248–262 *Nucleic Acids Research, 1997, Vol. 25, No. 1*

REBASE-restriction enzymes and methylases

Richard J. Roberts* and Dana Macelis

New England BioLabs, 32 Tozer Road, Beverly, MA 01915, USA

Received October 31, 1996; Accepted October 31, 1996

ABSTRACT

REBASE is a comprehensive database of information about restriction enzymes and their associated methylases, including their recognition and cleavage sites and their commercial availability. Information from REBASE is available via monthly electronic mailings as well as via anonymous ftp, WAIS/gopher and through the World Wide Web (http://www.neb.com/rebase). Specialized files are available that can be used directly by many software packages.

INTRODUCTION

The restriction enzyme database, REBASE, is a collection of information about restriction enzymes and methylases. Since the last description of the contents of REBASE (1), 131 new entries have been added. These include three Type II enzymes with new specificities and five new neoschizomers (isoschizomers that cleave at a position different from their prototype) shown in Table 1. Homing endonucleases (2) are also listed in REBASE and six of the 29 known are commercially available (Table 2). Among the 2821 restriction enzymes now known, there are 18 fully characterized Type I specificities, 210 different Type II specificities and four different Type III specificities. Table 2 contains a listing of all prototype restriction enzymes (Types I, II and III), together with their commercially available isoschizomers and neoschizomers. REBASE currently stores 3922 published references, with abstracts, and 446 unpublished observations.

REBASE is updated daily. Each month, a set of REBASE data files are released publically and distributed to the scientific community, at no charge, via e-mail. They can also be retrieved by anonymous ftp (www.neb.com, in pub/rebase), WAIS (www.neb.com, port 210) or through the World Wide Web (http://www.neb.com/rebase). These data files are ASCII text files, many of which are designed specifically for use with a variety of software packages such as GCG, IGSuite, GENEPRO, Staden, DNA Strider, Pro-Cite, PC/Gene, SEQAIDII, GENETYX, DNASIS, CAD Gene and Reference Manager. Other data files include a complete set of references, including abstracts, to papers on restriction enzymes and methylases; a list of all commercial suppliers of restriction enzymes and methylases, complete with contact information and a listing of enzymes they currently sell. New data files are constantly being added and each release of REBASE includes a monthly newsletter indicating that the files at the www and ftp sites have been updated, and listing new enzymes, newly available formats, enzyme name changes, etc. To join the mailing list or for more information, contact either author [e-mail: macelis@neb.com or roberts@neb.com; tel: +1 (508) 927-3382; fax +1 (508) 921-1527].

Since last year, we have added another commercial supplier of restriction enzymes, Advanced Biotechnologies Ltd, and REBASE now contains contact information for the distributors of each company listed.

Listed below are a number of name changes that have taken place during the last year.

Old name	**New name (Prototype)**	**Recognition sequence**
*Plo*214I	*Ple*214I (*Hae*III)	GGCC
*Eco*R124/3I	*Eco*R124II	GAANNNNNNNRTCG
*Sty*SBI	*Sty*LTIII	GAGNNNNNNRTAYG
*Rsp*519II	*Rma*495II (*Eco*RV)	GATATC
*Rsp*523I	*Rma*501I (*Mae*I)	CTAG
*Rsp*534I	*Rma*523I (*Asu*II)	TTCGAA
*Rsp*554I	*Uth*554I (*Ava*II)	GGWCC
*Rsp*556I	*Rma*376I (*Asu*II)	TTCGAA
*Tsp*515I	*Bsp*233I (*Bsp*MII)	TCCGGA

Microorganisms identified:

Old name	**New name (Prototype)**	**Recognition sequence**
*Uba*CI	*Tat*I	WGTACW
*Uth*506I	*Tsp*219I (*Bgl*I)	GCCNNNNNGGC
*Uth*511I	*Tsp*301I (*Ava*II)	GGWCC
*Uth*512I	*Tsp*358I (*Taq*I)	TCGA
*Uth*536I	*Bsp*H22I (*Asu*II)	TTCGAA
*Uth*537I	*Bsp*23I (*Hae*III)	GGCC
*Uth*558I	*Bsp*241I (*Asu*II)	TTCGAA
*Uth*559I	*Tsp*273I (*Eco*RV)	GATATC
*Uth*559II	*Tsp*273II (*Hae*III)	GGCC
*Uth*561I	*Tsp*266I (*Hae*III)	GGCC
*Uth*562I	*Tsp*281I (*Hae*III)	GGCC

In forming Tables 1 and 2, all endonucleases cleaving DNA at a specific sequence have been considered to be restriction enzymes, although in most cases there is no direct genetic evidence for the presence of a restriction–modification system. The endonucleases are named in accordance with the proposal of Smith and Nathans (3).

We request that authors wishing to cite REBASE use this article as a general reference for the complete listing that is available electronically.

---

* To whom correspondence should be addressed. Tel: +1 508 927 3382; Fax: +1 508 921 1527; Email: roberts@neb.com

Table 1.

Microorganism	Enzyme (prototype)	Recognition sequence	Reference
Bacillus circulans	*Bci*VI	GTATCC (6/5)	4
Stenotrophomonas maltophilia	*Sml*I	C↑TYRAG	5
Unidentified bacterium SS2-322	*Uba*EI	CACCTGC	7

Recognition sequences use the standard abbreviations [*Eur. J. Biochem.* (1985) **150**, 1–5] to represent ambiguity.
R = A or G
Y = C or T
M = A or C
K = G or T
S = G or C
W = A or T
B = not A (C or G or T)
D = not C (A or G or T)
H = not G (A or C or T)
V = not T (A or C or G)
N = A or C or G or T

Recognition sequences are written from 5′ to 3′, only one strand being given, and the point of cleavage is indicated by an arrow (↑). When no arrow appears, the precise cleavage site has not been determined. For enzymes such as *Bci*VI, which cleaves away from its recognition sequence, the sites of cleavage are indicated in parentheses. Thus GTATCC (6/5) indicates cleavage as shown below:

```
5′-GTATCCNNNNNN↑-3′
3′-CATAGGNNNNN↑-5′
```

In all cases the recognition sequences are oriented so that the cleavage sites lie on their 3′-side.

Table 2. Homing endonucleases

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
*Aat*II		GACGT↑C		ADEFKLMNOPRS
*Acc*I		GT↑MKAC	5(6)	ABDEGJKLMNOPQRS
*Ace*III		CAGCTC(7/11)		
*Aci*I		CCGC(-3/-1)		N
*Acl*I		AA↑CGTT		I
	*Psp*1406I	AA↑CGTT		DFMN
*Acy*I		GR↑CGYC		JM
	*Bbi*II	GR↑CGYC		K
	*Bsa*HI	GR↑CGYC		NS
	*Hin*1I	GR↑CGYC		EFO
	*Hsp*92I	GR↑CGYC		R
	*Msp*17I	GR↑CGYC		D
*Afl*II		C↑TTAAG		ABJKNOS
	*Bfr*I	C↑TTAAG		MO
	*Bsp*TI	C↑TTAAG		DF
	*Bst*98I	C↑TTAAG		R
	*Msp*CI	C↑TTAAG		C
	*Vha*464I	C↑TTAAG		I
*Afl*III		A↑CRYGT		ABMN
*Age*I		A↑CCGGT	?(5)	JNR
	*Asi*AI	A↑CCGGT		I
	*Pin*AI	A↑CCGGT		BM
*Aha*III		TTT↑AAA		
	*Dra*I	TTT↑AAA		ABDEFGIJKLMNOPQRS
*Alu*I		AG↑CT	3(5)	ABCDEFGHIJKLMNOPQRST
*Alw*NI		CAGNNN↑CTG		ABNS
*Apa*I		GGGCC↑C	4(5)	ABEGIJKLMNOPQRS
	*Bsp*120I	G↑GGCCC		DFN

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
	*Psp*OMI	G↑GGCCC		I
*Apa*BI		GCANNNNN↑TGC		
*Apa*LI		G↑TGCAC	?(5)	ADEKNS
	*Alw*44I	G↑TGCAC		FJMOR
	*Vne*I	G↑TGCAC		I
*Apo*I		R↑AATTY	?(6)	N
	*Acs*I	R↑AATTY		DIM
*Asc*I		GG↑CGCGCC	?(5)	N
*Asu*I		G↑GNCC		
	*Asp*S9I	G↑GNCC		I
	*Bsi*ZI	G↑GNCC		T
	*Cfr*13I	G↑GNCC	4(5)	DFKO
	*Fmu*I	GGNC↑C		
	*Sau*96I	G↑GNCC	4(5)	AJLMNORS
*Asu*II		TT↑CGAA		
	*Bpu*14I	TT↑CGAA		I
	*Bsi*CI	TT↑CGAA		T
	*Bsp*119I	TT↑CGAA		DF
	*Bst*BI	TT↑CGAA		NS
	*Cbi*I	TT↑CGAA		J
	*Csp*45I	TT↑CGAA		OR
	*Lsp*I	TT↑CGAA		L
	*Nsp*V	TT↑CGAA		ABJKO
	*Sfu*I	TT↑CGAA		M
*Ava*I		C↑YCGRG	?(4)	ABEGJKLMNOPQRS
	*Ama*87I	C↑YCGRG		DI
	*Bco*I	C↑YCGRG		AT
	*Bso*BI	C↑YCGRG	?(4)	N
	*Eco*88I	C↑YCGRG		F
	*Nli*387/7I	CYCGR↑G		
*Ava*II		G↑GWCC	?(5)	ABEGJKMNPQRS
	*Bme*18I	G↑GWCC		I
	*Eco*47I	G↑GWCC		FO
	*Hgi*EI	G↑GWCC	?(5)	D
	*Sin*I	G↑GWCC	4(5)	LR
	*Vpa*K11AI	↑GGWCC		
*Ava*III		ATGCAT		
	*Bfr*BI	ATG↑CAT		
	*Eco*T22I	ATGCA↑T		AKO
	*Mph*1103I	ATGCA↑T		F
	*Nsi*I	ATGCA↑T		BDEHJLMNRS
	*Ppu*10I	A↑TGCAT		DFN
	*Zsp*2I	ATGCA↑T		I
*Avr*II		C↑CTAGG		N

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
	*Bln*I	C↑CTAGG		AKM
	*Bsp*A2I	C↑CTAGG		I
*Bae*I[e]		(10/15)ACNNNNGTAYC(12/7)	1(6),–4(6)	
*Bal*I		TGG↑CCA	4(4)	AJKR
	*Mlu*NI	TGG↑CCA		M
	*Msc*I	TGG↑CCA		BDNOS
*Bam*HI		G↑GATCC	5(4)	ABCDEFGHIJKLMNOPQRST
*Bbv*I		GCAGC(8/12)	2(5),–2(5)	N
	*Bst*71I	GCAGC(8/12)		R
*Bbv*II		GAAGAC(2/6)		
	*Bbs*I	GAAGAC(2/6)		N
	*Bpi*I	GAAGAC(2/6)		DF
	*Bpu*AI	GAAGAC(2/6)		M
*Bcc*I		CCATC		
*Bce*83I		CTTGAG(16/14)		
*Bce*fI		ACGGC(12/13)		
*Bcg*I[e]		(10/12)GCANNNNNNTCG(12/10)	3(6),–3(6)	N
*Bci*VI		GGATAC		
*Bcl*I		T↑GATCA		BCDEFGJLMNOPRS
	*Bsi*QI	T↑GATCA		T
	*Fba*I	T↑GATCA		AK
	*Ksp*22I	T↑GATCA		I
*Bet*I		W↑CCGGW		
	*Bsa*WI	W↑CCGGW		N
*Bfi*I		ACTGGG(4/11)		F
*Bgl*I		GCCNNNN↑NGGC	?(4)	ABCDEFGHIJKLMNOPQRST
*Bgl*II		A↑GATCT	5(4)	ABCDEFGHIJKLMNOPQRST
*Bin*I		GGATC(4/5)		
	*Acl*WI	GGATC(4/5)		I
	*Alw*I	GGATC(4/5)		NS
*Bmg*I		GKGCCC		
*Bpl*I		GAGNNNNNCTC		
*Bpu*10I		CCTNAGC(–5/–2)		
*Bsa*AI		YAC↑GTR		N
	*Bst*BAI	YAC↑GTR		I
*Bsa*BI		GATNN↑NNATC		N
	*Bse*8I	GATNN↑NNATC		I
	*Bsh*1365I	GATNN↑NNATC		DF
	*Bsi*BI	GATNN↑NNATC		T
	*Bsr*BRI	GATNN↑NNATC		R
	*Mam*I	GATNN↑NNATC	?(6)	M
*Bsa*XI		ACNNNNNCTCC		
*Bsb*I		CAACAC		
*Bsc*GI		CCCGT		

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
*Bse*PI		G↑CGCGC		I
	*Bss*HII	G↑CGCGC	?(5)	ABDEJKLMNOPQR
	*Pau*I	G↑CGCGC		F
*Bse*RI		GAGGAG(10/8)		N
*Bsg*I		GTGCAG(16/14)		N
*Bsi*I		CACGAG(–5/–1)		
	*Bss*SI	CACGAG(–5/–1)		N
	*Bst*2BI	CACGAG(–5/–1)		DI
*Bsi*YI		CCNNNNN↑NNGG		MT
	*Bsc*4I	CCNNNNN↑NNGG		I
	*Bsl*I	CCNNNNN↑NNGG		N
*Bsm*I		GAATGC(1/–1)		ABDEJLMNOS
	*Bsa*MI	GAATGC(1/–1)		R
	*Bsc*CI	GAATGC(1/–1)		T
	*Mva*1269I	GAATGC(1/–1)		F
*Bsm*AI		GTCTC(1/5)		N
	*Alw*26I	GTCTC(1/5)	3(5),–4(6)	DFNR
*Bsp*24I[e]		(8/13)GACNNNNNNTGG(12/7)		
*Bsp*1407I		T↑GTACA		DF
	*Bsr*GI	T↑GTACA		N
	*Ssp*BI	T↑GTACA		M
*Bsp*GI		CTGGAC		
*Bsp*HI		T↑CATGA		AN
	*Rca*I	T↑CATGA		BM
*Bsp*LU11I		A↑CATGT		M
*Bsp*MI		ACCTGC(4/8)		N
*Bsp*MII		T↑CCGGA		
	*Acc*III	T↑CCGGA		EJKQR
	*Bse*AI	T↑CCGGA		CM
	*Bsi*MI	T↑CCGGA		AT
	*Bsp*13I	T↑CCGGA		I
	*Bsp*EI	T↑CCGGA		N
	*Kpn*2I	T↑CCGGA		BDF
	*Mro*I	T↑CCGGA		MO
*Bsr*I		ACTGG(1/–1)		N
	*Bse*1I	ACTGG(1/–1)		I
	*Bse*NI	ACTGG(1/–1)		DF
	*Bsr*SI	ACTGG(1/–1)		R
*Bsr*BI		CCGCTC(–3/–3)		N
	*Acc*BSI	CCGCTC(–3/–3)		DI
*Bsr*DI		GCAATG(2/0)		N
	*Bse*3DI	GCAATG(2/0)		I
*Bst*EII		G↑GTNACC		BDGHJLMNORS
	*Bst*PI	G↑GTNACC		K

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
	*Eco*91I	G↑GTNACC		F
	*Eco*O65I	G↑GTNACC		AK
	*Psp*EI	G↑GTNACC		I
*Bst*XI		CCANNNNN↑NTGG		ABDEFGHJKLMNOQRS
*Cac*8I		GCN↑NGC		N
*Cau*II		CC↑SGG		
	*Asu*C2I	CC↑SGG		I
	*Bcn*I	CC↑SGG	2(4)	ADFK
	*Eco*HI	↑CCSGG		
	*Nci*I	CC↑SGG		BEJLMNORST
*Cfr*I		Y↑GGCCR	4(5)	F
	*Eae*I	Y↑GGCCR	4(5)	ADKLMNS
*Cfr*10I		R↑CCGGY	2(5)	ADFKMO
	*Bse*118I	R↑CCGGY		I
	*Bsr*FI	R↑CCGGY		N
	*Bss*AI	R↑CCGGY		C
*Cje*I[e]		(8/14)CCANNNNNNGT(15/9)		
*Cje*PI[e]		(7/13)CCANNNNNNNTC(14/8)		
*Cla*I		AT↑CGAT	5(6)	ABKMNPRST
	*Ban*III	AT↑CGAT		O
	*Bsa*29I	AT↑CGAT		I
	*Bsc*I	AT↑CGAT		L
	*Bse*CI	AT↑CGAT	5(6)	C
	*Bsi*XI	AT↑CGAT		T
	*Bsp*106I	AT↑CGAT	5(6)	E
	*Bsp*DI	AT↑CGAT		N
	*Bsp*XI	AT↑CGAT		G
	*Bsu*15I	AT↑CGAT	5(6)	DF
*Cvi*JI		RG↑CY	3(5)	Q
*Cvi*RI		TG↑CA	4(6)	
*Dde*I		C↑TNAG	1(5)	ABEGLMNOR
	*Bst*DEI	C↑TNAG		DI
*Dpn*I[*]		GmA↑TC		ABLMNRS
*Dra*II		RG↑GNCCY		AGMS
	*Eco*O109I	RG↑GNCCY		ABDEFJKLN
	*Pss*I	RGGNC↑CY		
*Dra*III		CACNNN↑GTG		AEGMNS
*Drd*I		GACNNNN↑NNGTC		N
	*Dse*DI	GACNNNN↑NNGTC		I
*Drd*II		GAACCA		
*Dsa*I		C↑CRYGG		M
	*Bst*DSI	C↑CRYGG		DI
*Eam*1105I		GACNNN↑NNGTC		DFK
	*Ahd*I	GACNNN↑NNGTC		N

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
	*Asp*EI	GACNNN↑NNGTC		M
	*Ecl*HKI	GACNNN↑NNGTC		R
	*Nru*GI	GACNNN↑NNGTC		I
*Eci*I		TCCGCC		
*Eco*31I		GGTCTC(1/5)	4(5),–4(6)	DF
	*Bsa*I	GGTCTC(1/5)		N
*Eco*47III		AGC↑GCT		BDEFLMNOR
	*Afe*I	AGC↑GCT		I
	*Aor*51HI	AGC↑GCT		AK
*Eco*57I		CTGAAG(16/14)	5(6),–5(6)	DFNS
*Eco*NI		CCTNN↑NNNAGG		ANS
*Eco*RI		G↑AATTC	3(6)	ABCDEFGHIJKLMNOPQRST
*Eco*RII[f]		↑CCWGG	2(5)	BEJMOS
+	*Bsi*LI	CC↑WGG		T
+	*Bst*NI	CC↑WGG	2(4)	CENS
+	*Bst*OI	CC↑WGG		R
+	*Bst*2UI	CC↑WGG		I
+	*Mva*I	CC↑WGG	2(4)	ADFKMO
*Eco*RV		GAT↑ATC	2(6)	ABCDEGHIJKLMNOPQRST
	*Eco*32I	GAT↑ATC		F
*Esp*I		GC↑TNAGC		
	*Blp*I	GC↑TNAGC		N
	*Bpu*1102I	GC↑TNAGC		BDEFK
	*Bsp*1720I	GC↑TNAGC		I
	*Cel*II	GC↑TNAGC		ALM
*Esp*3I		CGTCTC(1/5)	4(5),–4(6)	DF
	*Bsm*BI	CGTCTC(1/5)		N
*Fau*I		CCCGC(4/6)		I
*Fin*I		GGGAC		
	*Bsm*FI	GGGAC(10/14)		N
*Fnu*DII		CG↑CG		
	*Acc*II	CG↑CG		AJKQ
	*Bsh*1236I	CG↑CG		DEF
	*Bst*UI	CG↑CG		NS
	*Mvn*I	CG↑CG		M
	*Sel*I	↑CGCG		
	*Tha*I	CG↑CG		B
*Fnu*4HI		GC↑NGC		N
	*Bso*FI	GC↑NGC		N
	*Fsp*4HI	GC↑NGC		I
	*Ita*I	GC↑NGC		M
*Fok*I		GGATG(9/13)	3(6),–2(6)	ADGIJKLMNRS
	*Bst*F5I	GGATG(2/0)		I
	*Sts*I	GGATG(10/14)	3(6),–2(6)	

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
*Fse*I		GGCCGG↑CC	?(5)	AKN
*Gdi*II		CGGCCR(–5/–1)		
*Gsu*I		CTGGAG(16/14)		DFS
	*Bpm*I	CTGGAG(16/14)		N
*Hae*I		WGG↑CCW		
*Hae*II		RGCGC↑Y		ABDEGJKLMNORS
	*Bsp*143II	RGCGC↑Y		F
	*Bst*H2I	RGCGC↑Y		I
	*Lpn*I	RGC↑GCY		
*Hae*III		GG↑CC	3(5)	ABCDGHIJKLMNOPQRST
	*Bsh*I	GG↑CC		T
	*Bsu*RI	GG↑CC	3(5)	FI
	*Pal*I	GG↑CC		E
*Hga*I		GACGC(5/10)	3(5),–2(5)	AN
*Hgi*AI		GWGCW↑C		
	*Alw*21I	GWGCW↑C		DF
	*Asp*HI	GWGCW↑C		M
	*Bbv*12I	GWGCW↑C		I
	*Bsi*HKAI	GWGCW↑C		N
*Hgi*CI		G↑GYRCC	5(5)	
	*Acc*B1I	G↑GYRCC		I
	*Ban*I	G↑GYRCC	?(5)	AEMNORS
	*Bsh*NI	G↑GYRCC		F
	*Eco*64I	G↑GYRCC		DF
*Hgi*EII		ACCNNNNNNGGT		
*Hgi*JII		GRGCY↑C		
	*Ban*II	GRGCY↑C	4(5)	ABKLMNOQRS
	*Eco*24I	GRGCY↑C		DF
	*Eco*T38I	GRGCY↑C		J
	*Fri*OI	GRGCY↑C		I
*Hha*I		GCG↑C	2(5)	ABGJKNOPRS
	*Asp*LEI	GCG↑C		I
	*Cfo*I	GCG↑C		BLMRS
	*Hin*6I	G↑CGC		DF
	*Hin*P1I	G↑CGC		N
	*Hsp*AI	G↑CGC		I
*Hin*4I		GABNNNNNVTC		
*Hin*dII		GTY↑RAC	5(6)	EM
	*Hin*cII	GTY↑RAC	5(6)	ABCDEFGHJKLNOPQRS
*Hin*dIII		A↑AGCTT	1(6)	ABCDEFGHIJKLMNOPQRST
*Hin*fI		G↑ANTC		ABDEFGHIJKLMNOPQRST
*Hpa*I		GTT↑AAC	5(6)	ABDEFGJKLMNOPQRST
	*Bst*HPI	GTT↑AAC		I
*Hpa*II		C↑CGG	2(5)	BDEFGILMNOPQRS

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
	*Bsi*SI	C↑CGG		C
	*Hap*II	C↑CGG	2(5)	AK
	*Hin*2I	C↑CGG		F
	*Msp*I	C↑CGG	1(5)	ABDEFGHIJKLMNOPQRS
*Hph*I		GGTGA(8/7)	5(6),–2(5)	AFNS
	*Asu*HPI	GGTGA(8/7)		I
*Kpn*I		GGTAC↑C	4(6)	ABCDEFGHIJKLMNOPQRST
	*Acc*65I	G↑GTACC		DFINR
	*Asp*718I	G↑GTACC		M
*Ksp*632I		CTCTTC(1/4)		M
	*Bsu*6I	CTCTTC(1/4)		I
	*Eam*1104I	CTCTTC(1/4)		DEF
	*Ear*I	CTCTTC(1/4)		N
*Mae*I		C↑TAG		M
	*Bfa*I	C↑TAG		N
*Mae*II		A↑CGT		M
	*Tai*I	ACGT↑		FN
	*Tsc*I	ACGT↑		L
*Mae*III		↑GTNAC		MP
*Mbo*I[g]		↑GATC	2(6)	ABCEFGKLNQRST
	*Bsc*FI	GATC		T
+	*Bsp*143I	↑GATC		DF
+	*Bsp*KT6I	GAT↑C	2(6)	
	*Cha*I	GATC↑		
	*Dpn*II	↑GATC	2(6)	N
	*Kzo*9I	↑GATC		I
	*Nde*II	↑GATC		BDGJM
+	*Sau*3AI	↑GATC	4(5)	ABEGHJKLMNOPQRS
*Mbo*II		GAAGA(8/7)	5(6)	ABDFGJKNOQRS
*Mcr*I		CGRY↑CG		
	*Bsa*OI	CGRY↑CG		R
	*Bsh*1285I	CGRY↑CG		DF
	*Bsi*EI	CGRY↑CG		N
	*Bst*MCI	CGRY↑CG		I
*Mfe*I		C↑AATTG		N
	*Mun*I	C↑AATTG	3(6)	BDEFKM
*Mlu*I		A↑CGCGT		ABDEFGIJKLMNOPQRS
*Mme*I		TCCRAC(20/18)		
*Mnl*I		CCTC(7/6)		AEFNQ
*Mse*I		T↑TAA		BN
	*Tru*1I	T↑TAA		F
	*Tru*9I	T↑TAA		DILMRS
*Msl*I		CAYNN↑NNRTG		N
*Mst*I		TGC↑GCA		

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
	*Acc*16I	TGC↑GCA		DI
	*Avi*II	TGC↑GCA		M
	*Fsp*I	TGC↑GCA		ABJKNOS
*Mwo*I		GCNNNNN↑NNGC	?(5)	N
*Nae*I		GCC↑GGC		ADEKLMNOR
	*Mro*NI	G↑CCGGC		I
	*Ngo*AIV	G↑CCGGC		B
	*Ngo*MI	G↑CCGGC	2(5)	NR
*Nar*I		GG↑CGCC		BEJMNOPRS
	*Bbe*I	GGCGC↑C		AK
	*Ege*I	GGC↑GCC		I
	*Ehe*I	GGC↑GCC		ADFNO
	*Kas*I	G↑GCGCC		N
	*Mly*113I	GG↑CGCC		I
*Nco*I		C↑CATGG	?(4)	ABCDEFGHJKLMNOPQRST
	*Bsp*19I	C↑CATGG		I
*Nde*I		CA↑TATG		ABDEFGJKLMNPRS
	*Fau*NDI	CA↑TATG		DI
*Nhe*I		G↑CTAGC		ABDEFGJKLMNOPRS
	*Ace*II	GCTAG↑C		
	*Asu*NHI	G↑CTAGC		I
*Nla*III		CATG↑	2(6)	ANOS
	*Cvi*AII	C↑ATG	2(6)	
	*Hsp*92II	CATG↑		R
*Nla*IV		GGN↑NCC		NS
	*Bsc*BI	GGN↑NCC		T
	*Bsp*LI	GGN↑NCC		F
	*Psp*N4I	GGN↑NCC		I
*Not*I		GC↑GGCCGC	?(4)	ABCDEFGJKLMNOPQRST
	*Cci*NI	GC↑GGCCGC		I
*Nru*I		TCG↑CGA		ABCDEGIJKLMNOQRST
	*Bsp*68I	TCG↑CGA		F
*Nsp*I		RCATG↑Y		ABKM
	*Bst*NSI	RCATG↑Y		DI
*Nsp*BII		CMG↑CKG		A
	*Msp*A1I	CMG↑CKG		INR
*Pac*I		TTAAT↑TAA		NO
*Pfl*1108I		TCGTAG		
*Pfl*MI		CCANNNN↑NTGG		ANS
	*Acc*B7I	CCANNNN↑NTGG		IR
	*Esp*1396I	CCANNNN↑NTGG		F
	*Van*91I	CCANNNN↑NTGG		DFKM
*Ple*I		GAGTC(4/5)		N
	*Mly*I	GACTC(5/5)		

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
*Pma*CI		CAC↑GTG		AK
	*Bbr*PI	CAC↑GTG		MO
	*Eco*72I	CAC↑GTG		DEFR
	*Pml*I	CAC↑GTG		N
*Pme*I		GTTT↑AAAC		N
*Ppu*MI		RG↑GWCCY		ANO
	*Psp*5II	RG↑GWCCY		BDF
	*Psp*PPI	RG↑GWCCY		I
*Psh*AI		GACNN↑NNGTC		AKN
*Pst*I		CTGCA↑G	5(6)	ABCDEFGHIJKLMNOPQRST
*Pvu*I		CGAT↑CG		ABDEFGKLMNOPQRS
	*Bsp*CI	CGAT↑CG		E
	*Ple*19I	CGAT↑CG		I
*Pvu*II		CAG↑CTG	4(4)	ABCDEFGHIJKLMNOPQRST
*Rle*AI		CCCACA(12/9)		
*Rsa*I		GT↑AC		ABCDEFGIJLMNOPQRST
	*Afa*I	GT↑AC		K
	*Csp*6I	G↑TAC		DFN
*Rsr*II		CG↑GWCCG		BMNS
	*Cpo*I	CG↑GWCCG		ADFK
	*Csp*I	CG↑GWCCG		EOR
*Sac*I		GAGCT↑C	5(?)	ACEFGHJKLMNOPQRST
	*Ecl*136II	GAG↑CTC		DFN
	*Eco*ICRI	GAG↑CTC		R
	*Psp*124BI	GAGCT↑C		DI
	*Sst*I	GAGCT↑C		BCS
*Sac*II		CCGC↑GG		ACEGHJKLNOPQRST
	*Cfr*42I	CCGC↑GG		DF
	*Ksp*I	CCGC↑GG		M
	*Mlu*113I	CC↑GCGG		
	*Sfr*303I	CCGC↑GG		I
	*Sst*II	CCGC↑GG		BS
*Sal*I		G↑TCGAC	5(6)	ABCDEFGHIJKLMNOPQRST
*San*DI		GG↑GWCCC		E
*Sap*I		GCTCTTC(1/4)		N
*Sau*I		CC↑TNAGG		
	*Aoc*I	CC↑TNAGG		M
	*Axy*I	CC↑TNAGG		J
	*Bse*21I	CC↑TNAGG		I
	*Bsu*36I	CC↑TNAGG		ENRS
	*Cvn*I	CC↑TNAGG		B
	*Eco*81I	CC↑TNAGG		ADFKO
*Sca*I		AGT↑ACT		ABCDEFGJKLMNOPQRS
	*Acc*113I	AGT↑ACT		I

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
	*Eco*255I	AGT↑ACT		F
*Scr*FI		CC↑NGG	?(5)	DJMNOS
	*Bss*KI	↑CCNGG		N
	*Msp*R9I	CC↑NGG		I
*Sdu*I		GDGCH↑C		F
	*Bmy*I	GDGCH↑C		M
	*Bsp*1286I	GDGCH↑C		ADJKNR
*Sec*I		C↑CNNGG		
	*Bsa*JI	C↑CNNGG		N
	*Bse*DI	C↑CNNGG		DF
*Sex*AI		A↑CCWGGT		M
*Sfa*NI		GCATC(5/9)		IN
	*Bsc*AI	GCATC(4/6)		
*Sfe*I		C↑TRYAG		
	*Bfm*I	C↑TRYAG		F
	*Bst*SFI	C↑TRYAG		DI
	*Sfc*I	C↑TRYAG		N
*Sfi*I		GGCCNNNN↑NGGCC	?(4)	ABCDEGIJKLMNOPQRST
*Sgf*I		GCGAT↑CGC		R
*Sgr*AI		CR↑CCGGYG		M
*Sim*I		GGGTC(-3/0)		
*Sma*I		CCC↑GGG	2(4)	ABCDEFGHIJKLMNOPQRST
	*Cfr*9I	C↑CCGGG	2(4)	FO
	*Psp*AI	C↑CCGGG		E
	*Xma*I	C↑CCGGG	?(4)	ADINPRS
	*Xma*CI	C↑CCGGG		M
*Sml*I		CTYRAG		
*Sna*I		GTATAC		
	*Bst*1107I	GTA↑TAC		DFKMN
	*Bst*Z17I	GTATAC		N
*Sna*BI		TAC↑GTA		ACEKLMNRS
	*Bst*SNI	TAC↑GTA		DI
	*Eco*105I	TAC↑GTA		FO
*Spe*I		A↑CTAGT		ABDEHJKLMNOPQRST
	*Acl*NI	A↑CTAGT		I
*Sph*I		GCATG↑C		ABCDEGHIJKLMNOPQRST
	*Bbu*I	GCATG↑C		R
	*Pae*I	GCATG↑C		F
*Spl*I		C↑GTACG		AK
	*Bsi*WI	C↑GTACG		MNO
	*Pfl*23II	C↑GTACG		DF
	*Psp*LI	C↑GTACG		I
	*Sun*I	C↑GTACG		B
*Srf*I		GCCC↑GGGC		EO

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
*Sse*8387I		CCTGCA↑GG		AK
	*Sbf*I	CCTGCA↑GG		DI
	*Sda*I	CCTGCA↑GG		F
*Sse*8647I		AG↑GWCCT		
*Ssp*I		AAT↑ATT	?(6)	ABCDEFIJKLMNOQR
*Stu*I		AGG↑CCT		ABEJKLMNPQRST
	*Aat*I	AGG↑CCT		O
	*Eco*147I	AGG↑CCT		DF
	*Pme*55I	AGG↑CCT		I
	*Sse*BI	AGG↑CCT		C
*Sty*I		C↑CWWGG		BCEJMNRS
	*Bss*T1I	C↑CWWGG		I
	*Eco*130I	C↑CWWGG		DF
	*Eco*T14I	C↑CWWGG		AK
	*Erh*I	C↑CWWGG		I
*Swa*I		ATTT↑AAAT		M
	*Smi*I	ATTT↑AAAT		DI
*Taq*I		T↑CGA	4(6)	ABCDEFGJLMNOPQRST
	*Tth*HB8I	T↑CGA	4(6)	AK
*Taq*II[h]		CACCCA(11/9)		
*Taq*II[h]		GACCGA(11/9)		
*Tat*I		WGTACW		
*Tau*I		GCSGC		
*Tfi*I		G↑AWTC		N
*Tse*I		G↑CWGC		N
*Tsp*45I		↑GTSAC		N
*Tsp*4CI		ACN↑GT		
*Tsp*EI		↑AATT		LO
	*Sse*9I	↑AATT	2(6)	DI
	*Tsp*509I	↑AATT		N
*Tsp*RI		CAGTG(2/-7)		N
*Tth*111I		GACN↑NNGTC		ADIKNQR
	*Asp*I	GACN↑NNGTC		M
	*Pfl*FI	GACNNNGTC		N
*Tth*111II		CAARCA(11/9)		
*Uba*DI		GAACNNNNNNTCC		
*Uba*EI		CACCTGC		
*Vsp*I		AT↑TAAT	5(6)	BDEFIR
	*Ase*I	AT↑TAAT		AJNO
	*Asn*I	AT↑TAAT		MS
	*Psh*BI	AT↑TAAT		K
*Xba*I		T↑CTAGA	6(6)	ABCDEFGHIJKLMNOPQRST
*Xcm*I		CCANNNNN↑NNNNTGG		AN
*Xho*I		C↑TCGAG		ABCDEFGHJKLMNOPQRST

Table 2. *continued*

Prototype[a]	Isoschizomers	Recognition sequence[b]	Me site[c]	Commercial source[d]
	*Pae*R7I	C↑TCGAG	5(6)	N
	*Sci*I	CTC↑GAG		
	*Sfr*274I	C↑TCGAG		I
*Xho*II		R↑GATCY		EMR
	*Bst*X2I	R↑GATCY		DI
	*Bst*YI	R↑GATCY		BN
	*Mfl*I	R↑GATCY		AK
*Xma*III		C↑GGCCG	4(5)	BE
	*Bst*ZI	C↑GGCCG		R
	*Eag*I	C↑GGCCG	?(5)	NPS
	*Ecl*XI	C↑GGCCG		M
	*Eco*52I	C↑GGCCG		ADFKOR
*Xmn*I		GAANN↑NNTTC		DENR
	*Asp*700I	GAANN↑NNTTC		M
	*Mro*XI	GAANN↑NNTTC		I

Type I enzymes

Prototype enzyme	Recognition sequence[b]	Me site[c]
*Cfr*AI	GCANNNNNNNNGTGG	
*Eco*AI	GAGNNNNNNNGTCA	2(6),–3(6)
*Eco*BI	TGANNNNNNNNTGCT	3(6),–4(6)
*Eco*DI	TTANNNNNNNGTCY	
*Eco*DR	TCANNNNNNGTCG	
*Eco*DR3	TCANNNNNNNATCG	
*Eco*DXXI	TCANNNNNNNRTTC	
*Eco*EI	GAGNNNNNNNATGC	
*Eco*KI	AACNNNNNNGTGC	2(6),–3(6)
*Eco*R124I	GAANNNNNNRTCG	3(6),–3(6)
*Eco*R124II	GAANNNNNNNRTCG	–3(6)
*Eco*RD2	GAANNNNNNRTTC	
*Eco*RD3	GAANNNNNNNRTTC	
*Eco*prrI	CCANNNNNNNRTGC	
*Sty*LTIII	GAGNNNNNNRTAYG	2(6),–4(6)
*Sty*SJ	GAGNNNNNNGTRC	
*Sty*SPI	AACNNNNNNGTRC	2(6),–3(6)
*Sty*SQ	AACNNNNNNRTAYG	

Type III enzymes

Prototype enzyme	Recognition sequence[b]	Me site[c]
*Eco*PI	AGACC	3(6)
*Eco*P15I	CAGCAG(25/27)	5(6)
*Hinf*III	CGAAT	
*Sty*LTI	CAGAG	4(6)

Homing endonucleases

Enzyme	Sequence around one cleavage site[i]	Commercial source[d]
I-CeuI	CTATAACGGTC_CTAA↑GGTAGCGAGG	N
I-PpoI	CTATGACTCTC_TTAA↑GGTAGCCAAAA	NR
I-SceI	TTACGCTAGGG_ATAA↑CAGGGTAATA	M
PI-PspI	AAAATCCTGGCAAACAGCTATTATGGGTAT	N
PI-SceI	ATCTATGTCGG_GTGC↑GGAGAAAGAGGTAAT	N
PI-TliI	TCTTTATGCGG_ACAG↑TGACGGCTTTTAT	N

[a] *Signifies that *Dpn*I requires the presence of 6-methyladenosine within the recognition sequence GmATC.

[b] Recognition sequence nomenclature is described in the legend to Table 1.

[c] The site of methylation by the cognate methylase when known is indicated as follows. The first number shows the base within the recognition sequence that is modified. A negative number indicates the complementary strand, numbered from the 5′ base of that strand. The number in parentheses indicates the specific methylation involved. (6), N6-methyladenosine; (5), 5-methylcytosine; (4), N4-methylcytosine. Thus for *Alw*26I the designation 3(5),–4(6) signifies the following pattern of methylation:

```
          1     2     3     4     5
   5′     G     T   5mC     T     C     3′
   3′     C   6mA     G     A     G     5′
         -1    -2    -3    -4    -5
```

[d] Commercial sources of restriction enzymes are abbreviated as follows: A, Amersham Life Sciences-USB (10/96); B, Life Technologies Inc, Gibco-BRL (10/96); C, Minotech Molecular Biology Products (10/96); D, Angewandte Gentechnologie Systeme (10/96); E, Stratagene (10/96); F, Fermentas MBI (10/96); G, Appligene Oncor (10/96); H, American Allied Biochemical, Inc. (10/96); I, SibEnzyme Ltd (10/96); J, Nippon Gene Co., Ltd (10/96); K, Takara Shuzo Co. Ltd (10/96); L, NBL Gene Sciences Limited (10/96); M, Boehringer-Mannheim (7/96); N, New England BioLabs (10/96); O, Toyobo Biochemicals (10/96); P, Pharmacia Biotech Inc. (10/96); Q, CHIMERx (5/96); R, Promega Corporation (10/96); S, Sigma (10/96); T, Advanced Biotechnologies Ltd (10/96). The date in parentheses indicates the most recent update.

[e] Enzymes that cut on both sides of their recognition sequence, such as *Bae*I, *Bcg*I, *Bsp*24I, *Cje*I and *Cje*PI, have four cleavage sites each instead of two. For example, *Bsp*24I cleaves as indicated:

```
        5′ ↑NNNNNNNNGACNNNNNNTGGNNNNNNNNNNNN↑ 3′
   3′ ↑NNNNNNNNNNNNNCTGNNNNNNACCNNNNNNN↑ 5′
```

This is described in Table 2 and some REBASE reports as *Bsp*24I (8/13)GACNNNNNNTGG(12/7). Thus the recognition site is excised in a fragment, 32 base pairs long, with five-base 3′-extensions at each end.

[f] *Eco*RII isoschizomers fall into two classes based upon their sensitivity to methylation. *Eco*RII will not cleave when the second cytosine in the recognition sequence is methylated to 5-methylcytosine whereas *Mva*I will cleave such a sequence. Isoschizomers of *Eco*RII that are like *Mva*I are indicated by +.

[g] *Mbo*I isoschizomers fall into two classes based upon their sensitivity to methylation. *Mbo*I will not cleave when the recognition sequence contains 6-methyladenosine whereas *Sau*3AI will not cleave when its recognition sequence contains 5-methylcytosine. Isoschizomers of *Mbo*I that are like *Sau*3AI are indicated by +.

[h] *Taq*II differs from other restriction enzymes in recognizing two distinct sequences: GACCG and CACCCA.

[i] For the homing endonucleases the cleavage sites, where known, are indicated by ↑ on the strand presented and by _ on the complementary strand.

ACKNOWLEDGEMENTS

Special thanks are due to the many individuals who have so kindly communicated their unpublished results for inclusion in this compilation. This database is supported by the National Library of Medicine (LM04971).

REFERENCES

1 Roberts,R.J. and Macelis,D. (1996) *Nucleic Acids Res*. **24**, 223–235.
2 Mueller,J.E., Bryk,M., Loizes,N. and Belfort,M. (1993) in Linn,S.M., Lloyd,R.S. and Roberts,R.J. (eds), *Nucleases*, 2nd Ed. Cold Spring Harbor Laboratory Press, pp. 111–143.
3 Smith,H.O. and Nathans,D.J. (1973) *Mol. Biol*. **81**, 419–423.
4 Le,T.K.T., Vu,T.K.L., Vu,H.N., Polisson,C. and Morgan,R., unpublished observations.
5 Le,T.K.T., Vu,H.N., Vu,T.K.L., Polisson,C. and Morgan,R. unpublished observations.
6 Grigaite,R., Lazareviciute,L., Kiuduliene,L., Trinkunaite,L., Butkus,V. and Janulaitis,A., unpublished observations.