

1. **Restriction endonucleases:** एक *endonuclease* एक enzyme है जो nucleic acid के phosphodiester bonds को एक internal site (एक *exonuclease* से नहीं काटा जाता, जो केवल दोनों सिरों में से एक सिरा के nucleotides को काटता है) कुछ endonucleases DNAs और RNAs के internal bond (अन्तः बन्धों) को randomly (यादृच्छिक) काटता है। हालांकि, **restriction endonucleases specific** (विशिष्ट) restriction sites पर dsDNA के दोनों strands को काटते हैं। Restriction endonucleases कई प्रकार के हैं और प्रत्येक restriction endonucleases एक restriction site के लिए अत्यधिक विशिष्ट हैं, जो ज्यादातर 4, 6 या 8 base pairs के होते हैं। कुछ अपवाद हैं जो बाद में discuss किये जायेंगे।

Three types of restriction endonucleases

Type I

- Type I enzymes को सबसे पहले खोजा गया था। इनकी recognition site अलग होती है और इनकी cleavage site से काफी दूर होती है।
- इन restriction enzymes को कार्य करने के लिए Magnesium, ATP (adenosine triphosphate) और AdoMet (S-adenosyl methionine) की आवश्यकता होती है।
- Type I restriction enzymes में 3 subunits होती हैं।

Type II

- 3000 से भी ज्यादा अलग-अलग प्रकार के type II restriction endonuclease enzymes हैं।
- Type II restriction enzymes की recognition sites 4 से 8 base pairs लम्बी होती है और Mg^{2+} की उपस्थिति में सक्रिय होती है।
- इनको कार्य करने के लिए ATP या AdoMet की आवश्यकता नहीं होती है।
- ये restriction endonuclease ज्यादातर उसी site पर cut करते हैं, जहां वे DNA को पहचानते हैं और ये एक subunit के बने होते हैं।

Type III

- Type III restriction endonucleases recognition site से approximately 25 base pairs पर cut करते हैं।
- Type I और Type II से अलग, Type III restriction endonucleases दो अलग-अलग nucleotide sequences को recognition sites की तरह पहचानकर कार्य करता है।
- ये दो nucleotide sequences Type I और Type II द्वारा पहचाने गये sequences से अलग होते हैं, क्योंकि दो sequences non-palindromic और इनकी base pairing में एक दूसरे के विपरीत होते हैं।
- Type III restriction endonucleases को DNA restriction और methylation को पूर्ण करने के लिए AdoMet और ATP की आवश्यकता होती है।

Applications

► Restriction endonucleases को प्राथमिक रूप से DNA replication और protein expression experiments में genes को plasmid vectors में insert कराने में उपयोग किया जाता है।

- सामान्य रूप से उपयोग में लाया जाने वाला restriction endonuclease type II restriction endonucleases है।
- Vector में एक gene को insert कराने के लिए same restriction enzyme को plasmid DNA और gene insert पर use करते हैं। Restriction endonuclease से काटे जाने के बाद, इन्हें DNA ligase के द्वारा जोड़ दिया जाता है।
- इनको single nucleotide polymorphism को पहचानने में भी उपयोग किया जाता है, जो gene alleles से अन्तर पता करने में काम आती है।
- ये एक DNA की genotyping बिना sequencing की cost के allow कर सकता है। Restriction enzymes को डालने के बाद फिर gel electrophoresis कराने पर, gel में उपस्थित bands की संख्या samples के genotype को बताता है।

- Restriction endonucleases, DNA fragmentation के लिए भी उपयोगी है। यह technique DNA को specific sites पर कटने के लिए allow करती है, जो एक निश्चित sequence का fragment में होना सुनिश्चित करती है और सभी टूकड़ें समान आकार के होते हैं। Fragmented होने के बाद, DNA कई प्रकार की analytic techniques के लिए ज्यादा उपयुक्त हो जाता है। यह genetic fingerprinting जैसे प्रक्रियाओं में भी उपयोग किया जाता है।

Restriction sites: सभी restriction sites **palindromes** हैं (इनमें same double-stranded DNA base sequence दोनों दिशाओं में होते हैं)। उदाहरण के लिए अधिकतर काम में लिया जाने वाला **Eco RI** enzyme GAATT (5' से 3') sequence को पहचानता है। जब 5'- से 3'- पढ़ जाता है तो, complementary strand का sequence भी GAATTC होता है।

Bacteria restriction endonucleases को defense mechanisms (रक्षा तंत्र) (जैसे: viral invasion के विरुद्ध) की तरह उपयोग करता है। प्रत्येक प्रकार के bacterium में अपने स्वयं के restriction enzymes और specific recognition sites होती हैं। Foreign DNA को प्रभावी रूप से recognition sites पर cut लगाकर नष्ट कर दिया जाता है।

Bacteria अपने genome को, अपने DNA की restriction sites में base-modification से (सामान्यतः methylation से) बचाता है। अतः प्रत्येक strain के पास एक restriction endonuclease और DNA methylase, समान target विशिष्टताओं के साथ होता है,

“Restriction” endonuclease नाम इन अत्यधिक विशिष्ट nucleases को इसलिए दिया था क्योंकि ये foreign DNAs से invasion को जैसे कि bacterial viruses, को रोकता है।

Frequency of cutting: Restriction site specificity के कारण, restriction endonucleases DNA को fragments में cut करता है, जिनकी औसत लम्बाई का पता restriction site में उपस्थित base pairs की संख्या से (and to a lesser extent by the ratio of bases in the DNA) चलता है।

DNA जिसमें चारों bases समान मात्रा में होती हैं, DNA sense strand पर किसी स्थान पर प्रत्येक base की probability (होने की संभावना) (प्रायिकता) $1/4$ होती है। 4 base pairs की एक restriction site के लिए उस sequence की random occurrence की प्रायिकता $(1/4)(1/4)(1/4)(1/4) = 1/4^4 = 1/256$ है। 6 base pairs के लिए प्रायिकता $= 1/4^6 = 1/4096$ है और 8 base pairs के लिए यह प्रायिकता $= 1/4^8 = 1/65536$ है। अतः एक 6 base pairs restriction site वाला restriction endonuclease 4,096 base pairs औसत लम्बाई के fragments बनाएगा। ये fragment एक पूरे gene को रखने के लिए काफी बड़े होते हैं (अतः इसमें gene में इस restriction endonuclease के लिए कोई cut site नहीं है)।

Effect of base composition: DNA जिसका base composition 50% GC 50% AT से अलग है (जो चारों bases के बराबर संख्या के समकक्ष हैं), एक site के लिए probability calculate करना आवश्यक है जो इसके प्रत्येक घटक की probabilities का गुणन होता है। अगर एक DNA 66.7% GC है (मतलब $2/3$ base pairs GC है) और base pairs का random orientation (दिशात्मकता) होती है तो प्रत्येक A एवं T दोनों की probabilities $1/6$ होगी और G & C की probabilities $1/3$ होगी। अतः GAATTC की probability $(1/3)(1/6)(1/6)(1/6)(1/6)1/3 = 1/11,664$ हा गह, जो $1/4096$ से अलग है जब चारों bases बराबर मात्रा में उपस्थित थे। अतः **Eco RI** द्वारा बनाये गये fragment की औसत लम्बाई DNA जिसमें higher GC content हैं उसमें ज्यादा होगी।

Naming of restriction endonucleases: Restriction endonucleases के नाम bacteria की species और strains जिनसे ये derive हुए हैं पर रखा जाता है। **1st letter genus** के लिए होता है और अगले **2 letters (2nd and 3rd) species** के लिए **4th strain** के लिए और **Roman numeral जो strain के enzyme को बताता है।** 1st तीन letters, जो genus और species के लिए होते हैं, उन्हें italics में लिखते हैं। अतः **Eco RI**, *E. coli* strain RY13 से प्राप्त पहला restriction endonuclease है।

नीचे दी हुई list में, केवल एक strand के palindromes का sequence दिया जा रहा है (दूसरा इसका reverse complement है और 5' से 3' पढ़े जाने पर identical होगा)।

Cut site को vertical line (|) जो कि bases के बीच होती है, से समझाया गया है

Sticky ends (नीचे स) बनते हैं, जब भी cut sequence के exact center पर नहीं लगता। Pu मतलब कोई purine (A या G), Py मतलब कोई pyrimidine (C या T). (A/T) मतलब A या T (an AT base pair in either orientation).

Note: **Eco RII** असामान्य है क्योंकि उसमें recognition sequence एक (विषम) odd number के bases का होता है।

- *Aac65* I G|GTACC
- *Alu* I AG|CT
- *Bam* HI G|GATCC
- *Bgl* II A|GATCT
- *Cla* I AT|CGAT
- *Eco* RI G|AATTC
- *Eco* RII |CC(A/T)GG
- *Hae* III GG|CC
- *Hin* dII GTPy|PuAC
- *Hin* dIII A|AGCTT
- *Hpa* II C|CGG
- *Kpn* I GGTAC|C
- *Mbo* I |GATC
- *Not* I GC|GGCCGC
- *Nst* I ATGCA|T
- *Pst* I CTGCA|G
- *Pvu* I CGAT|CG
- *Sac* I GAGCT|C
- *Sal* I G|TCGAC
- *Sma* I CCC|GGG
- *Xma* I C|CCGGG

Isoschizomers: कुछ मामलों में, 2 या ज्यादा अलग-अलग enzymes identical sites की पहचान करते हैं। अलग-अलग source (स्रोत) से प्राप्त enzymes जो same site को पहचानते हैं और चाहें वे समान या असमान तरीके से काटते हों, उन्हें isoschizomers कहते हैं। *Sma* I और *Xma* I जो कि list में दिये हैं, ये एक Isoschizomers का उदाहरण है, जो एक ही site को अलग-अलग तरीके से cut करते हैं।

Neoschizomer: एक enzyme जो same sequence को पहचानता है लेकिन अलग-अलग तरह से काटता है। Neoschizomers एक विशिष्ट प्रकार (subset) के Isoschizomers हैं। उदाहरण के लिए *Sma* I (CCC/GGG) और *Xma* I (C/CCGGG) एक दूसरे के neoschizomers हैं।

Isocaudomer: एक enzyme जो थोड़े अलग sequence को पहचानता है, लेकिन एक समान सिरे बनाता है।

2. DNA methyltransferases

DAM Methyltransferase: *E. coli* के dam⁺ strains में, 5 ...GATC...3 sequence के adenine residues के N6 atom पर methyl group जुड़ा रहता है। DNA adenine methylase एक single-subunit nucleotide-independent (type II) DNA methyltransferase है जो recognition sequence 5 GATC 3 में adenine residues पर S-adenosylmethionine से methyl group का transfer करता है।

- *E. coli* में, efficient DNA mismatch repair के लिए dam methylation आवश्यक होता है।
- oriC पर सटीक DNA replication के initiation के लिए oriC पर उपस्थित chromosomes के segregation और partition के लिए और gene expression के modulation में भी आवश्यक है।
- Methyl group के adenine के N6 atom पर transfer से एक bulky alkyl residue को B-form DNA के major groove में रख दिया जाता है, जिनकी recognition sites में sequence 5 GATC 3 होता है।
- इसके विपरीत कुछ restriction enzymes को GATC-residues पर DNA को cleave करने के लिए methylation की आवश्यकता होती है।

DCM Methyltransferase

dcm 5 ...CCAGG...3 या 5 ...CCTGG...3 sequence में उपस्थित cytosine के C5 position पर methyl groups जोड़ता है और इस प्रकार EcoRI की तरह restriction enzymes से cleavage को रोकता है।

- 3. Restriction Mapping:** Restriction endonucleases से DNA के बड़े pieces को छोटे टुकड़ों में तोड़ा जा सकता है। यदि 2 restriction endonucleases उपयोग किये जाते हैं, तो प्रत्येक fragment जो 1st enzyme से बना है उसे भी दूसरे enzyme द्वारा और भी छोटे टुकड़ों में काट दिया जाता है। Electrophoresis और fragments की mobilities के मुकाबले में (fragments with those of "markers" of known size) fragments के relative size का पता चलता है। जिनमें cut बनाये हैं, उस क्रम को उल्टा करने पर, overlapping fragments को align करके मूल DNA का एक पूर्ण restriction map बना सकते हैं।
- 4. Vectors:** Cloned genes का replication प्राप्त करने के लिए इनको self-replicating genomes में डालना आवश्यक है, जिन्हें vectors कहते हैं। हालांकि bacterial plasmids को vectors की तरह उपयोग किया जाता है, जैसा कि नीचे बताया गया है। इसको अलावा भी कई प्रकार के vectors होते हैं। एक vector में निम्न properties होने चाहिए:

1. Vector में एक origin of replication (*ori*⁺) होना चाहिए जो DNA को अपनेआप replication के लिए allow करे और DNA में स्वतंत्र रूप host DNA होता है।
2. Cloning के लिए उपयोगी, circular vector को बिना किसी vector DNA नुकसान के restriction endonucleases से open करते हैं। इसके लिए विशिष्ट restriction site का होना आवश्यक है जो पूरे circular vector DNA में केवल 1 होनी चाहिए। Restriction endonuclease जो कि site विशिष्ट है, से काटने पर circular DNA linearize हो जाता है। Cut सिरों को वापस जोड़ने पर पूरा vector फिर से बन जाता है। यह cut सिरों के बीच foreign DNA sequence को अन्दर जाने के लिए allow करके एक नया, बड़ा circular vector बनाता है, जिसमें एक inserted sequence होता है जैसे: एक cloned coding sequence यह ध्यान दिया जाना आवश्यक है कि कुछ प्रकार के vectors का genome circular (गोलाकार) होता है।
3. ज्यादातर vectors कुछ selectable marker (जैसे कि antibiotic resistance) को code करते हैं। अतः host organism को विशेष परिस्थितियों में जीवित रहने के लिए vector की उपस्थिति आवश्यक है।
4. इसमें एक second marker भी होना चाहिए जो cloned DNA वाले vectors और cloned DNA रहित vectors के बीच अन्तर कर सके।

Positive selection- Positive selection में, host strain में कार्यात्मक जीन की कमी मीडिया पर नहीं बढ़ती है, लेकिन recombinant clone के साथ परिवर्तित होस्ट मीडिया में बढ़ने के लिए आवश्यक जीन उत्पाद की आपूर्ति करने में सक्षम हो सकता है।

- **Negative selection -** Negative selection में, एक रासायनिक यौगिक मीडिया में जोड़ा जाता है जो जीन उत्पाद की उपस्थिति में cytotoxic agent में परिवर्तित हो जाएगा, और परिणामस्वरूप, यह wild प्रकार की कोशिकाओं के विकास की अनुमति नहीं देता है। लेकिन पुनः संयोजक क्लोन के साथ तब्दील host strain एक गैर-कार्यात्मक जीन उत्पाद है और मीडिया में यौगिक की उपस्थिति में बढ़ता है।

- 5. Plasmids:** Bacterial plasmids छोटे circular DNAs होते हैं, जिनमें उनके स्वयं के origins of replication होते हैं और bacterial cells में autonomous (स्वायत्त) replication में समर्थ होता है। Plasmids जिनमें उचित genes होते हैं, वे antibiotics प्रतिरोधी bacteria बनाने में सक्षम होते हैं, जो उन bacterias का चयन करने में संभव होते हैं जिन्होंने plasmid लिया है। आकार छोटा होने के कारण, plasmid में एक restriction endonuclease के लिए केवल 1 cut site होती है, जो foreign DNA के समाकलन (Integration) के लिए circle को open करता है और plasmid के हिस्सों को खत्म होने से बचाता है। इनको कई तरीकों से modify भी किया जा सकता है। कई multiple cloning sites(MCS) के addition से बदलाव कर सकते हैं।

एक अन्य trick जिसका उपयोग प्राकृतिक रूप से उपस्थित cut sites को हटाने या बनाने में करते हैं। Codon के 3rd base को बदल कर restriction endonuclease की cut site को हटाता है। Amino acid के सम्बन्ध में ऐसे mutations silent हो हैं, लेकिन restriction endonucleases से DNA के cleavage के सम्बन्ध में नहीं।

Subcloning

Cloning experiment का सबसे सरल तरीका जो DNA cloning की कई basic पद्धतियों को प्रदर्शित करता है, cloned DNA का एक vector से दूसरे vector में transfer है, उपयोग जिससे subcloning कहा जाता है। यह एक large cloned fragment के short region को अधिक विस्तार में जाँच करने के लिए उपयोग किया जा सकता है या एक gene को एक ऐसे vector में transfer करने के लिए जो इसे particular species में express करने के लिए design किया गया है। उदाहरण के लिए E.coli में plasmid vectors के मामले में, जो सबसे सामान्य स्थिति है, प्रक्रिया को निम्न चरणों में बाँटा जा सकता है।

- इच्छित cloned sequence रखने वाले plasmid DNA का isolation.
- Restriction endonucleases के साथ plasmid का discrete fragment में digestion (cutting).
- Agarose gel electrophoresis द्वारा fragments का separation.
- Desired target fragment का purification.
- एक नया recombinant molecule बनाने के लिए एक नए plasmid में fragment का ligation (joining).
- Ligated plasmid का एक E. coli strain में transfer (transformation).
- Transformed bacteria का selection.
- Recombinant plasmids का analysis.

Plasmids: एक plasmid, bacterial chromosome से अलग एक दूसरा DNA molecule है जो स्वतंत्र रूप से replication और transmission के लिए सक्षम है। Plasmids circular है और या तो bacterial chromosome से अलग या bacterial chromosome के साथ जुड़े हुए होते हैं। विशिष्ट परिस्थितियों को छोड़कर वे host cell के लिए आवश्यक नहीं होते हैं। कई तरह के bacterial plasmids हैं, लेकिन इनमें 3 सबसे अधिक अध्ययन किए गए प्रकार हैं :-

- (1) F plasmids (conjugation के लिए उत्तरदायी हैं)
- (2) R plasmids (antibiotics के प्रति resistance के लिए genes रखता है) और
- (3) Col plasmids (colicins के लिए code करते हैं, वे proteins जो sensitive E.coli cells को kill (खत्म) करते हैं, वे particular colicin के प्रति immunity प्रदान करने वाले genes भी रखते हैं). Plasmids या तो conjugative (transmissible) हो सकते हैं (conjugation के द्वारा DNA transfer की मध्यस्तता करते हैं और इस प्रकार एक population के bacterial cells के बीच तेजी से वृद्ध करता है) e.g., F plasmids, कई R plasmids और कुछ Col plasmids, या nonconjugative (conjugation द्वारा DNA transfer mediate नहीं होता है) e.g., कई R plasmids और अधिकतर Col plasmids.

Stringent and Relaxed Replication: Mainly इसके replication control system के कारण bacterial cell में प्रत्येक plasmid एक विशिष्ट copy number कायम होता है।

इस संबंध में plasmids 2 प्रकार के होते हैं:- (1) single copy और (2) multicopy plasmids. Single copy plasmids का replication control उनके bacterial host cells की तरह है, इसलिए वे bacterial chromosome के साथ replicate और segregate होते हैं, इसे **stringent replication** कहा जाता है। इसके contrast में, multicopy plasmids का replication control bacterial host genome से अलग होता है, इसलिए host genome के प्रत्येक replication के लिए वे एक से अधिक बार replication करते हैं। इसे **relaxed replication** कहा जाता है।

Stringent Plasmid	Relaxed Plasmid
Low #	High #
Depend on protein synthesis of system (Host)	Independent (self)
Stop replicating upon using protein synthesis inhibitors.	Get amplify upon using protein synthesis inhibitors.
Replicates with Genomic DNA	Independent of genomic DNA (Only DNA Poly. Required)

TABLE: Antibiotic resistance genes found in R plasmids, their proteins and mechanism of antibiotic resistance

Antibiotic gene	Protein produced by the gene conferring resistance	Mechanism of resistance
Ampicillin (amp)	Penicillase or β -lactamase	Hydrolysis of C-N bond in β -lactam ring
Kanamycin (kan)	Kanamycin acetyltransferase*	N-acetylation of the antibiotic
Neomycin (neo)	Aminoglycoside phospho transferase	O-phosphorylation of the antibiotic
Streptomycin (str)	Streptomycin phospho transferase	Phosphorylation of OH on the antibiotic
Streptomycin	adenylate synthetase	Adenylation of the OH on the antibiotic

Some of major reporter genes/ protein used in RDT

Protein	Activity & Measurement
CAT (chloramphenicol acetyltransferase)	Transfers radioactive acetyl groups to chloramphenicol; detection by thin layer chromatography and autoradiography
GAL (β -galactosidase)	Hydrolyzes colorless o-nitrophenol- β -galactoside to o-nitrophenol and galactose (colored products.)
GUS (β -glucuronidase)	Hydrolyzes colorless X-gluc (5-bromo, 4-chloro, 3-indoyl β -glucuronide) to yield colored products.
LUC (luciferase) enzyme from firefly (<i>Photinus pyralis</i>),	Oxidizes Luciferin to oxyluciferin emitting photons.
GFP (green fluorescent protein) from the jelly-fish <i>Aequoria Victoria</i>	Fluoresces on irradiation with UV.

There are now several variants of GFP with altered fluorescence emission to give different colours – YFP (yellow), BFP (blue), CFP (cyan).

VECTORS

एक vector एक DNA molecule जो appropriate host cell में replicate होने की ability रखता है और जिसमें cloning के लिए clone किया जाने वाला DNA fragment (इसे DNA insert कहा जाता है) integrated होता है। इस प्रकार, एक vector को एक origin of DNA replication (जिसे ori से denote किया जाता है) रखना चाहिए जो host cell में function करता है। किसी extra-chromosomal small genome, e.g., plasmid, phage और virus को vector की तरह use किया जा सकता है।

Properties of A Good Vector

एक अच्छे vector को निम्न properties रखनी चाहिए: -

1. इसे autonomously replicate होने में सक्षम होना चाहिए। जब cloning का उद्देश्य DNA insert की कई copies प्राप्त करना है, vector replicon को relaxed control में होना चाहिए, जिससे यह single host cell में अपनी multiple copies generate कर सके।
2. इसे isolate और purify करना आसान हो।
3. इसे host cells में easily introduce किया जा सके, i.e., vector के साथ host का transformation easy होना चाहिए।
4. Vector में suitable marker genes होने चाहिए जो transformed host cells के easy selection को allow करता है।
5. जब उद्देश्य gene transfer है, इसमें स्वयं का DNA insert जो यह carry करता है, को host cell के genome में integrate करने की क्षमता होनी चाहिए।
6. DNA insert रखने वाले vector molecules के साथ transformed cells (recombinant या chimaeric vector) सिर्फ vector molecules द्वारा transformed cells में selectable या identifiable होने चाहिए।
7. एक vector को अधिक से अधिक restriction enzymes के लिए विशिष्ट target sites रखनी चाहिए, जहां DNA आवश्यक कार्यों को बाधित किये बिना जोड़ा जा सकता है।
8. जब DNA insert के expression की आवश्यकता होती है, vector को at least suitable control element रखने चाहिए, promoter, operator और ribosome binding sites; कई दूसरे factors भी important हो सकते हैं।

Cloning and Expression Vectors

- एक suitable host में DNA inserts के propagation के लिए use किये गये सभी vectors को cloning vectors कहा जाता है। लेकिन जब एक vector DNA insert के expression के लिए या इसके द्वारा specified protein के production के लिए design किया जाता है, इसे expression vector कहा जाता है।
- As a rule, इस प्रकार के expression vector at least सभी regulatory sequence i.e., promoters, operators, ribosomal binding sites etc. रखते हैं, जो चुने गए host में optimum function रखती हैं।
- सभी cloning vectors relaxed replication control रखते हैं, इसलिए वे per host cell अनेक प्रतियाँ उत्पन्न करते हैं।
- जब एक eukaryotic gene को एक prokaryote में express किया जाता है, eukaryotic coding sequence को prokaryotic promoter या ribosome binding site के बाद place किया जाना चाहिए क्योंकि eukaryotic regulatory sequences prokaryotes में recognize नहीं होती हैं।
- साथ ही, eukaryotic genes, as a rule, उनके coding regions में उपस्थित introns (noncoding regions) रखते हैं। Eukaryotic genes के उचित expression के लिए इन introns को हटाया जाना चाहिए क्योंकि prokaryotes में RNA transcript