20181130 12 KUSHAGRA AGARWAL

DATE
PAGE

Assignment - 2

Q1) BamHI → G|GATCC
Bal II → A|GATCT

i) NO, the two enzymes will not result in the same number of fragments in a random DNA seq. ~~Only the probability~~ The reason is the difference in the 6-mers. For BamHI to recognize a site "GGATCC" must be present while for Bal II "AGATCT" must be present. This difference ~~results~~ of recognition sites (irrespective of the fact that the sticky ends are same) results in the diff number of fragments.

ii) As the sticky ends formed are same, ~~such~~ the 2 can be used to cut plasmid vectors & DNA of interest interchangeably without the need of an adapter/linker.

For example, let's say our DNA of interest has a ~~restrict~~ recognition site for BamHI within itself (making it unsuitable to cut with BamHI) while our plasmid only has a recognition site for BamHI. In such a case, the general workflow is to cut our DNA of interest using same other RE & then ligate it with an adaptor/linker. But considering some sticky ends for Bal II & BamHI, we can instead cut our DNA of interest using Bal II & directly ligate with the plasmid to form the recombinant DNA.

Q2) Advantage/Disadv. of cloning vs PCR.

Cloning	PCR
i) Allows DNA manipulation	Only amplification is possible (sufficient only for sequencing)
ii) Once we have a good cloned products, the error rate for plasmid replic-ation is much lower than PCR error	Polymerases have error rates for incorporating wrong bases often near the primer (at low temperatures)
iii) At least a microgram of DNA is required	A nanogram of DNA is enough for PCR
iv) Time taken is 2-4 days	Time taken is maximum 4 hours
v) Labour intensive as automation is difficult	Automated kits available
vi) Absence of ~~suf~~ precise RE can limit the process	NO need of RE

Q3) Roles of

i) Primer → These are (15-20 bp) short fragments which can attach to our DNA region of interest They are single stranded and complementary to ~~Our~~ DNA strand. Taq Polymerases need a starting point to begin extension by adding dNTPs. Primers act as these start points.

ii) Taq polymerase → Taq polymerase is a thermostable DNA polymerase I named after thermophillic eubacterial microorganism. (Thermus Aquaticus) from which it was originally isolated by (Chien et.al)
It extends ~~DNA strands~~ complementary strand to our DNA strand of interest by starting from the primer and adding dNTPs. Its thermostability is useful as PCR is performed at high temperatures.

iii) ddNTPs → Dideoxy nucleotides are useful in PCR sequencing as they lack the ability to extend upon. They result in termination of DNA extension by as they lack hydroxyl groups at both 2' and 3' positions. They are marked with fluorescent markers. These are useful as on extension of multiple copies of DNA of interest on termination will take place at different lengths (due to addition of these ddNTPs instead of dNTPs by Taq polymerase). Then we can use gel electrophoresis to separate the diff lengths and scan using a laser to determine the Nucleotide.

Q4) Unknown DNA sequence is cloned in the vector. How would you design primers to sequence the insert?

sol7 first we use approprate RE to extract our DNA of interest and then

sol7 for the unknow DNA sequence we can choose a primer in 3 ways.

i) If the DNA of interest is from a species then looking for close relatives of the species may help us decide the primer as inter-species diversity for closely related species is very less. Ex: To sequence E N gene of SARS-CoV-2, we could use the primer used for N gene of SARS-CoV or MERS-CoV.

ii) we can attach adaptors/linkers to our isolated DNA of interest and then use these adaptor regions as primers for sequencing.

iii) if RNA transcribed mRNA is present for the DNA (gene encoding) then the poly-A tail can also be used as the primer binding region.

Q5) Adv of RNA seq by NGS approach compared to cDNA sequencing.

i) NGS is significantly cheaper
ii) Time taken is significantly lesser
iii) very less amount of input RNA is required
iv) Uses massive parallel sequencing which allows us to get greater coverage therefore reducing error significantly even though shorter reads are formed.
v) Hundreds to thousands of genes/gene regions can be sequenced simultaneously allowing us to sequence entire genomes.