SWAYAM NPTEL COURSE ON

COMPREHENSIVE MOLECULAR DIAGNOSTICS AND ADVANCED GENE EXPRESSION ANALYSIS

By

Prof. Aritri Bir

Dr. B. C. Roy Multi-Speciality Medical Research Centre
Indian Institute of Technology Kharagpur

Module 4: Tools of Molecular diagnostics and Gene expression Analysis (III)

Lecture 16: DNA Microarray

CONCEPTS COVERED

DNA microarray

- Concept & Principle
- Mechanism
- Application

DNA microarray

- DNA chips
- Gene chips
- DNA arrays
- Gene arrays
- Biochips

- DNA microarray is the technology in which specific DNA sequences are orderly arranged on a surface to detect and quantify genes.

Principle: Nucleic acid hybridization based on complementarity (Chargaff's rule)

- Evolved from Southern blotting
- The use of miniaturized microarrays for gene expression profiling was first reported in 1995
- The complete eukaryotic genome (Saccharomyces cerevisiae) on a microarray was published in 1997

Principle: Nucleic acid hybridization based on complementarity (Chargaff's rule)

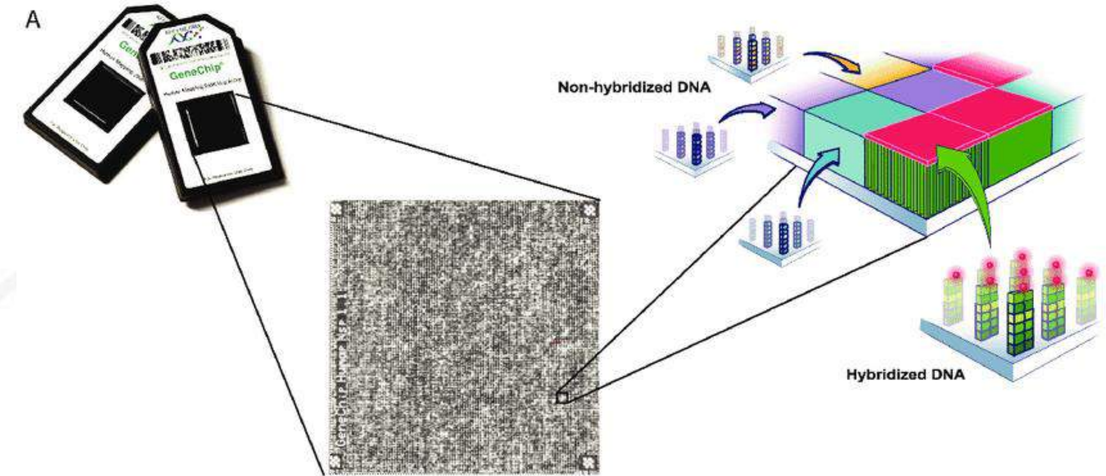
Steps:

- Manufacturing of DNA Chip containing synthetic single stranded DNA sequences.

Principle: Nucleic acid hybridization based on complementarity (Chargaff's rule)

Steps:

- Manufacturing of DNA Chip containing synthetic single stranded DNA sequences.



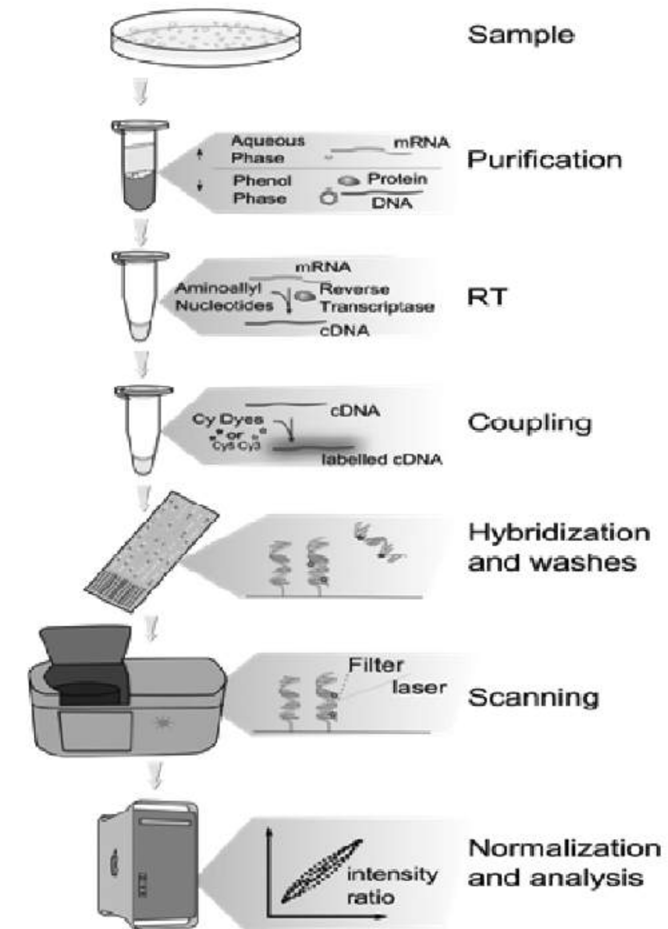
Kryh, Hanna. Molecular characterization of neuroblastoma tumors-a basis for personalized medicine. 2012.

Each spot (feature) contain a few million copies of identical DNA strands (probe) representing a single gene.

Principle: Nucleic acid hybridization based on complementarity (Chargaff's rule)

Steps:

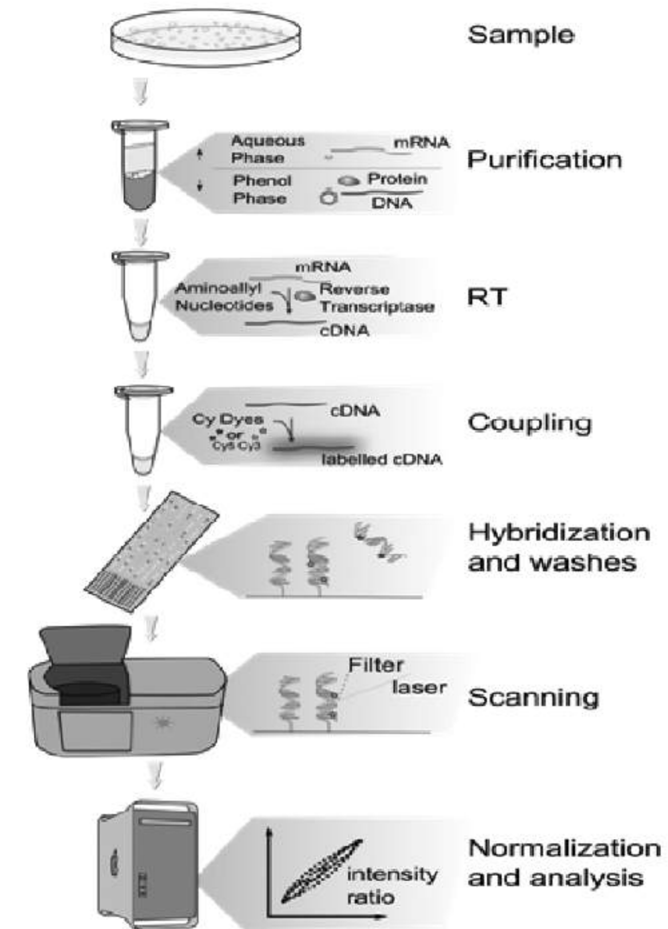
- Manufacturing of DNA Chip containing synthetic single stranded DNA sequences.
- Isolation and preparation of Sample DNA - labeled with a fluorescent dye.
- Hybridization into the chip.
- The bound DNA will be detected by its fluorescent dye and analyzed by a computer.



Principle: Nucleic acid hybridization based on complementarity (Chargaff's rule)

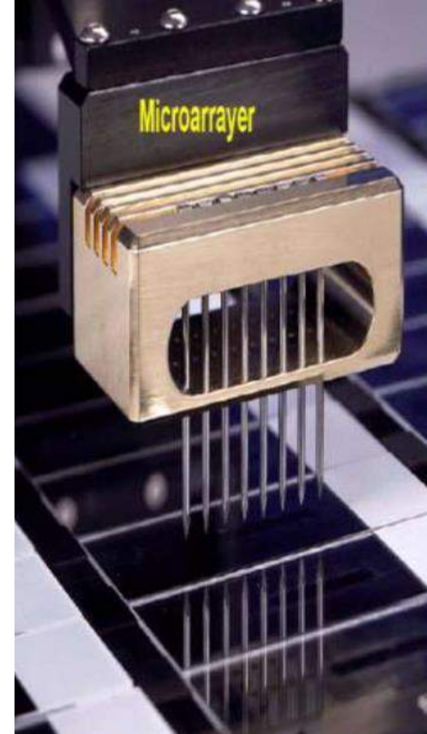
A mixture of labeled nucleic acids (target) binds to the probes by hybridization on the array surface.

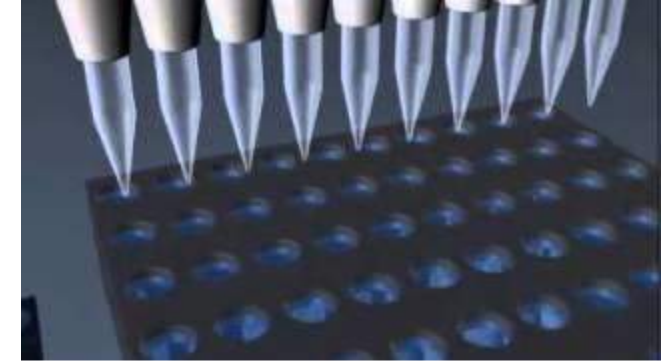
The presence/absence and relative concentrations of the nucleic acid species in solution can be measured.

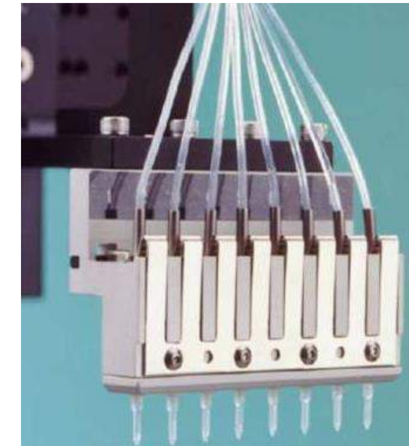


DNA Microarray types based on Fabrication process

Spotted	• Probes are oligonucleotides, cDNA • PCR products • The pre-synthesized probes are deposited or "spotted" on the array surface. • Pen tip method or Inkjet method • Robotic arm
In-situ	• Printing/synthesizing the probe directly onto the array surface • Photolithography / Inkjet • Short oligonucleotide sequences • High density



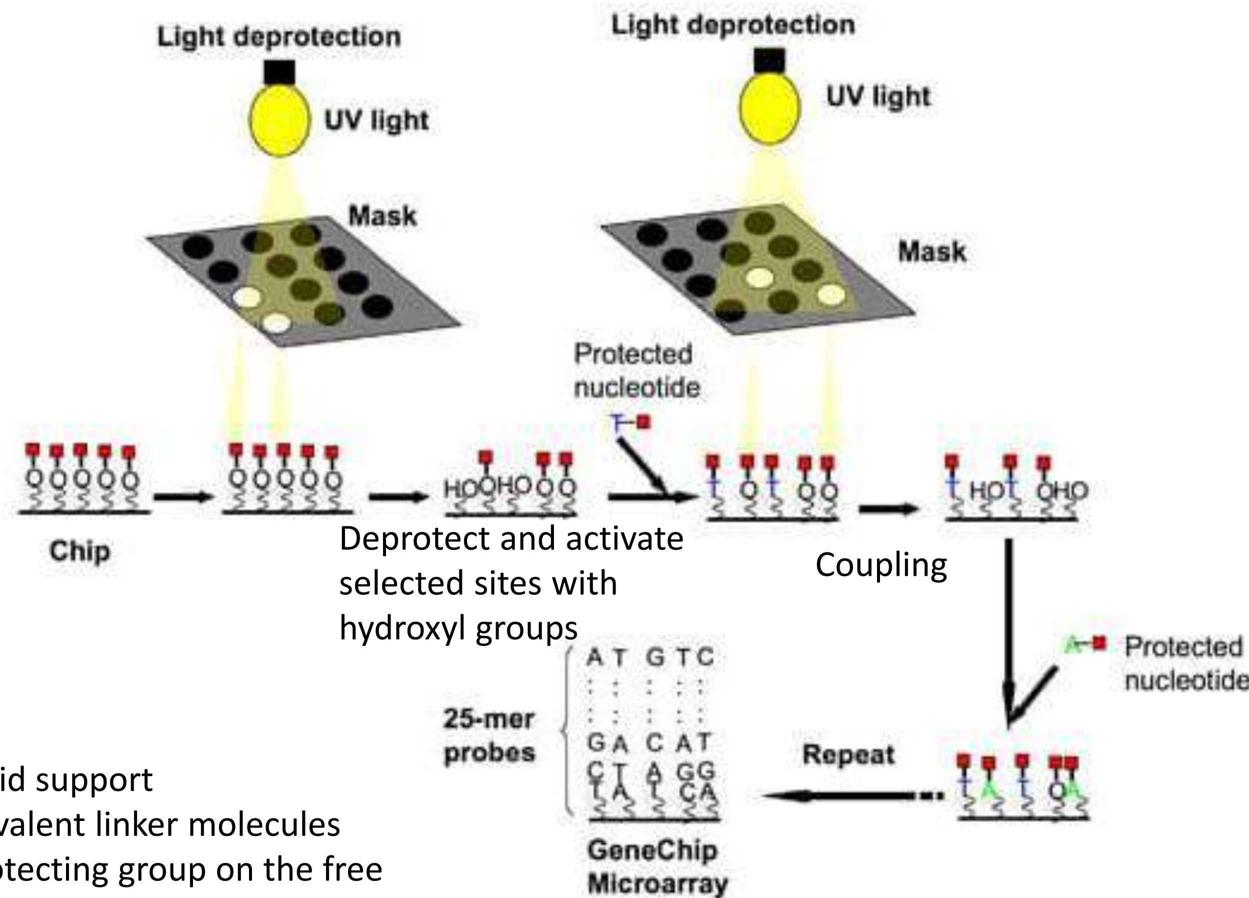




DNA Microarray types based on Fabrication process

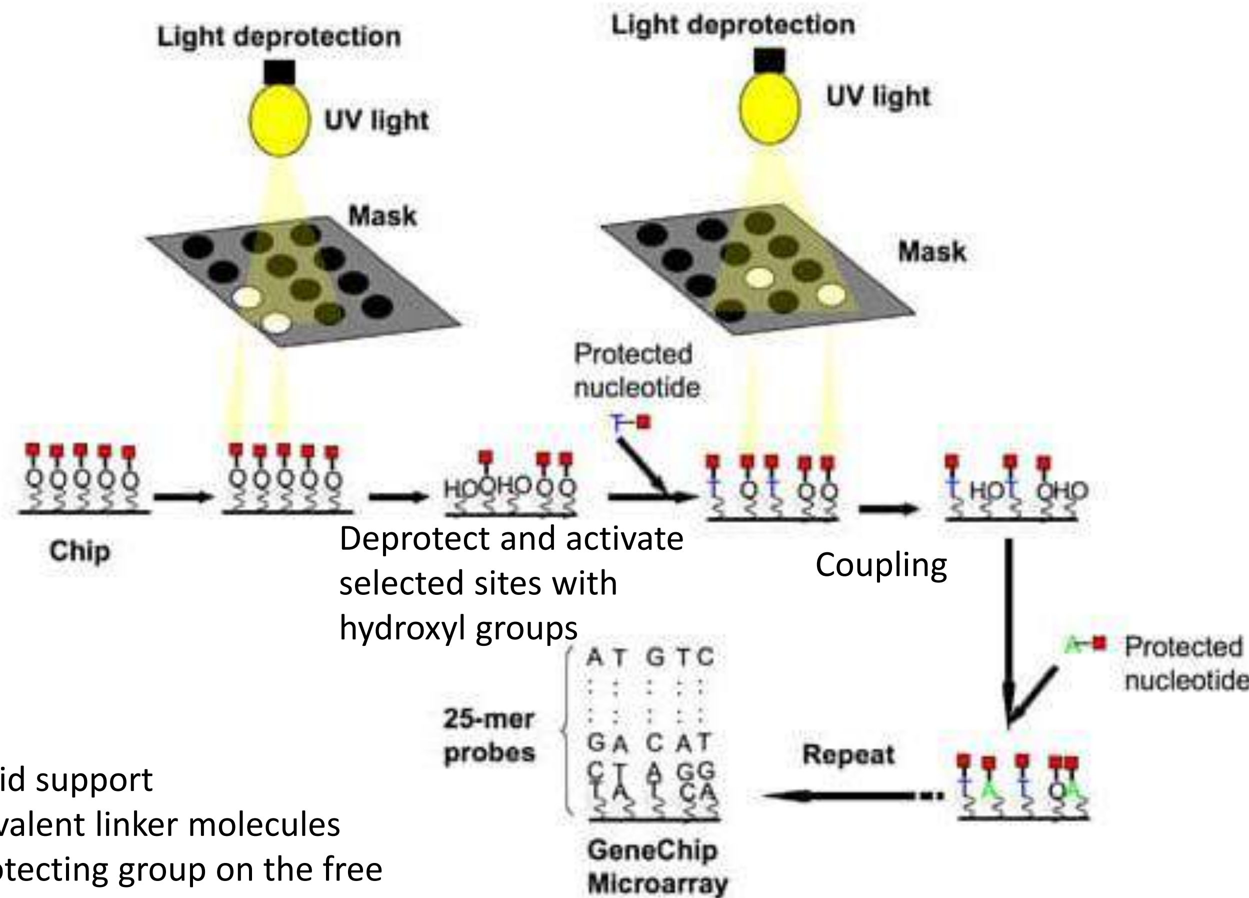
Spotted	• Probes are oligonucleotides, cDNA • PCR products • The pre-synthesized probes are deposited or "spotted" on the array surface. • Pen tip method or Inkjet method • Robotic arm	• Relatively inexpensive. • Flexible
In-situ	• Printing/synthesizing the probe directly onto the array surface • Photolithography / Inkjet • Short oligonucleotide sequences • High density	• Less time consuming • High specificity, less flexibility • costly • specialized machinery to synthesize the arrays

In-situ DNA microarray - Photolithography



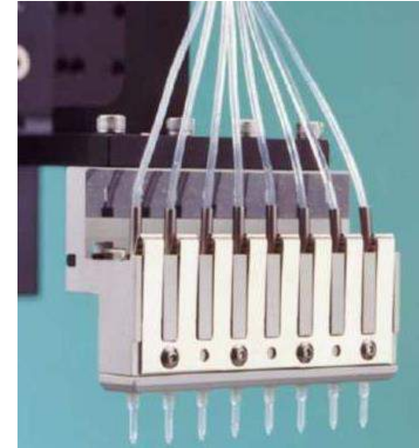
- In-situ synthesized arrays are high-density oligonucleotide probe DNA microarrays
- Synthesizes probes directly on the surface of the array
- one nucleotide at a time per spot, for many spots simultaneously
- Affymetrix GeneChip arrays most common.
- UV light is directed through a photolithographic mask
- The mask is designed in such a way that the exposure sites can be chosen as per requirement
- Addition of nucleotide occurs only on the chosen unprotected/unmasked sites.

In-situ DNA microarray - Photolithography



- The process is repeated
- New mask is applied activating different sets of sites and coupling different bases
- Each probe on the chip requires four masks per round of synthesis: one mask to allow addition of the required base and three other masks to prevent light from deprotecting the same spot while the other three nucleotides are being added.

In-situ DNA microarray –
Inkjet spotting process



- Five “ink” printing
  - 4 nucleotide precursors plus catalyst
  - combined with coupling and deprotection steps
  - Photolithographic masks not required.
- Printing picoliter volumes of nucleotides on the array surface in repeated rounds of base-by-base printing that extends the length of specific oligonucleotide probes.
- Synthesis of longer molecules (60-mer length oligos) for their probes.

JAN 2024

cDNA probe vs Oligonucleotide probe

Array preparation

cDNA collection

Insert amplification by PCR
Vector-specific primers
Gene-specific primers

Printing
Coupling
Denaturing

Ratio Cy5/Cy3

Hybridization
mixing

Target preparation

Cy3

Cy5

Cy3 or Cy5
labeled cDNA

First-strand cDNA
synthesis

Total RNA

Cells/tissue

mRNA reference
sequence

Perfect match
Mismatch

Probe set

In situ synthesis
by photolithography

Array 2

Array 1

Ratio array 1/array 2

Staining
hybridization

Biotiin-labeled
cRNA

In vitro transcription

Double-stranded
cDNA

cDNA synthesis

PolyA+ RNA

Cells/tissue

Wilson, Ann S., et al. "The microarray: potential applications for ophthalmic research." Mol Vis 8 (2002): 259-270.

Applications of DNA Microarray

Gene expression and profiling:

- Presence and absence of specific genes
- compare genes from different sources and to know how the external genes affect the external stimuli. gene expression pattern of different diseases such as cancer.
- Effect of certain treatments, diseases and developmental stage on gene expression.
    - Dynamic changes in gene expression under different conditions help to understand cellular processes, signaling pathways, and the regulatory mechanisms governing gene activity.

Comparative Genomic Hybridization: Genomic DNA is fluorescently labeled and used to determine the presence of gene loss or amplification

Applications of DNA Microarray

Disease Diagnosis and Classification: specific gene signatures associated with different diseases

Biomarker discovery

Cancer Research and Profiling:

- Molecular signatures of different types of cancer
- Identify genetic variations and expression patterns associated with specific cancer subtypes

Infectious Disease Monitoring:

- Monitor and study the genetic variations of infectious agents
- Identification and classification of pathogens
- Tracking disease outbreaks
- Understanding drug resistance mechanisms

Genetic Counseling and Prenatal Screening:

- Analysis of fetal DNA
- Genetic counseling

Functional Genomics:

- systematic analysis of gene functions.
- Identification of genes involved in specific biological processes
- Gene interactions
- Functional roles of genes within cellular pathways

Pharmacogenomics:

- Comparative analysis of genes to identify specific proteins produced by diseased cell
- Synthesis of drug
- Personalized medicine

Limitations of DNA Microarrays:

- Measurement is dependent on assay kinetics and linearity range
- Chances of cross-hybridization
- Difficulty in detection of individual members of gene families or splice variants
- Difficulty in detection on non-coding regions
- Reference strain bias
- Limited genome coverage

JAN 2024

Summary:

- Nucleotide hybridization
- In-situ and spotted technique
- Role in cancer diagnostics and personalized medicine.

REFERENCES

- Wilson, K. and Walker, J. (2010) Principles and Techniques of Biochemistry and Molecular Biology. Cambridge University Press, Cambridge.
- Nelson, D. L., & Cox, M. M. (2017). Lehninger principles of biochemistry (7th ed.). W.H. Freeman.
- Sambrook, J. and Russell, D.W. (2001) Molecular Cloning: A Laboratory Manual. 3rd Edition, Vol. 1, Cold Spring Harbor Laboratory Press, New York.
- Molecular Diagnostics: Techniques and Applications for the Clinical Laboratory" by Wayne W. Grody and Robert M. Nakamura
- Govindarajan R, Duraiyan J, Kaliyappan K, Palanisamy M. Microarray and its applications. J Pharm Bioallied Sci. 2012;4(Suppl 2):S310-S312. doi:10.4103/0975-7406.100283
- Bumgarner R. Overview of DNA microarrays: types, applications, and their future. Curr Protoc Mol Biol. 2013;Chapter 22:Unit-22.1.. doi:10.1002/0471142727.mb2201s101

THANK YOU

SWAYAM NPTEL COURSE ON

COMPREHENSIVE MOLECULAR DIAGNOSTICS AND ADVANCED GENE EXPRESSION ANALYSIS

By

Prof. Aritri Bir

Dr. B. C. Roy Multi-Speciality Medical Research Centre
Indian Institute of Technology Kharagpur

Module 4: Tools of Molecular diagnostics and Gene expression Analysis (III)

Lecture 17: FISH (Fluorescence in situ Hybridization)

CONCEPTS COVERED

- Fluorescence in situ Hybridization
  - Concept & Principle
  - Mechanism
  - Application

Fluorescence in-situ Hybridization

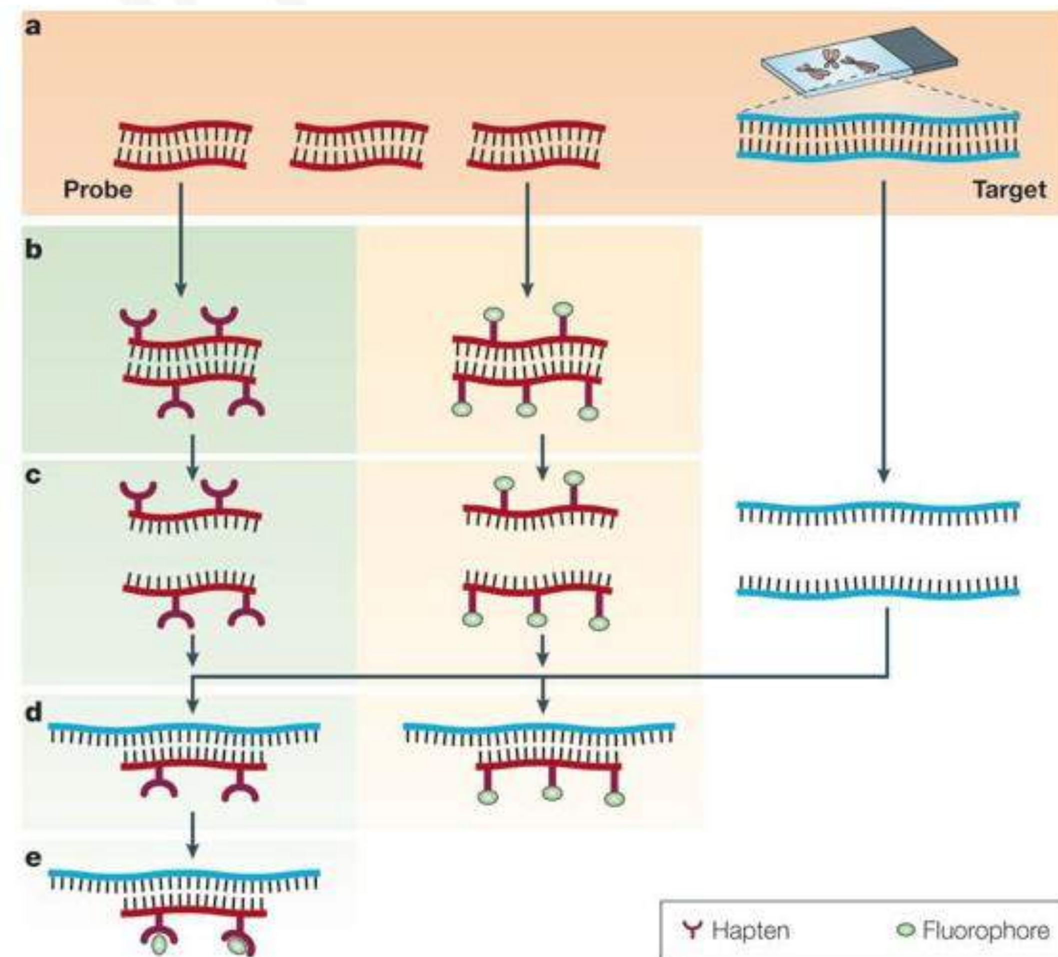
- In situ hybridizations are molecular cytogenetic method used for detection and localization of a specific DNA sequence inside a cell (in situ).
- First in situ hybridization experiments in 1969 (Gall & Pardue, 1969)
- FISH - In situ hybridization technique where the probe is tagged with fluorophore

Principle:

"Probe"— small piece of purified DNA tagged with a fluorescent dye finds and then binds to its matching (complementary) sequence within the set of chromosomes.

JAN 2024

Principle of fluorescence in situ hybridization



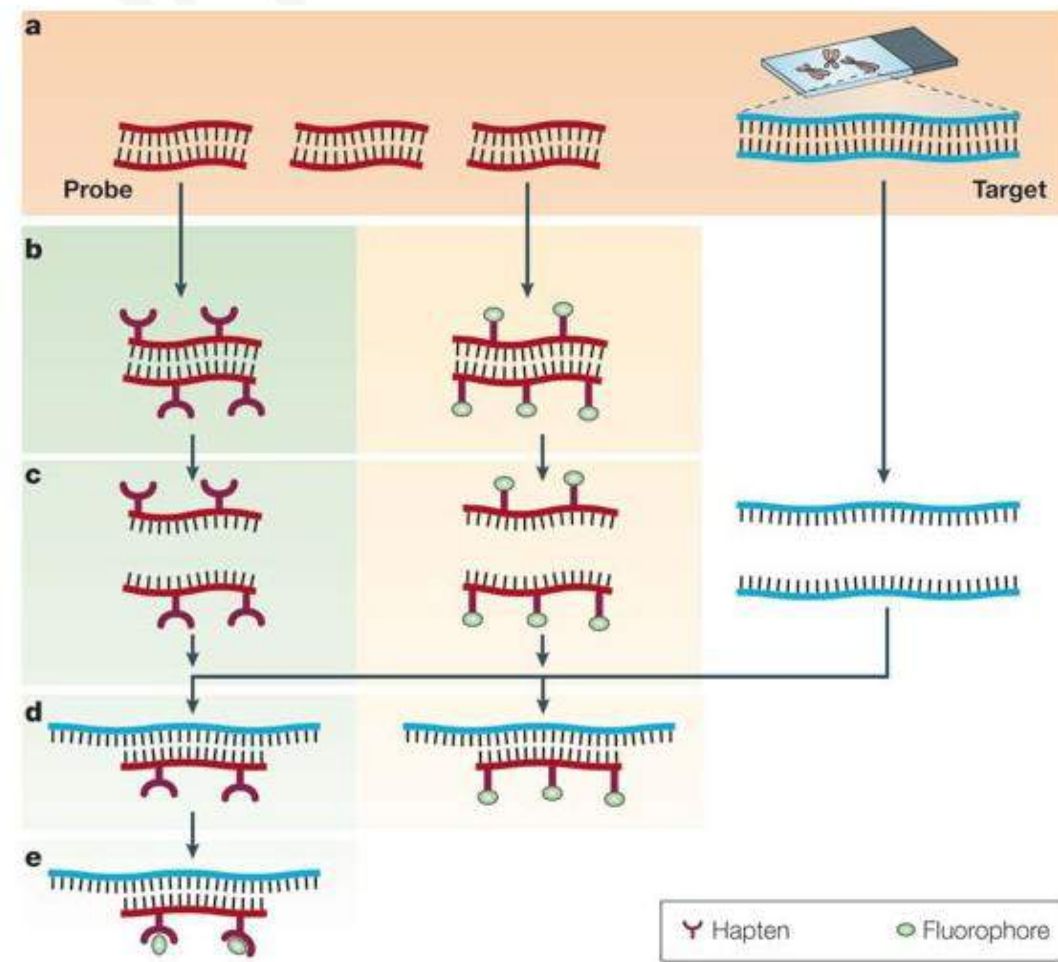
PROBE (Labelled) indirectly with a hapten/ directly with fluorophore

- Direct labelling: fluorescein (fluorescein isothiocyanate, FITC), rhodamine, Cy2, Cy3, Cy5, and AMCA.
- Indirect labelling: biotin, digoxigenin, and dinitrophenol.

- Nick translation
- Random primer labeling
- PCR using tagged nucleotide

Speicher MR, Carter NP. The new cytogenetics: blurring the boundaries with molecular biology. Nat Rev Genet. 2005;6(10):782-792. doi:10.1038/nrg1692

Principle of fluorescence in situ hybridization



PROBE (Labelled)
indirectly with a hapten/
directly with fluorophore

SAMPLE
DNA of either interphase cells
or of metaphase chromosomes

Denaturation

annealing of complementary DNA sequences

Fluorescence microscopy

Indirect labelling :

- ✓ Enzymatic : fluorochrome emitting colored signals at the hybridization site
- ✓ Immunological : based on binding of antibodies to specific antigens.

Speicher MR, Carter NP. The new cytogenetics: blurring the boundaries with molecular biology. Nat Rev Genet. 2005;6(10):782-792. doi:10.1038/nrg1692

The process is done in 3 main procedures:

- Tissue preparation (pre-hybridization)
- Hybridization
- Washing (post-hybridization) and Detection

- Frozen sections
- Paraffin embedded sections
- Cells in suspension

Hybridization and Washing:-

- The hybridization is done by mixing the single strand probes with the denaturated target DNA.
- Denaturation by heating
- Overnight incubation
- Serial washing to remove the unbound probe.

Type of Probes

Chromosome painting probes

- Chromosome-specific composite probe pools
- DNA probes derived from a single type of chromosome (PCR amplified & labelled)
- Homogenous “paint” to highlights the entire chromosome.
- Visualization of individual chromosomes in metaphase or interphase cells
- Identification of chromosomal aberrations.
- Simultaneous painting of every human chromosome in 24 colors.

Type of Probes

Chromosome painting probes	Repetitive sequence probes
• Chromosome-specific composite probe pools • DNA probes derived from a single type of chromosome (PCR amplified & labelled) • Homogenous “paint” to highlights the entire chromosome. • Visualization of individual chromosomes in metaphase or interphase cells • Identification of chromosomal aberrations. • Simultaneous painting of every human chromosome in 24 colors.	• Specific chromosomal regions or structures that contain short sequences • Pan-telomeric probes target the tandemly repeated (TTAGGG) sequences • Centromeric probes target the α- and β-satellite sequences flanking the centromeres ▪ Monosomy, trisomy, and other aneuploidies in leukemias and solid tumors

Type of Probes

Chromosome painting probes	Repetitive sequence probes	Locus-specific probes
• Chromosome-specific composite probe pools • DNA probes derived from a single type of chromosome (PCR amplified & labelled) • Homogenous “paint” to highlights the entire chromosome. • Visualization of individual chromosomes in metaphase or interphase cells • Identification of chromosomal aberrations. • Simultaneous painting of every human chromosome in 24 colors.	• Specific chromosomal regions or structures that contain short sequences • Pan-telomeric probes target the tandemly repeated (TTAGGG) sequences • Centromeric probes target the α- and β-satellite sequences flanking the centromeres ▪ Monosomy, trisomy, and other aneuploidies in leukemias and solid tumors	• Genomic cloned probes to particular region of chromosome • Detection of translocations, inversions, and deletions in both metaphase and interphase.

APPLICATION OF FISH

Gene Mapping and Detection of Chromosome Aberration:

- ✓ Target DNA intact
- ✓ Information about the correct positions
- ✓ Target nuclear RNA and genes within cells (Transcription and transport) by RNA-FISH
- ✓ Nuclear organization - high-resolution information about gene position and relationships during different cell cycle stages.

- Chromosomal deletions, disomy, imprinting, translocations
- Comparative Genomic Hybridization (CGH)

JAN 2024

APPLICATION OF FISH

Cancer Diagnosis and Subtyping:

- Applicability in wide range of tissue types
- Identify specific chromosomal abnormalities or gene amplifications associated with various types of cancer.
- Monitoring Minimal Residual Disease by detecting residual cancer cells with specific genetic abnormalities.

Prenatal Genetic Testing:

- Visualization of specific chromosomal abnormalities in fetal cells.
- Down syndrome, trisomy 18, and trisomy 13

Microbial Detection and Identification:

- **R**apid and specific detection of microbial pathogens in clinical samples by targeting the nucleic acid sequences
- Human Papillomavirus (HPV) Detection: early diagnosis of cervical cancer.

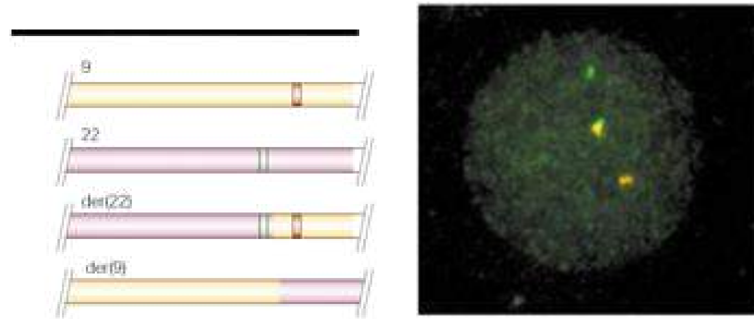
Pharmacogenomics Research: Study genetic variations affecting drug response.

Multiplex-FISH (M-FISH):

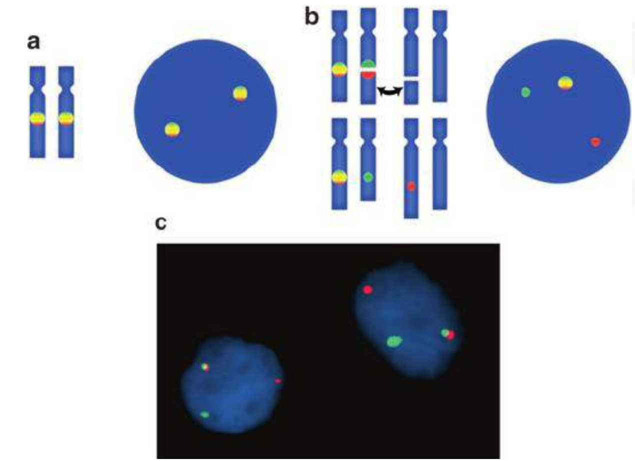
- Simultaneous identification of multiple chromosomal regions or genes using distinct colors for each target.
- Unique combination of fluorophores for each probe, allowing for comprehensive chromosomal analysis.
- Each probe is labeled with a specific combination of five spectrally separable fluorochromes
- Initially, M-FISH was designed for the simultaneous detection of all 24 human chromosomes
- Adapted to specific chromosomal subregions, such as centromeres and sub-centromeres

Multilocus FISH (ML-FISH) - simultaneous use of multiple probes in multicolor FISH.

D- FISH (Double colour FISH) - Dual colour probes can be used to develop secondary colour on overlapping.



O'Connor, C. (2008) Fluorescence in situ hybridization (FISH). Nature Education 1(1):171

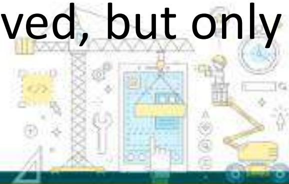


S.E. Pambuccian and R.H. Bardales, Lymph Node Cytopathology,Essentials in Cytopathology 10, DOI 10.1007/978-1-4419-6964-4_2,

D- FISH (Double colour FISH) - Dual colour probes can be used to develop secondary colour on overlapping.

Presence of secondary colour : detection of BCR/ABL translocations, where the secondary color indicates disease.

Absence of secondary color is pathological - **"Break-apart FISH"** - Locus-specific probes for one side of the breakpoint and the other intact chromosome. In normal cells secondary color is observed, but only the primary colors are observed when the translocation occurs.



Quantitative Fluorescent in situ Hybridization (Q-FISH):

- Quantitative information about the abundance or characteristics of specific sequences.
- Commonly used to measure the length of telomeres (cellular aging, cancer)
- Peptide nucleic acid (PNA) probes - synthetic DNA analog with a neutral peptide backbone that enhances probe stability and binding affinity.

RNA-FISH:

- Simultaneous detection, localization, and quantification of individual mRNA molecules either in the nucleus or cytoplasm at the cellular level in fixed samples.
- allelic-specific expression on a single-cell basis

Stellaris RNA FISH (Single-Molecule RNA FISH) – single molecule of mRNA detection and quantification.

Immuno-FISH is a combination of standard FISH and indirect or direct immunofluorescence.

Flow-FISH - the in situ hybridization is combined with flow cytometry for measurement of the telomeric signals from cells in suspension. The PNA-labeled telomere probes.

Fiber fish: High resolution mapping of chromosome fibers - stretched over slides.

Summary:

- Detection of gene/DNA sequence in situ
- Different types of probes
- Cancer diagnostics.
- Chromosomal anomaly detection

REFERENCES

- Wilson, K. and Walker, J. (2010) Principles and Techniques of Biochemistry and Molecular Biology. Cambridge University Press, Cambridge.
- Nelson, D. L., & Cox, M. M. (2017). Lehninger principles of biochemistry (7th ed.). W.H. Freeman.
- Sambrook, J. and Russell, D.W. (2001) Molecular Cloning: A Laboratory Manual. 3rd Edition, Vol. 1, Cold Spring Harbor Laboratory Press, New York.
- Molecular Diagnostics: Techniques and Applications for the Clinical Laboratory" by Wayne W. Grody and Robert M. Nakamura
- Speicher MR, Carter NP. The new cytogenetics: blurring the boundaries with molecular biology. Nat Rev Genet. 2005;6(10):782-792. doi:10.1038/nrg1692
- O'Connor, C. (2008) Fluorescence in situ hybridization (FISH). Nature Education 1(1):171
- S.E. Pambuccian and R.H. Bardales, Lymph Node Cytopathology,Essentials in Cytopathology 10, DOI 10.1007/978-1-4419-6964-4_2,
- Shakoori AR. Fluorescence In Situ Hybridization (FISH) and Its Applications. Chromosome Structure and Aberrations. 2017;343-367. Published 2017 Feb 10. doi:10.1007/978-81-322-3673-3_16

THANK YOU

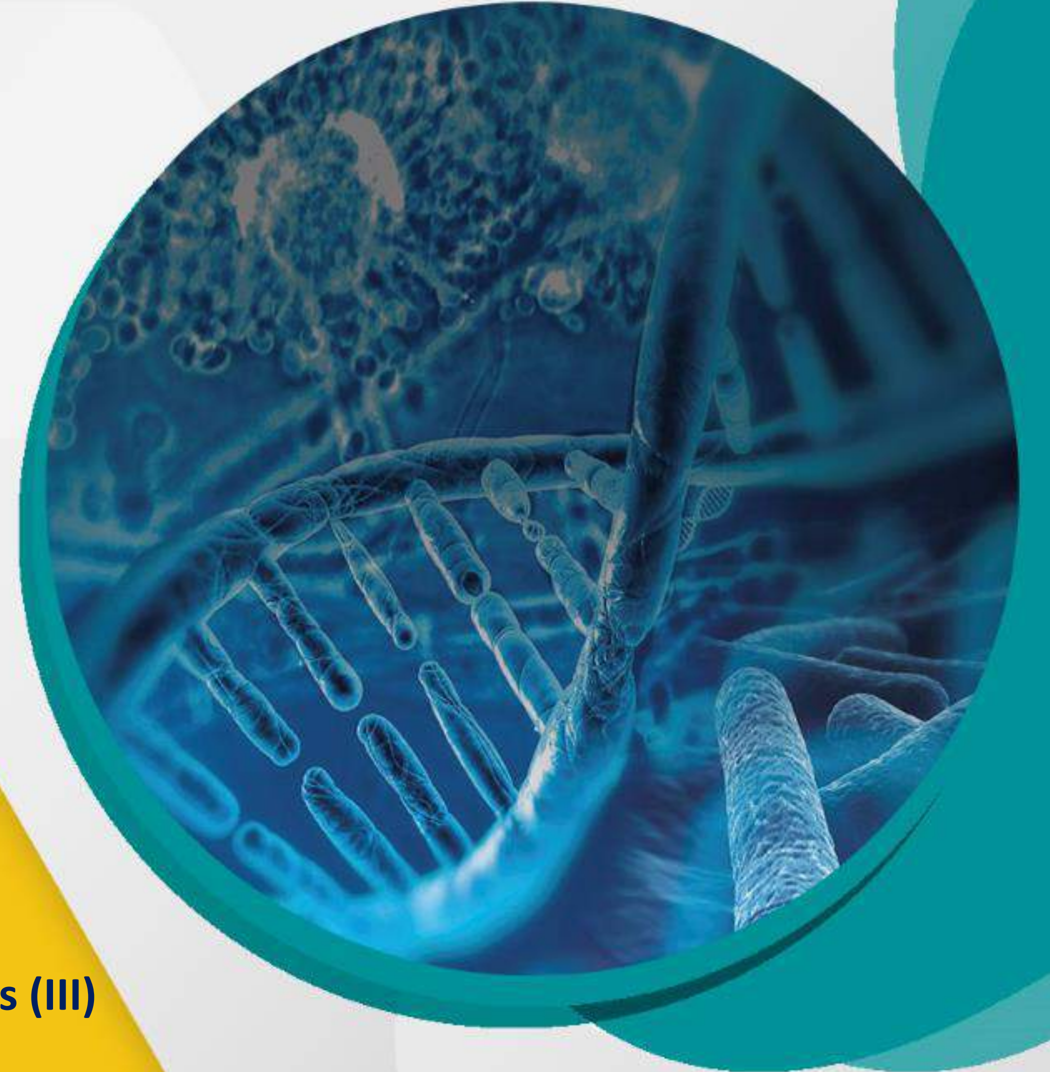
SWAYAM NPTEL COURSE ON

COMPREHENSIVE MOLECULAR DIAGNOSTICS AND ADVANCED GENE EXPRESSION ANALYSIS

By

Prof. Aritri Bir

Dr. B. C. Roy Multi-Speciality Medical Research Centre
Indian Institute of Technology Kharagpur



Module 04: Tools of Molecular diagnostics and Gene expression Analysis (III)

Lecture 18: Methods to study DNA-protein interaction I

CONCEPTS COVERED

- In vitro techniques to study DNA–protein interaction
- In vivo techniques to study DNA–protein interaction

DNA –Protein Interaction:

- Fundamental Understanding of Gene Regulation
- Cellular Processes and Function: DNA replication, transcription, translation, and repair.
- Aberrant DNA-protein interactions innumerous diseases, including cancer, neurodegenerative disorders, and genetic diseases.
- Many drugs target proteins involved in DNA-protein interactions.
- Biotechnology and Genetic Engineering: Techniques like gene editing (e.g., CRISPR-Cas9) and gene expression manipulation rely on the precise control of DNA-protein interactions.

JAN 2024

Methods

In vitro techniques to study DNA–protein interaction

- Footprinting assay
- Electrophoretic mobility shift assay (EMSA)
- Southwestern blotting
- Yeast one-hybrid assay (Y1H)
- Phage display for DNA-binding proteins

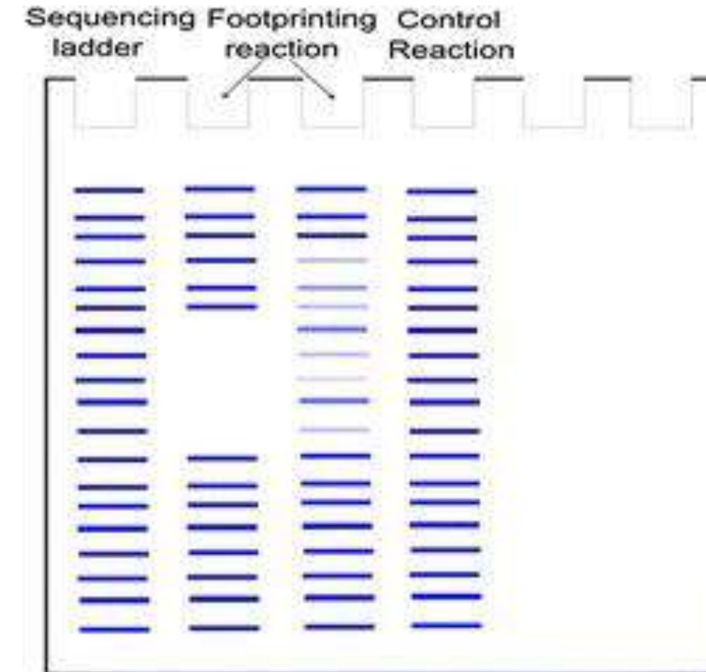
In vivo techniques to study DNA–protein interaction

- Chromatin immunoprecipitation (ChIP)
- X-ChIP
- Native-ChIP (N-ChIP)
- Carrier ChIP

NPTEL

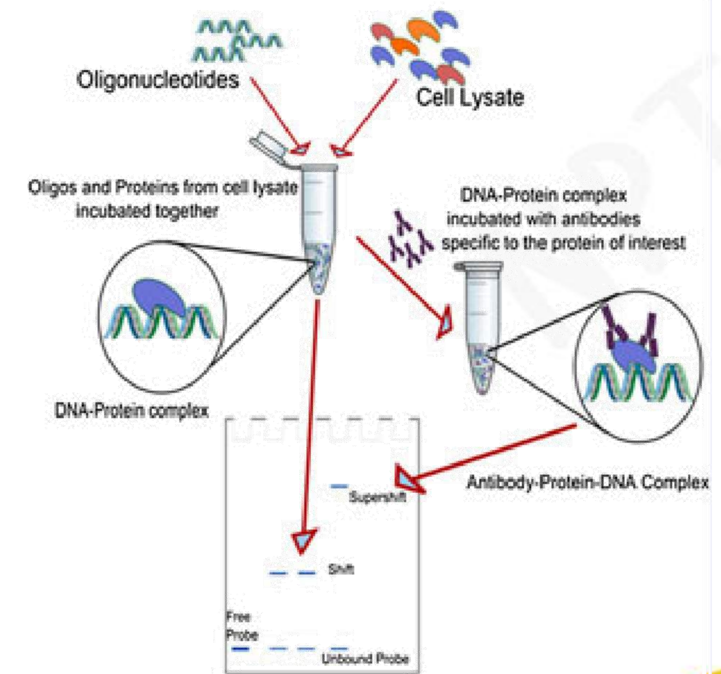
Foot printing assay

- Principle: Protein bound to DNA protects from degradation.
- Identification of specific sequence to which a DNA-binding protein or molecule binds.
- Chemical or enzymatic (DNAse) digestion of naked and protein bound-DNA.
- Both the reactions are then compared using gel electrophoresis.
- The segment of the DNA bound by the protein appears as an empty stretch 'footprint' in the protein-bound reaction when compared to the continuous fragments produced by naked DNA digestion.
- By varying the concentration of the DNA-binding protein, the binding affinity of the protein can be estimated according to the minimum concentration of protein at which a footprint is observed.



Electrophoretic mobility shift assay (EMSA)

- Electrophoretic separation of a protein–DNA or protein–RNA mixture on a polyacrylamide or agarose gel.
- Principle: DNA–protein complexes are heavier and move slowly when subjected to non-denaturing polyacrylamide or agarose gel electrophoresis as compared to unbound free probe.
- Gel shift or gel retardation assay.
- The DNA probes used may be radiolabeled or dyes specific to stain DNA and protein
- The control lane (DNA probe without protein present) will contain a single band corresponding to the unbound DNA.
- The lane with a protein bound DNA will produce band that represents the larger, less mobile complex of nucleic acid probe bound to protein which is 'shifted' up on the gel (since it has moved more slowly).

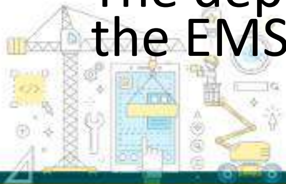


Electrophoretic mobility shift assay (EMSA)

- Supershift assay: an antibody specific to the protein of interest.

- Capillary electrophoretic mobility shift assay (CEMSA):
  - ✓ separation and quantitation of DNA–protein interactions in uncoated capillaries with no gel matrixes
  - ✓ High-sensitivity laser-induced-fluorescence detection of fluorescein-labeled DNA.
  - ✓ Capillary electrophoresis separates analytes on the basis of their mass-to-charge ratio and elutes complexes in the order of free protein, protein/DNA complex, and lastly DNA.

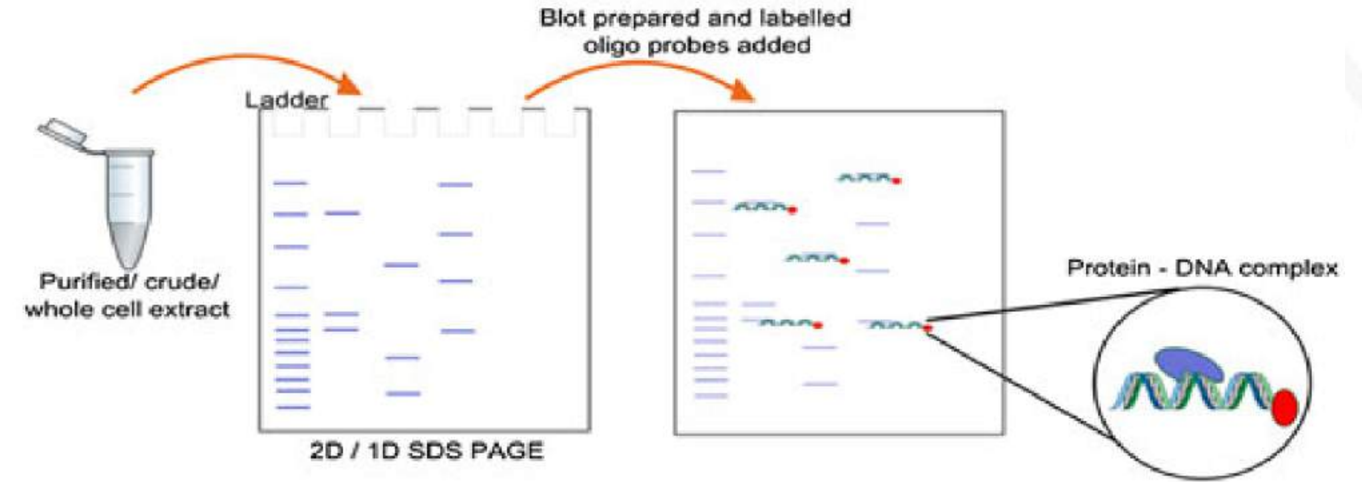
- IDEMSA: (EMSA combined with immunodepletion and supershift assays)
  - ✓ The nuclear or cytoplasmic extracts are depleted of the specific protein by incubation with the relevant antibody and protein A-sepharose.
  - ✓ The depleted extracts are then analyzed for the presence of protein by the EMSA and supershift assay.



Southwestern blotting

- Principle: Combines principles of Southern and Western blotting.
  - Used to elucidate the molecular weight of proteins in protein-DNA complexes.
  - Particularly useful when no antibodies are known for the bound protein.

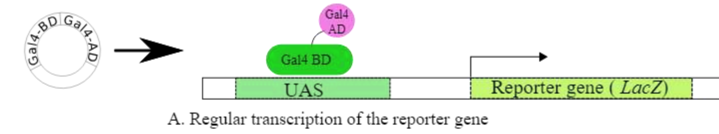
Southwestern blotting



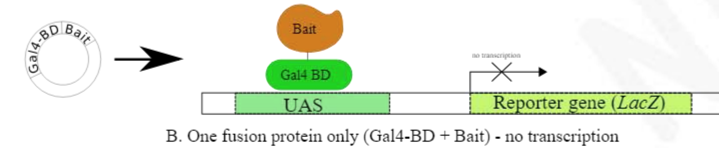
- Experimental Procedure:
  - Labeled oligonucleotides instead of antibodies as probes.
  - Crude or purified extracts are resolved on SDS-PAGE.
  - Proteins transferred to a membrane under conditions favoring renaturation.
  - Membrane-bound proteins incubated with oligonucleotides to which the protein of interest putatively binds.
  - Developed membrane photographed; only the band corresponding to the bound oligo appears in the final picture.
- Using differently labeled oligos on the same blot provides information on the binding affinity of various mutants.

Yeast one-hybrid assay (Y1H)

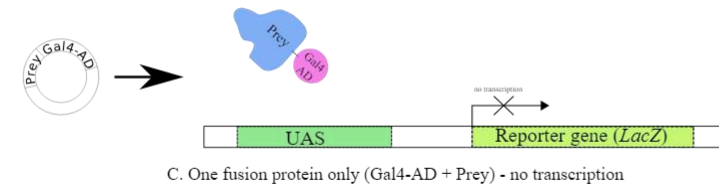
- Modification of the yeast-two hybrid (Y2H) assay
- Most eukaryotic transcription factors - activation domain (AD) and the DNA-binding domain (DB) (physically separable).
- AD and DB if separated : functionally inactive transcription factor (No RNA polymerase recruitment at promoter site)
- Protein X is translationally fused to AD (Prey)
- Protein Y is translationally fused to DB (Bait)
- Both are expressed in the same yeast cell.
- On X-Y interaction - the AD and DB in close physical proximity to reconstitute the functionally active transcription factor
    - expression of a downstream reporter gene.



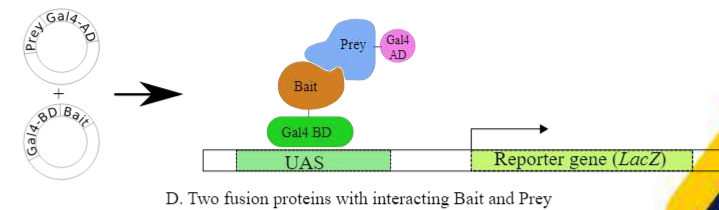
A. Regular transcription of the reporter gene



B. One fusion protein only (Gal4-BD + Bait) - no transcription



C. One fusion protein only (Gal4-AD + Prey) - no transcription

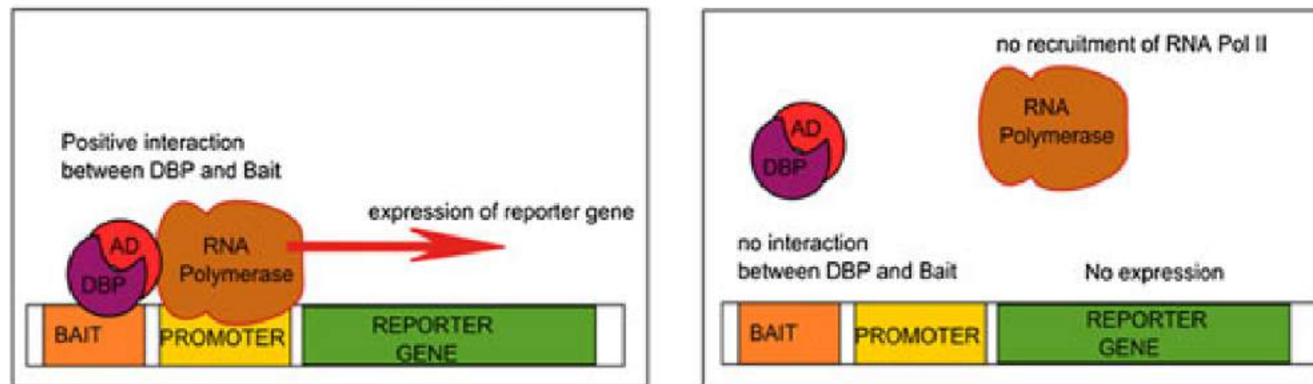


D. Two fusion proteins with interacting Bait and Prey

JAN 2024

Yeast one-hybrid assay (Y1H)

- The bait - DNA sequence of our interest
- Interaction of a protein X with the bait sequence is assayed.
- If X interacts with the bait DNA sequence, it results in bringing AD-X fusion close to the promoter, allowing AD to activate the RNA polymerase and result in the expression of the downstream reporter gene.
- Since this assay contains only one-hybrid encoded on a vector, it is called the 'Y1H'.
- Detection of the DNA–protein interaction occurs while proteins are in their native configurations – high sensitivity



Phage display for DNA-binding proteins

- Phage display refers to the method of expressing a peptide or protein domain on a bacteriophage capsid by genetically fusing its amino acid sequence to that of the coat proteins encoded by the phage.
- Phage-display library - affinity purification using an appropriate ligand
- Affinity purification - dsDNA oligos (with the binding sequence specific to a protein) bound to a solid matrix.
- The phage that displays a protein that binds to dsDNA oligo will remain bound to matrix while others are removed by washing.
- Those that remain can be eluted and will be used to produce more phage (by bacterial infection with helper phage)
- The identity of the selected clones can be obtained by sequencing the phage genome.

Summary:

- In vitro techniques to study DNA–protein interaction
- Footprinting assay - DNA-binding proteins protect specific DNA sequences from degradation; identification achieved through chemical or enzymatic digestion, and gel electrophoresis reveals footprints indicating protein binding.
- Electrophoretic mobility shift assay (EMSA) - Electrophoretic separation of protein–DNA or protein–RNA complexes on polyacrylamide or agarose gel using gel shift assay, where slower migration indicates protein binding, visualized with radiolabeled or dye-stained DNA probes.
- Southwestern blotting -The method employs labeled oligonucleotide probes in a hybrid approach, integrating SDS-PAGE and membrane transfer, to discern protein-DNA interactions and determine the molecular weight of protein-DNA complexes, especially beneficial in the absence of specific antibodies.
- Yeast one-hybrid assay (Y1H)- assessing the activation of reporter genes in yeast
- Phage display for DNA-binding proteins - Phage display involves expressing a peptide or protein on a bacteriophage capsid, facilitating affinity purification through dsDNA oligos, enabling selection of binding clones via washing, and subsequent identification through phage genome sequencing.

REFERENCES

- Biochemistry by Lubert Stryer 9th Editon
- Lehninger Principles of Biochemistry 8th Edition
- Harper's Illustrated Biochemistry 31st Edition
- Molecular Cloning: A Laboratory Manual by David W. Russell and Joseph Sambrook
- Wilson, K. and Walker, J. (2010) Principles and Techniques of Biochemistry and Molecular Biology. Cambridge University Press, Cambridge.
- Dey, B., Thukral, S., Krishnan, S. et al. DNA–protein interactions: methods for detection and analysis. Mol Cell Biochem 365, 279–299 (2012). https://doi.org/10.1007/s11010-012-1269-z
- Ferraz, R.A.C., Lopes, A.L.G., da Silva, J.A.F. et al. DNA–protein interaction studies: a historical and comparative analysis. Plant Methods 17, 82 (2021). https://doi.org/10.1186/s13007-021-00780-z
- Cozzolino F, Iacobucci I, Monaco V, Monti M. Protein-DNA/RNA Interactions: An Overview of Investigation Methods in the -Omics Era. J Proteome Res. 2021;20(6):3018-3030. doi:10.1021/acs.jproteome.1c00074

THANK YOU

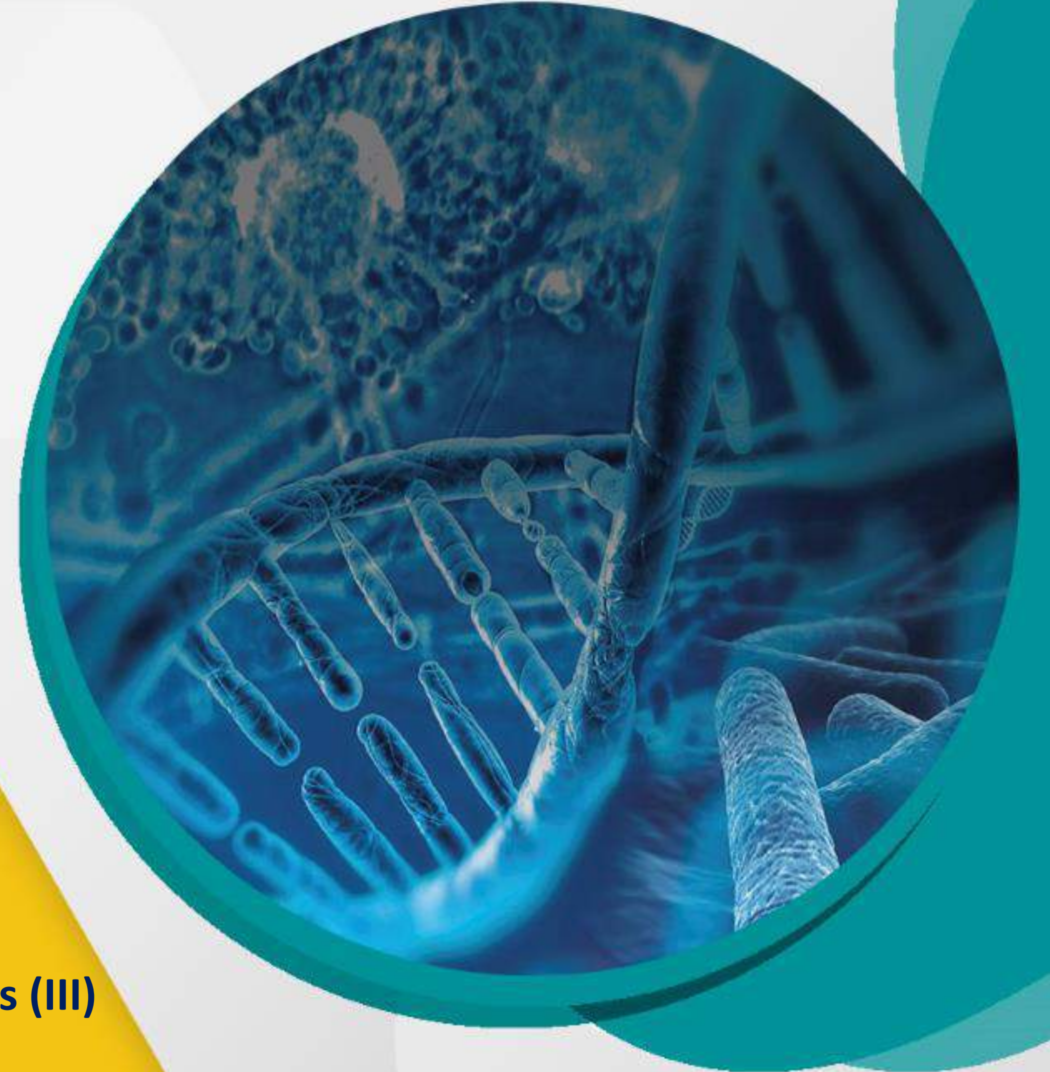
SWAYAM NPTEL COURSE ON

COMPREHENSIVE MOLECULAR DIAGNOSTICS AND ADVANCED GENE EXPRESSION ANALYSIS

By

Prof. Aritri Bir

Dr. B. C. Roy Multi-Speciality Medical Research Centre
Indian Institute of Technology Kharagpur



Module 04: Tools of Molecular diagnostics and Gene expression Analysis (III)

Lecture 19: Methods to study DNA-protein interaction II

CONCEPTS COVERED

- In vitro techniques to study DNA–protein interaction
- In vivo techniques to study DNA–protein interaction

Methods

In vitro techniques to study DNA–protein interaction

- Footprinting assay
- Electrophoretic mobility shift assay (EMSA)
- Southwestern blotting
- Yeast one-hybrid assay (Y1H)
- Phage display for DNA-binding proteins

In vivo techniques to study DNA–protein interaction

- Chromatin immunoprecipitation (ChIP)
- X-ChIP
- Native-ChIP (N-ChIP)
- Carrier ChIP
- ChIP-Chip
- ChIP Sequencing

JAN 2024

Chromatin immunoprecipitation

- Classical antibody-based ChIP format –
- shearing of protein associated chromatin into smaller fragments (into ~500 bp DNA fragments by sonication or nuclease digestion)
- immunoprecipitating the DNA–protein complex using protein-specific antibody.
- The isolated DNA–protein complexes are then dissociated and the specifically enriched DNA segment is analyzed using PCR amplification methods.
- Antibody-free format by the use of HaloTag Technology –
- transfection of Halo tag vectors containing halo tags fused to proteins of interest
- followed by their expression in mammalian cell lines.
- The cells are then cross-linked, lysed and sonicated and the DNA–protein complexes are captured onto a HaloLink Resin.
- This is followed by the standard decross-linking, DNA purification and PCR amplification of enriched DNA

JAN 2024

X-ChIP

- Freezing of all DNA-associated proteins by cross-linking (formaldehyde).
- Formaldehyde forms covalent cross-link between the specific proteins to the DNA on which they are situated (reacts with primary amines located on amino acids and the bases on DNA or RNA molecules)
    - ✓The protein–DNA complex is immunoprecipitated
    - ✓DNA–protein cross-links are reversed by heating.
    - ✓Removal of protein by proteinase K.
    - ✓The DNA portion of the complex is then purified and identified by PCR

- Minimization of nucleosome rearrangements
- Analysis of proteins that are weakly or indirectly associated to DNA.

Native-ChIP (N-ChIP)

- Natural DNA–protein interactions where the proteins are tightly associated to chromatin in their native state such as histones due to their high-affinity for DNA. Hence, these interactions do not require cross-linking with formaldehyde.

Fast ChIP

- Suitable for large cell numbers (less time consuming)
- Conventional chip assays
  - high cell number (due to low recovery rate of cross-linked DNA from total cellular DNA)
  - Multiple washes during the procedure leading to loss of specific interactions.
- Lesser time requirement sonicating water bath to improve the rate of antibody-antigen binding
- Increased recovery efficiency by using a Chelex resin to combine cross-linking reversal and DNA purification .
- After incubation, the tubes are spun and DNA containing supernatant can be directly used in PCR.
- These simple modifications reduced the amount of time required for ChIP assay from 2–3 days to 4 hours.

Carrier ChIP

- Immunoprecipitation from very few cells up to 100 cells.
- histone modifications associated with developmentally regulated genes.
- Immunoprecipitation of such a small amount of chromatin is facilitated by the addition of carrier

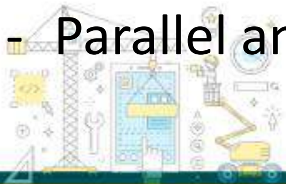
  - Heterogeneous chromatin (Drosophila or any other species which is evolutionarily distant from the species being investigated)

- The low amount of chromatin is detected by radioactive PCR and phosphorimaging.
- This technique, however, requires the primers to be designed with high specificity to prevent any spurious amplification of carrier DNA instead of the target chromatin.

ChIP-Chip

- Combination of “Chromatin Immunoprecipitation” with “Microarray technology”.
  - Labelling the immunoprecipitated DNA fragments with a fluorescent dye such as Cy5 or Alexa 647 and combining it with the genomic DNA labelled with Cy3 (reference DNA).
  - This probe mixture is then applied for hybridization over the microarray chip (whole genome)
  - Characterization of the regions of the DNA enriched by immunoprecipitation

- Advantages:
  - Probing of a large number of genomic regions in a single experiment
  - Commercially available platforms to study the localization of protein binding
  - Parallel analysis of different genes



JAN 2024

ChIP sequencing

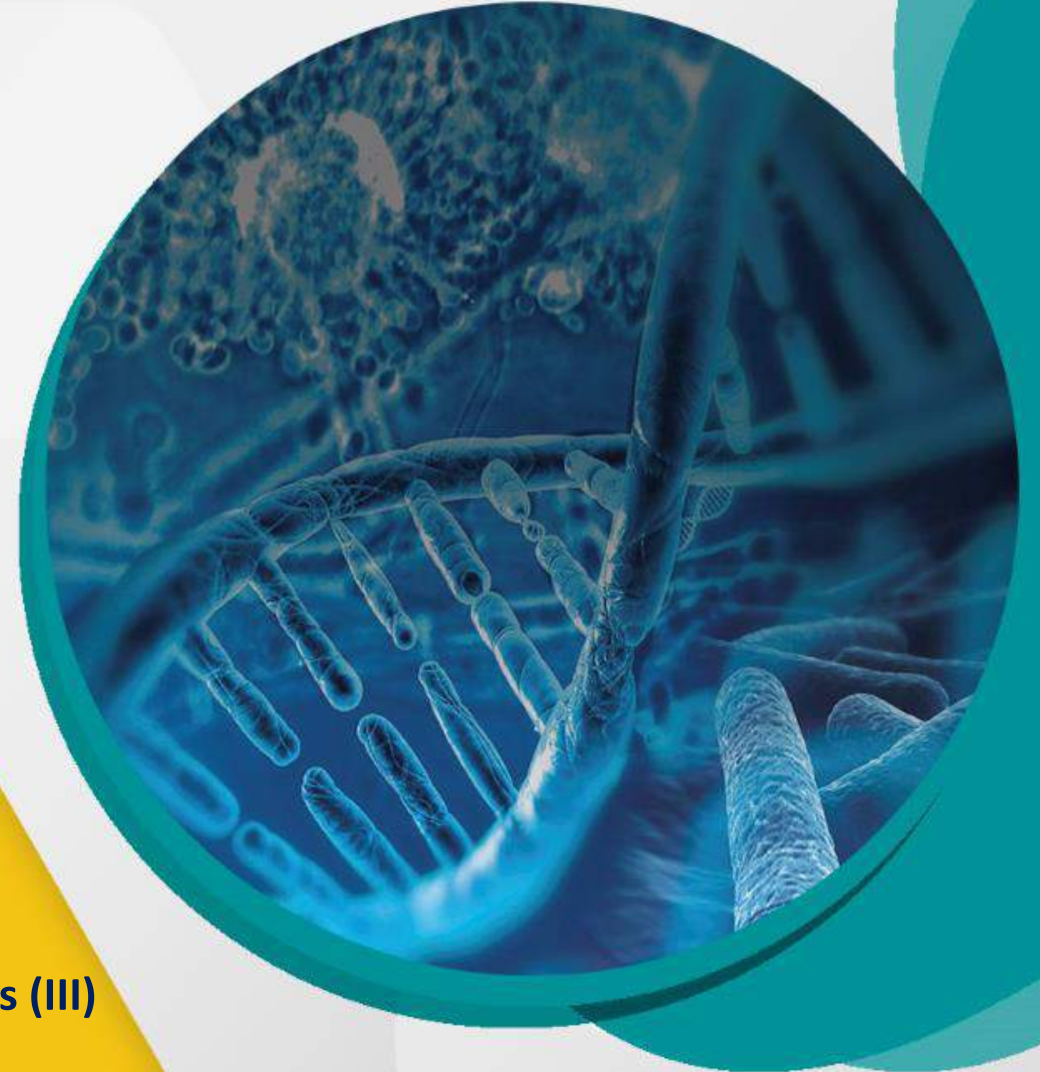
- Combination of chromatin immunoprecipitation and DNA
- ChIP-seq library –
  - The ends of enriched DNA fragments are blunted and phosphorylated using T4 kinase.
  - Adenine is added using Taq and an adapter is ligated to both the ends of the fragment
- PCR based amplification of the library
- DNA fragments of length 100–300 bp (Tag) are selected and sequenced.
- Tags are analyzed computationally with the help of alignment tools using a particular genome as reference to identify the enriched sites
- Detection of histone modifications, nucleosome positioning and mapping of binding sites of various DNA-binding proteins, distinguishing alleles on the basis of difference in SNPs

JAN 2024

REFERENCES

- Biochemistry by Lubert Stryer 9th Editon
- Lehninger Principles of Biochemistry 8th Edition
- Harper's Illustrated Biochemistry 31st Edition
- Molecular Cloning: A Laboratory Manual by David W. Russell and Joseph Sambrook
- Wilson, K. and Walker, J. (2010) Principles and Techniques of Biochemistry and Molecular Biology. Cambridge University Press, Cambridge.
- Dey, B., Thukral, S., Krishnan, S. et al. DNA–protein interactions: methods for detection and analysis. Mol Cell Biochem 365, 279–299 (2012). https://doi.org/10.1007/s11010-012-1269-z
- Ferraz, R.A.C., Lopes, A.L.G., da Silva, J.A.F. et al. DNA–protein interaction studies: a historical and comparative analysis. Plant Methods 17, 82 (2021). https://doi.org/10.1186/s13007-021-00780-z
- Cozzolino F, Iacobucci I, Monaco V, Monti M. Protein-DNA/RNA Interactions: An Overview of Investigation Methods in the -Omics Era. J Proteome Res. 2021;20(6):3018-3030. doi:10.1021/acs.jproteome.1c00074

THANK YOU



SWAYAM NPTEL COURSE ON

COMPREHENSIVE MOLECULAR DIAGNOSTICS AND ADVANCED GENE EXPRESSION ANALYSIS

By

Prof. Aritri Bir

Dr. B. C. Roy Multi-Speciality Medical Research Centre
Indian Institute of Technology Kharagpur

Module 04: Tools of Molecular diagnostics and Gene expression Analysis (III)

Lecture 20: Epigenetics & DNA methylation analysis

CONCEPTS COVERED

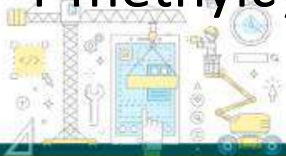
- Basic concepts of epigenetic mechanisms
- Detection methods for
    - DNA methylation
    - Histone modifications
    - Non-coding RNAs (ncRNAs)

- Epigenetics encompasses heritable structural and biochemical alterations of the chromatin without changing DNA sequence.
- Epigenetic mechanisms manipulate various physiological and pathological processes through regulations of relevant gene expressions via changing the accessibility of epigenetic codes to the chromatin locally and globally.
- Three primary epigenetic
    - DNA methylation
    - Histone modifications
    - Non-coding RNAs (ncRNAs)

JAN 2024

DNA methylation

- DNA methylation involves the transfer of a methyl group to the C5 position of cytosine within CpG dinucleotides via *DNA methyltransferases* (DNMTs).
- CpG dinucleotides often cluster in regulatory gene regions, forming CpG islands - crucial for gene transcriptional regulation by modulating chromatin structure.
- Hypermethylation of CpG islands in promoters leads to transcriptional silencing, while hypomethylation activates gene transcription.
- Biological processes
- Aberrant DNA methylation in promoter regions of tumor-related genes can silence tumor suppressor genes or activate oncogenes
- 5-methylcytosine (5mC) - most widely distributed methylation type in eukaryotes
    - N6-methyladenine (6mA) and
    - 4-methylcytosine (4mC)



Bisulfite Treatment-Based Methods:

- Sodium bisulfite treatment deaminates unmethylated cytosines to uracil, leaving methylated cytosines intact.
- Gold standard for DNA methylation detection, offering single base-pair resolution.
- Locus specific analysis:
  - Direct bisulfite PCR product sequencing
  - Methylation-specific PCR (MSP) –
    - analysis of methylation patterns in CpG islands.
    - bisulfite-converted DNA used as the template for subsequent PCR processes
    - methylation status in the specific loci of interest by using specific primers for recognizing methylated or unmethylated DNA template
  - Pyrosequencing - quantitatively monitor the real-time incorporation of nucleotides through bioluminometric detection of pyrophosphate

Bisulfite Treatment-Based Methods:

- Genome-wide DNA methylation profiling:
  - Hybridization-based microarray followed by bisulfite conversion - using predesigned probes for methylated and unmethylated CpGs
  - Whole genome bisulfite sequencing (WGBS) - comprehensive methylation state, including low CpG-density regions such as intergenic regions, partially methylated domains and distal regulatory elements
  - Sophisticated bioinformatics analysis
  - Enrichment techniques followed by bisulfite sequencing such as reduced representation bisulfite sequencing (RRBS)
    - Integrated with restriction enzyme digestion, bisulfite conversion and NGS

Evaluation of oxidized forms of DNA methylation

- Ten-Eleven Translocation enzymes (TET1, 2, and 3) - DNA demethylation
  - TET enzymes generate oxidized forms of 5mC, including 5hmC, 5-formylcytosine, and 5-carboxylcytosine.
- oxidative bisulfite sequencing (OxBS-seq) :modified bisulfite sequencing technique
  - identifies 5hmC by oxidizing it to 5fC, leaving 5mC intact during bisulfite treatment.
- hmTOP-seq (5hmC-specific tethered oligonucleotide-primed sequencing) _ high-resolution bisulfite-free method
  - directly reads sequences primed at covalently labeled 5hmC sites from an in situ tethered DNA oligonucleotide.

Affinity Enrichment

- Specifically capture methylated DNA sequences from a complex genomic sample.
- Based on affinity of certain proteins or antibodies for methylated DNA -selective isolation and analysis of methylated DNA fragments.
    - Genomic DNA Shearing: either by sonication or enzymatic digestion.
    - Immunoprecipitation:
        - The fragmented genomic DNA is denatured to generate single-stranded DNA.
        - Antibodies specific to 5-methylcytosine (5mC) are added to the denatured DNA sample.
    - Magnetic Bead Capture : antibodies bound to methylated DNA fragments are captured onto the magnetic beads through protein A/G interaction.
    - The resulting enriched fraction and input control are differentially labeled and hybridized to an array platform.
- Methylated DNA immunoprecipitation (MeDIP)-chip method

Using Restriction Enzymes

- Methylation-sensitive restriction enzymes (MSREs) recognize specific DNA sequences and cleave them only if the recognition site is unmethylated.
    - HpaII - recognizes 5'-CCGG-3' but cleaves it only if the internal cytosine is unmethylated.
- Methylation-dependent restriction enzymes (MDREs) - methylation of specific DNA sequences for cleavage.
    - MspI recognizes 5'-CCGG-3' and cleaves it if the cytosine is methylated.
- Following size selection, fragments of interest can be labeled and hybridized to pre-designed arrays.
- Comparative approaches
    - *Hpa*II tiny-fragment enrichment by ligation-mediated PCR (HELP)
        - *Hpa*II enzyme (methylation sensitive) and its isoschizomer *Msp*I (methylation insensitive)
    - Methylated CpG island (CGI) amplification in combination with microarrays (MCAM)
    - Differential methylation hybridization (DMH)
    - Comprehensive high-throughput arrays for relative methylation (CHARM)

Chromatin modification

- Posttranslational modifications of histone proteins, including acetylation and methylation of specific lysine residues on core histone tails.
- Increased histone acetylation is associated with transcriptional activation, while deacetylation leads to gene silencing.
- Histone methylation patterns are site-specific and methylation-dependent
  - methylation of H3K4 activates transcription
  - H3K9 methylation represses it
- Histone modifications alter chromatin structure, affecting DNA accessibility to transcription factors and epigenetic modulators at gene regulatory regions.
- Enzymes
  - histone acetyltransferases (HATs)
  - histone deacetylases (HDACs)
  - histone methyltransferases (HMTs)

JAN 2024

Chromatin modification

Chromatin immunoprecipitation (ChIP)

- specific antibody-directed towards the specific marker causing histone modifications or epigenetic modulators in conjunction with specific DNA fragments
- ChIP integrated with other techniques - interaction of histone modifications with other chromatin regulators at the specific loci
- If the target histone modification and DNA regulatory region are specific - ChIP followed by qPCR to reveal the enrichments of specific histone modifications or binding ability of a remodeling complex to the specific DNA region
- If the specific modification patterns are undefined - sequencing-based ChIP methods such as ChIP-chip or ChIP-seq
    - ✓analysis of protein-DNA binding events and histone modification enrichment at large numbers of loci simultaneously

Integrative DNA Methylation and Histone Modification Analysis

- Dynamic Interactions between DNA Methylation and Histone Modification
- Challenge to determine the complex interactions between methylation and chromatin.
- ChIP-bisulfite methylation sequencing (ChIP-BMS)- determine the methylation status of ChIP DNA pulled-down by a specific antibody (histone markers or transcription factors)
- BisChIP-seq (bisulfite-treated chromatin immunoprecipitated DNA) – global profiling
- SCAN (Single Chromatin molecule Analysis in Nanochannels) combines epigenetic modification-specific antibodies with single-molecule imaging techniques for high-throughput assessment
- Advanced Methods for Characterizing 3D Chromatin Organization and the Epigenome

Techniques for Evaluating Histone Modifications

Detect histone modifications (type and abundance):

- Immunoblot analysis including dot blots and WB
- Whole cell lysates / enriched samples for histone proteins (nuclei or chromatin isolation)

Identify modified-histone interacting proteins:

- Immunoprecipitation and WB
- Enzymatic processing of chromatin to generate single nucleosome units.
- Antibodies specific to the chromatin binding protein - to pull down the complex.
- Subsequent WB analysis with antibodies to unmodified and modified histones

Determine the genomic location of histone marks:

Chromatin immunoprecipitation (ChIP) is used in combination with

- PCR
- NGS (ChIP-seq):
- Microarray (ChIP—chip)

Molecular diagnostic for ncRNAs

- ncRNAs - group of RNA transcripts that do not encode proteins like mRNAs.
- ncRNAs can be divided into two categories based on their regulatory roles:
  - Housekeeping ncRNAs (rRNA & tRNA) - regulate generic cellular functions
  - Regulatory ncRNAs that actively mediate gene expression during complex molecular and cellular processes.
  - small ncRNAs that are normally less than 200 nucleotides - miRNA, siRNA and piRNA
  - lncRNA exceeds 200 nucleotides in lengths.

Methods for Small ncRNAs Detection

- Methods for profiling:
  - qRT-PCR
  - hybridization-based microarrays
  - high-throughput RNA sequencing

- Investigation of ncRNA-based targets:
  - high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP) - identification of functional protein-RNA interaction sites
  - ligation of interacting RNA followed by high-throughput sequencing (LIGR-seq) - profiling of miRNA-mRNA interaction

Methods for lncRNAs Detection

- lncRNA: non-coding transcripts greater than 200 base pairs.
    - transcribed by RNA Pol II from independent promoters
    - caps and poly(A) tails
    - exhibit extensive mechanistic diversity through interaction with RNA binding proteins (RBPs) at specific DNA regions

- **Detection and profiling methods:**
    - qPCR
    - Genome-wise chromatin binding site oflncRNA: ChIRP (Chromatin isolation by RNA purification) and CHART (Capture Hybridization Analysis of RNA Targets)
    - RAP (RNA antisense purification) - map the localization of a given lncRNA across the genome
    - protein-RNA interaction : RIP (RNA Immunoprecipitation)-associated array or sequencing such as RIP-chip and RIP-seq
    - RNA pull-down followed by (LC-MS/MS) - identify the interaction proteome of target lncRNA

Summary

- DNA methylation, histone modifications, and non-coding RNAs orchestrate gene expression and cellular processes through epigenetic regulation.
- Bisulfite sequencing stands as the gold standard for DNA methylation detection, offering single-base resolution by converting unmethylated cytosines to uracil while leaving methylated cytosines intact.
- Bisulfite sequencing enables locus-specific analysis through techniques like direct bisulfite PCR product sequencing, Methylation-Specific PCR (MSP), and pyrosequencing.
- ChIP can be integrated with NGS. Microarray etc .
- Techniques such as Methylated DNA Immunoprecipitation (MeDIP) and Methylated CpG Island Amplification with Microarrays (MCAM) enrich methylated DNA fragments for comprehensive DNA methylation analysis.
- RNA profiling techniques, including qRT-PCR and high-throughput RNA sequencing, allow for the investigation of non-coding RNAs (ncRNAs) and their regulatory roles in gene expression.
- Understanding the interplay between DNA methylation, histone modifications, and ncRNAs offers insights into the complex regulatory networks governing cellular processes and disease pathogenesis.
- Integration of multiple epigenetic profiling techniques with advanced bioinformatics analysis enhances the efficiency

REFERENCES

- Biochemistry by Lubert Stryer 9th Editon
- Lehninger Principles of Biochemistry 8th Edition
- Molecular Cloning: A Laboratory Manual by David W. Russell and Joseph Sambrook
- Wilson, K. and Walker, J. (2010) Principles and Techniques of Biochemistry and Molecular Biology. Cambridge University Press, Cambridge.
- Dey, B., Thukral, S., Krishnan, S. et al. DNA–protein interactions: methods for detection and analysis. Mol Cell Biochem 365, 279–299 (2012). https://doi.org/10.1007/s11010-012-1269-z
- Cozzolino F, Iacobucci I, Monaco V, Monti M. Protein-DNA/RNA Interactions: An Overview of Investigation Methods in the -Omics Era. J Proteome Res. 2021;20(6):3018-3030. doi:10.1021/acs.jproteome.1c00074
- Alexander Meissner, Andreas Gnirke, George W. Bell, Bernard Ramsahoye, Eric S. Lander and Rudolf Jaenisch. 2005. "Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis". Nucleic Acids Res. 33(18):5868-77
- Li Y. Modern epigenetics methods in biological research. Methods. 2021 Mar;187:104-113. doi: 10.1016/j.ymeth.2020.06.022. Epub 2020 Jul 6. PMID: 32645449; PMCID: PMC7785612.
- Mehrmohamadi M, Sepehri MH, Nazer N, Norouzi MR. A Comparative Overview of Epigenomic Profiling Methods. Front Cell Dev Biol. 2021;9:714687. Published 2021 Jul 22. doi:10.3389/fcell.2021.714687

THANK YOU