) Genetic and physical mapping of the mcrA (rglA) and mcrB (rglB) loci of *Escherichia coli* K-12. *Genetics*, **122**, 279–296.

162. Heitman,J. and Model,P. (1987) Site-specific methylases induce the SOS DNA repair response in *Escherichia coli*. *J. Bacteriol.*, **169**, 3243–3250.

163. Janosi,L., Yonemitsu,H., Hong,H. and Kaji,A. (1994) Molecular cloning and expression of a novel hydroxymethylcytosine-specific restriction enzyme (PvuRts1I) modulated by glucosylation of DNA. *J. Mol. Biol.*, **242**, 45–61.

164. Bair,C.L. and Black,L.W. (2007) A type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. *J. Mol. Biol.*, **366**, 768–778.

165. Zhou,X., He,X., Liang,J., Li,A., Xu,T., Kieser,T., Helmann,J.D. and Deng,Z. (2005) A novel DNA modification by sulphur. *Mol. Microbiol.*, **57**, 1428–1438.

166. Liu,G., Ou,H.Y., Wang,T., Li,L., Tan,H., Zhou,X., Rajakumar,K., Deng,Z. and He,X. (2010) Cleavage of phosphorothioated DNA and methylated DNA by the type IV restriction endonuclease ScoMcrA. *PLoS Genet.*, **6**, e1001253.

167. Raleigh,E.A., Murray,N.E., Revel,H., Blumenthal,R.M., Westaway,D., Reith,A.D., Rigby,P.W., Elhai,J. and Hanahan,D. (1988) McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning. *Nucleic Acids Res.*, **16**, 1563–1575.

168. Woodcock,D.M., Crowther,P.J., Doherty,J., Jefferson,S., DeCruz,E., Noyer-Weidner,M., Smith,S.S., Michael,M.Z. and Graham,M.W. (1989) Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res.*, **17**, 3469–3478.

169. Graham,M.W., Doherty,J.P. and Woodcock,D.M. (1990) Efficient construction of plant genomic libraries requires the use of *mcr*- host strains and packaging mixes. *Plant Mol. Biol. Rep.*, **8**, 33–42.

170. Blumenthal,R.M., Gregory,S.A. and Cooperider,J.S. (1985) Cloning of a restriction-modification system from Proteus vulgaris and its use in analyzing a methylase-sensitive phenotype in *Escherichia coli*. *J. Bacteriol.*, **164**, 501–509.

171. Palmer,L.E., Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V.S., Nascimento,L.U., Dike,S., de la Bastide,M., Martienssen,R.A. and McCombie,W.R. (2003) Maize genome sequencing by methylation filtration. *Science*, **302**, 2115–2117.

172. Penn,N.W., Suwalski,R., O'Riley,C., Bojanowski,K. and Yura,R. (1972) The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem. J.*, **126**, 781–790.

173. Penn,N.W. (1976) Modification of brain deoxyribonucleic acid base content with maturation in normal and malnourished rats. *Biochem. J.*, **155**, 709–712.

174. Kriaucionis,S. and Heintz,N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*, **324**, 929–930.

175. Tahiliani,M., Koh,K.P., Shen,Y., Pastor,W.A., Bandukwala,H., Brudno,Y., Agarwal,S., Iyer,L.M., Liu,D.R., Aravind,L. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930–935.

176. Ordway,J.M., Bedell,J.A., Citek,R.W., Nunberg,A., Garrido,A., Kendall,R., Stevens,J.R., Cao,D., Doerge,R.W., Korshunova,Y. *et al.* (2006) Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets. *Carcinogenesis*, **27**, 2409–2423.

177. Irizarry,R.A., Ladd-Acosta,C., Carvalho,B., Wu,H., Brandenburg,S.A., Jeddeloh,J.A., Wen,B. and Feinberg,A.P. (2008) Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res.*, **18**, 780–790.

178. Loenen,W.A., Daniel,A.S., Braymer,H.D. and Murray,N.E. (1987) Organization and sequence of the hsd genes of *Escherichia coli* K-12. *J. Mol. Biol.*, **198**, 159–170.

179. Szomolanyi,E., Kiss,A. and Venetianer,P. (1980) Cloning the modification methylase gene of Bacillus sphaericus R in *Escherichia coli*. *Gene*, **10**, 219–225.

180. Liu,Y., Ichige,A. and Kobayashi,I. (2007) Regulation of the EcoRI restriction-modification system: Identification of ecoRIM gene promoters and their upstream negative regulators in the ecoRIR gene. *Gene*, **400**, 140–149.

181. Liu,Y. and Kobayashi,I. (2007) Negative regulation of the EcoRI restriction enzyme gene is associated with intragenic reverse promoters. *J. Bacteriol.*, **189**, 6928–6935.

182. Mruk,I., Liu,Y., Ge,L. and Kobayashi,I. (2011) Antisense RNA associated with biological regulation of a restriction-modification system. *Nucleic Acids Res.*, **39**, 5622–5632.

183. Ward,D.F. and Murray,N.E. (1979) Convergent transcription in bacteriophage lambda: interference with gene expression. *J. Mol. Biol.*, **133**, 249–266.

184. Heitman,J., Zinder,N.D. and Model,P. (1989) Repair of the *Escherichia coli* chromosome after *in vivo* scission by the EcoRI endonuclease. *Proc. Natl Acad. Sci. USA*, **86**, 2281–2285.

185. Tao,T., Bourne,J.C. and Blumenthal,R.M. (1991) A family of regulatory genes associated with type II restriction-modification systems. *J. Bacteriol.*, **173**, 1367–1375.

186. Ives,C.L., Nathan,P.D. and Brooks,J.E. (1992) Regulation of the BamHI restriction-modification system by a small intergenic open reading frame, bamHIC, in both *Escherichia coli* and *Bacillus subtilis*. *J. Bacteriol.*, **174**, 7194–7201.

187. Sohail,A., Ives,C.L. and Brooks,J.E. (1995) Purification and characterization of C.BamHI, a regulator of the BamHI restriction-modification system. *Gene*, **157**, 227–228.

188. Bart,A., Dankert,J. and van der Ende,A. (1999) Operator sequences for the regulatory proteins of restriction modification systems. *Mol. Microbiol.*, **31**, 1277–1278.

189. Anton,B.P., Heiter,D.F., Benner,J.S., Hess,E.J., Greenough,L., Moran,L.S., Slatko,B.E. and Brooks,J.E. (1997) Cloning and characterization of the Bg/II restriction-modification system reveals a possible evolutionary footprint. *Gene*, **187**, 19–27.

190. Rimseliene,R., Vaisvila,R. and Janulaitis,A. (1995) The eco72IC gene specifies a trans-acting factor which influences expression of both DNA methyltransferase and endonuclease from the Eco72I restriction-modification system. *Gene*, **157**, 217–219.

191. Zheleznaya,L.A., Kainov,D.E., Yunusova,A.K. and Matvienko,N.I. (2003) Regulatory C protein of the EcoRV modification-restriction system. *Biochemistry (Mosc.)*, **68**, 105–110.

192. Cesnaviciene,E., Mitkaite,G., Stankevicius,K., Janulaitis,A. and Lubys,A. (2003) Esp1396I restriction-modification system: structural organization and mode of regulation. *Nucleic Acids Res.*, **31**, 743–749.

193. Heidmann,S., Seifert,W., Kessler,C. and Domdey,H. (1989) Cloning, characterization and heterologous expression of the SmaI restriction-modification system. *Nucleic Acids Res.*, **17**, 9783–9796.

194. Semenova,E., Minakhin,L., Bogdanova,E., Nagornykh,M., Vasilov,A., Heyduk,T., Solonin,A., Zakharova,M. and Severinov,K. (2005) Transcription regulation of the EcoRV restriction-modification system. *Nucleic Acids Res.*, **33**, 6942–6951.

195. McGeehan,J.E., Streeter,S.D., Papapanagiotou,I., Fox,G.C. and Kneale,G.G. (2005) High-resolution crystal structure of the restriction-modification controller protein C.AhdI from Aeromonas hydrophila. *J. Mol. Biol.*, **346**, 689–701.

196. Sawaya,M.R., Zhu,Z., Mersha,F., Chan,S.H., Dabur,R., Xu,S.Y. and Balendiran,G.K. (2005) Crystal structure of the restriction-modification system control element C.Bcll and mapping of its binding site. *Structure.*, **13**, 1837–1847.

197. Bogdanova,E., Djordjevic,M., Papapanagiotou,I., Heyduk,T., Kneale,G. and Severinov,K. (2008) Transcription regulation of the type II restriction-modification system AhdI. *Nucleic Acids Res.*, **36**, 1429–1442.

198. Callow,P., Sukhodub,A., Taylor,J.E. and Kneale,G.G. (2007) Shape and subunit organisation of the DNA methyltransferase M.AhdI by small-angle neutron scattering. *J. Mol. Biol.*, **369**, 177–185.

199. Marks,P., McGeehan,J., Wilson,G., Errington,N. and Kneale,G. (2003) Purification and characterisation of a novel DNA methyltransferase, M.AhdI. *Nucleic Acids Res.*, **31**, 2803–2810.

200. Marks,P., McGeehan,J. and Kneale,G. (2004) A novel strategy for the expression and purification of the DNA methyltransferase, M.AhdI. *Protein Expr. Purif.*, **37**, 236–242.

201. McGeehan,J.E., Papapanagiotou,I., Streeter,S.D. and Kneale,G.G. (2006) Cooperative binding of the C.AhdI controller protein to the C/R promoter and its role in endonuclease gene expression. *J. Mol. Biol.*, **358**, 523–531.

202. McGeehan,J.E., Streeter,S., Cooper,J.B., Mohammed,F., Fox,G.C. and Kneale,G.G. (2004) Crystallization and preliminary X-ray analysis of the controller protein C.AhdI from Aeromonas hydrophilia. *Acta Crystallogr. D. Biol. Crystallogr.*, **60**, 323–325.

203. Papapanagiotou,I., Streeter,S.D., Cary,P.D. and Kneale,G.G. (2007) DNA structural deformations in the interaction of the controller protein C.AhdI with its operator sequence. *Nucleic Acids Res.*, **35**, 2643–2650.

204. Streeter,S.D., Papapanagiotou,I., McGeehan,J.E. and Kneale,G.G. (2004) DNA footprinting and biophysical characterization of the controller protein C.AhdI

205. Ball,N., Streeter,S.D., Kneale,G.G. and McGeehan,J.E. (2009) Structure of the restriction-modification controller protein C.Esp1396I. *Acta Crystallogr. D. Biol. Crystallogr.*, **65**, 900–905.

206. McGeehan,J.E., Streeter,S.D., Thresh,S.J., Ball,N., Ravelli,R.B. and Kneale,G.G. (2008) Structural analysis of the genetic switch that regulates the expression of restriction-modification genes. *Nucleic Acids Res.*, **36**, 4778–4787.

207. Ball,N.J., McGeehan,J.E., Streeter,S.D., Thresh,S.J. and Kneale,G.G. (2012) The structural basis of differential DNA sequence recognition by restriction-modification controller proteins. *Nucleic Acids Res.*, **40**, 10532–10542.

208. McGeehan,J.E., Ball,N.J., Streeter,S.D., Thresh,S.J. and Kneale,G.G. (2012) Recognition of dual symmetry by the controller protein C.Esp1396I based on the structure of the transcriptional activation complex. *Nucleic Acids Res.*, **40**, 4158–4167.

209. Brooks,J.E., Benner,J.S., Heiter,D.F., Silber,K.R., Sznyter,L.A., Jager-Quinton,T., Moran,L.S., Slatko,B.E., Wilson,G.G. and Nwankwo,D.O. (1989) Cloning the BamHI restriction modification system. *Nucleic Acids Res.*, **17**, 979–997.

210. Moran,C.P. Jr, Lang,N., LeGrice,S.F., Lee,G., Stephens,M., Sonenshein,A.L., Pero,J. and Losick,R. (1982) Nucleotide sequences that signal the initiation of transcription and translation in Bacillus subtilis. *Mol. Gen. Genet.*, **186**, 339–346.

211. Nakayama,Y. and Kobayashi,I. (1998) Restriction-modification gene complexes as selfish gene entities: roles of a regulatory system in their establishment, maintenance, and apoptotic mutual exclusion. *Proc. Natl Acad. Sci. USA*, **95**, 6442–6447.

212. Segal,D.J. and Meckler,J.F. (2013) Genome engineering at the dawn of the golden age. *Annu. Rev. Genomics Hum. Genet.*, **14**, 135–158.

213. Rambach,A. and Tiollais,P. (1974) Bacteriophage lambda having EcoRI endonuclease sites only in the nonessential region of the genome. *Proc. Natl Acad. Sci. USA*, **71**, 3927–3930.

214. Thomas,M., Cameron,J.R. and Davis,R.W. (1974) Viable molecular hybrids of bacteriophage lambda and eukaryotic DNA. *Proc. Natl Acad. Sci. USA*, **71**, 4579–4583.

215. Kelley,W.S., Chalmers,K. and Murray,N.E. (1977) Isolation and characterization of a lambdapolA transducing phage. *Proc. Natl Acad. Sci. USA*, **74**, 5632–5636.

216. Murray,N.E. and Kelley,W.S. (1979) Characterization of lambdapolA transducing phages; effective expression of the *E. coli* polA gene. *Mol. Gen. Genet.*, **175**, 77–87.

217. Wilson,G.G. and Murray,N.E. (1979) Molecular cloning of the DNA ligase gene from bacteriophage T4. I. Characterisation of the recombinants. *J. Mol. Biol.*, **132**, 471–491.

218. Murray,N.E., Bruce,S.A. and Murray,K. (1979) Molecular cloning of the DNA ligase gene from bacteriophage T4. II. Amplification and preparation of the gene product. *J. Mol. Biol.*, **132**, 493–505.

219. Midgley,C.A. and Murray,N.E. (1985) T4 polynucleotide kinase; cloning of the gene (pseT) and amplification of its product. *EMBO J.*, **4**, 2695–2703.

220. Sutcliffe,J.G. (1978) pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. *Nucleic Acids Res.*, **5**, 2721–2728.

221. Sutcliffe,J.G. (1979) Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. *Cold Spring Harb. Symp. Quant. Biol.*, **43(Pt.1)**, 77–90.

222. Klein,B. and Murray,K. (1979) Phage lambda receptor chromosomes for DNA fragments made with restriction endonuclease I of Bacillus amyloliquefaciens H. *J. Mol. Biol.*, **133**, 289–294.

223. Loenen,W.A. and Brammar,W.J. (1980) A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes. *Gene*, **10**, 249–259.

224. Karn,J., Brenner,S., Barnett,L. and Cesareni,G. (1980) Novel bacteriophage lambda cloning vector. *Proc. Natl Acad. Sci. USA*, **77**, 5172–5176.

225. Maniatis,T., Hardison,R.C., Lacy,E., Lauer,J., O'Connell,C., Quon,D., Sim,G.K. and Efstratiadis,A. (1978) The isolation of

structural genes from libraries of eucaryotic DNA. *Cell*, **15**, 687–701.

226. Maniatis,T., Fritsch,A. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

227. Hendrix,R.W., Roberts,J.W., Stahl,F.W. and Weisberg,R.A. (eds), (1983) *Lambda II*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

228. Stahl,F.W., Crasemann,J.M. and Stahl,M.M. (1975) Rec-mediated recombinational hot spot activity in bacteriophage lambda. III. Chi mutations are site-mutations stimulating rec-mediated recombination. *J. Mol. Biol.*, **94**, 203–212.

229. Moir,A. and Brammar,W.J. (1976) The use of specialised transducing phages in the amplification of enzyme production. *Mol. Gen. Genet.*, **149**, 87–99.

230. Hohn,B. and Murray,K. (1977) Packaging recombinant DNA molecules into bacteriophage particles *in vitro*. *Proc. Natl Acad. Sci. USA*, **74**, 3259–3263.

231. Sternberg,N., Tiemeier,D. and Enquist,L. (1977) *In vitro* packaging of a lambda Dam vector containing EcoRI DNA fragments of *Escherichia coli* and phage P1. *Gene*, **1**, 255–280.

232. Collins,J. and Hohn,B. (1978) Cosmids: a type of plasmid gene-cloning vector that is packageable *in vitro* in bacteriophage lambda heads. *Proc. Natl Acad. Sci. USA*, **75**, 4242–4246.

233. Muller-Hill,B. and Kania,J. (1974) Lac repressor can be fused to beta-galactosidase. *Nature*, **249**, 561–563.

234. Kalnins,A., Otto,K., Ruther,U. and Muller-Hill,B. (1983) Sequence of the lacZ gene of *Escherichia coli*. *EMBO J.*, **2**, 593–597.

235. Messing,J., Crea,R. and Seeburg,P.H. (1981) A system for shotgun DNA sequencing. *Nucleic Acids Res.*, **9**, 309–321.

236. Messing,J. and Vieira,J. (1982) A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene*, **19**, 269–276.

237. Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–119.

238. Roberts,T.M., Kacich,R. and Ptashne,M. (1979) A general method for maximizing the expression of a cloned gene. *Proc. Natl Acad. Sci. USA*, **76**, 760–764.

239. Guarente,L., Roberts,T.M. and Ptashne,M. (1980) A technique for expressing eukaryotic genes in bacteria. *Science*, **209**, 1428–1430.

240. Guarente,L., Lauer,G., Roberts,T.M. and Ptashne,M. (1980) Improved methods for maximizing expression of a cloned gene: a bacterium that synthesizes rabbit beta-globin. *Cell*, **20**, 543–553.

241. O'Farrell,P., Polisky,B. and Gelfand,D.H. (1978) Regulated expression by readthrough translation from a plasmid-encoded beta-galactosidase. *J. Bacteriol.*, **134**, 645–654.

242. Casadaban,M.J., Chou,J. and Cohen,S.N. (1980) *In vitro* gene fusions that join an enzymatically active beta-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. *J. Bacteriol.*, **143**, 971–980.

243. Fried,L., Lassak,J. and Jung,K. (2012) A comprehensive toolbox for the rapid construction of lacZ fusion reporters. *J. Microbiol. Methods.*, **91**, 537–543.

244. Linn,T. and Ralling,G. (1985) A versatile multiple- and single-copy vector system for the *in vitro* construction of transcriptional fusions to lacZ. *Plasmid*, **14**, 134–142.

245. Linn,T. and St,P.R. (1990) Improved vector system for constructing transcriptional fusions that ensures independent translation of lacZ. *J. Bacteriol.*, **172**, 1077–1084.

246. Pourcel,C., Marchal,C., Louise,A., Fritsch,A. and Tiollais,P. (1979) Bacteriophage lambda-*E. coli* K12 vector-host system for gene cloning and expression under lactose promoter control: I. DNA fragment insertion at the lacZ EcoRI restriction site. *Mol. Gen. Genet.*, **170**, 161–169.

247. Uhlich,G.A. and Chen,C.Y. (2012) A cloning vector for creation of *Escherichia coli* lacZ translational fusions and generation of linear template for chromosomal integration. *Plasmid*, **67**, 259–263.

248. Windle,B.E. (1986) Phage lambda and plasmid expression vectors with multiple cloning sites and lacZ alpha-complementation. *Gene*, **45**, 95–99.

249. Yu,D. and Court,D.L. (1998) A new system to place single copies of genes, sites and lacZ fusions on the *Escherichia coli* chromosome. *Gene*, **223**, 77–81.

250. Huynh,T.V., Young,R.A. and Davis,R.W. (1985) Constructing and screening cDNA libraries in lambda gt10 and lambda gt11. In: Glover,D.M. (ed.), *DNA Cloning*, Vol. 1. IRL Press, Oxford, U.K., pp. 49–78.

251. Ruther,U., Koenen,M., Sippel,A.E. and Muller-Hill,B. (1982) Exon cloning: immunoenzymatic identification of exons of the chicken lysozyme gene. *Proc. Natl Acad. Sci. USA*, **79**, 6852–6855.

252. Cupples,C.G. and Miller,J.H. (1988) Effects of amino acid substitutions at the active site in *Escherichia coli* beta-galactosidase. *Genetics*, **120**, 637–644.

253. Cupples,C.G. and Miller,J.H. (1989) A set of lacZ mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. *Proc. Natl Acad. Sci. USA*, **86**, 5345–5349.

254. Cupples,C.G., Cabrera,M., Cruz,C. and Miller,J.H. (1990) A set of lacZ mutations in *Escherichia coli* that allow rapid detection of specific frameshift mutations. *Genetics*, **125**, 275–280.

255. Gossen,J.A., de Leeuw,W.J. and Vijg,J. (1994) LacZ transgenic mouse models: their application in genetic toxicology. *Mutat. Res.*, **307**, 451–459.

256. Shwed,P.S., Crosthwait,J., Douglas,G.R. and Seligy,V.L. (2010) Characterisation of MutaMouse lambdagt10-lacZ transgene: evidence for *in vivo* rearrangements. *Mutagenesis*, **25**, 609–616.

257. Grodzicker,T., Williams,J., Sharp,P. and Sambrook,J. (1975) Physical mapping of temperature-sensitive mutations of adenoviruses. *Cold Spring Harb. Symp. Quant. Biol.*, **39(Pt 1)**, 439–446.

258. Hutchison,C.A. III, Newbold,J.E., Potter,S.S. and Edgell,M.H. (1974) Maternal inheritance of mammalian mitochondrial DNA. *Nature*, **251**, 536–538.

259. Potter,S.S., Newbold,J.E., Hutchison,C.A. III and Edgell,M.H. (1975) Specific cleavage analysis of mammalian mitochondrial DNA. *Proc. Natl Acad. Sci. USA*, **72**, 4496–4500.

260. Heller,R. and Arnheim,N. (1980) Structure and organization of the highly repeated and interspersed 1.3 kb EcoRI-Bg1II sequence family in mice. *Nucleic Acids Res.*, **8**, 5031–5042.

261. Jeffreys,A.J. (1979) DNA sequence variants in the G gamma-, A gamma-, delta- and beta-globin genes of man. *Cell*, **18**, 1–10.

262. Jeffreys,A.J., Craig,I.W. and Francke,U. (1979) Localisation of the G gamma-, A gamma-, delta- and beta-globin genes on the short arm of human chromosome 11. *Nature*, **281**, 606–608.

263. Kan,Y.W. and Dozy,A.M. (1978) Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. *Proc. Natl Acad. Sci. USA*, **75**, 5631–5635.

264. Kan,Y.W. and Dozy,A.M. (1978) Antenatal diagnosis of sickle-cell anaemia by D.N.A. analysis of amniotic-fluid cells. *Lancet*, **2**, 910–912.

265. Botstein,D., White,R.L., Skolnick,M. and Davis,R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, **32**, 314–331.

266. Gusella,J.F., Wexler,N.S., Conneally,P.M., Naylor,S.L., Anderson,M.A., Tanzi,R.E., Watkins,P.C., Ottina,K., Wallace,M.R., Sakaguchi,A.Y. *et al.* (1983) A polymorphic DNA marker genetically linked to Huntington's disease. *Nature*, **306**, 234–238.

267. Wexler,N.S. (2012) Huntington's disease: advocacy driving science. *Annu. Rev. Med.*, **63**, 1–22.

268. Gill,P., Jeffreys,A.J. and Werrett,D.J. (1985) Forensic application of DNA 'fingerprints'. *Nature*, **318**, 577–579.

269. Lindsay,J.A. (2010) Genomic variation and evolution of *Staphylococcus aureus*. *Int. J. Med. Microbiol.*, **300**, 98–103.

270. Otto,M. (2012) MRSA virulence and spread. *Cell Microbiol.*, **14**, 1513–1521.
271. Arber,W. (1977) What is the function of restriction enzymes? *Trends Biochem. Sci.*, **2**, N176–N178.
272. Ishikawa,K., Handa,N. and Kobayashi,I. (2009) Cleavage of a model DNA replication fork by a Type I restriction endonuclease. *Nucleic Acids Res.*, **37**, 3531–3544.
273. Ishikawa,K., Handa,N., Sears,L., Raleigh,E.A. and Kobayashi,I. (2011) Cleavage of a model DNA replication fork by a methyl-specific endonuclease. *Nucleic Acids Res.*, **39**, 5489–5498.
274. McNeil,E.M. and Melton,D.W. (2012) DNA repair endonuclease ERCC1-XPF as a novel therapeutic target to overcome chemoresistance in cancer therapy. *Nucleic Acids Res.*, **40**, 9990–10004.