PAGE NO.
DATE: 02/03/2020

Genomics & Proteomics
Assignment No: 02

Yogita Salgar
Msc II Biotech
72155

Question :-

(1) What is protein-protein interaction? Explain various methods to study protein-protein interaction? Explain in detail yeast two hybrid system

Ans 1) Protein-Protein interaction

- Protein-protein interactions are central to virtually every biological process and many different analytical methods have been developed to study them.
- Most of these methods are suitable for studying interactions within a small group of proteins, but cannot be employeed on a proteomic scale.
- Protein-Protein interaction plays a key role in predicting the protein function of target protein and drug ability of molecules.
- There are in vitro, invivo and insilico methods to study protein-protein interaction.
- Experimental determination of PPI i.e invitro or inviwo includes Yeast-two hybrid and phage display methods.
- Insilico methods include databases for PPI like database of ligand protein (DIP), database of ligand recepter partners (DLRP), Biomolecular interaction network database (BIND) etc.

1) Phage Display-

- Phage display allows the presentation of large peptide & protein libraries on the surface of filamentous phage which leads to the selection of peptides and proteins including absorbance with increase in affinity and specificity to almost any target.

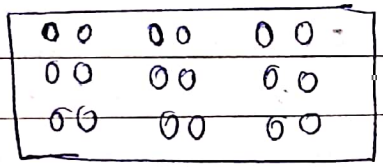
- The Technology involves introduction of exogenous peptides sequences into a location in the genome of phage capsid proteins. This encoded peptides are expressed or displayed on phage surface as a fusion product with one of phage coat proteins.
- The strength of phage display is its ability to identify interactive regions of proteins and other molecules without preexisting notions about nature of interaction
- It can be used for epitope peptides, peptide ligands, enzyme substrates / single chain Ab-fragments.

Bacterisally expressed GST (Glutathione-S-Transferase) Fusion proteins or domains

↓

Immobilize individually into wells

↓



↓

Wash unbound particles of phage

↓

Amplify phage particles

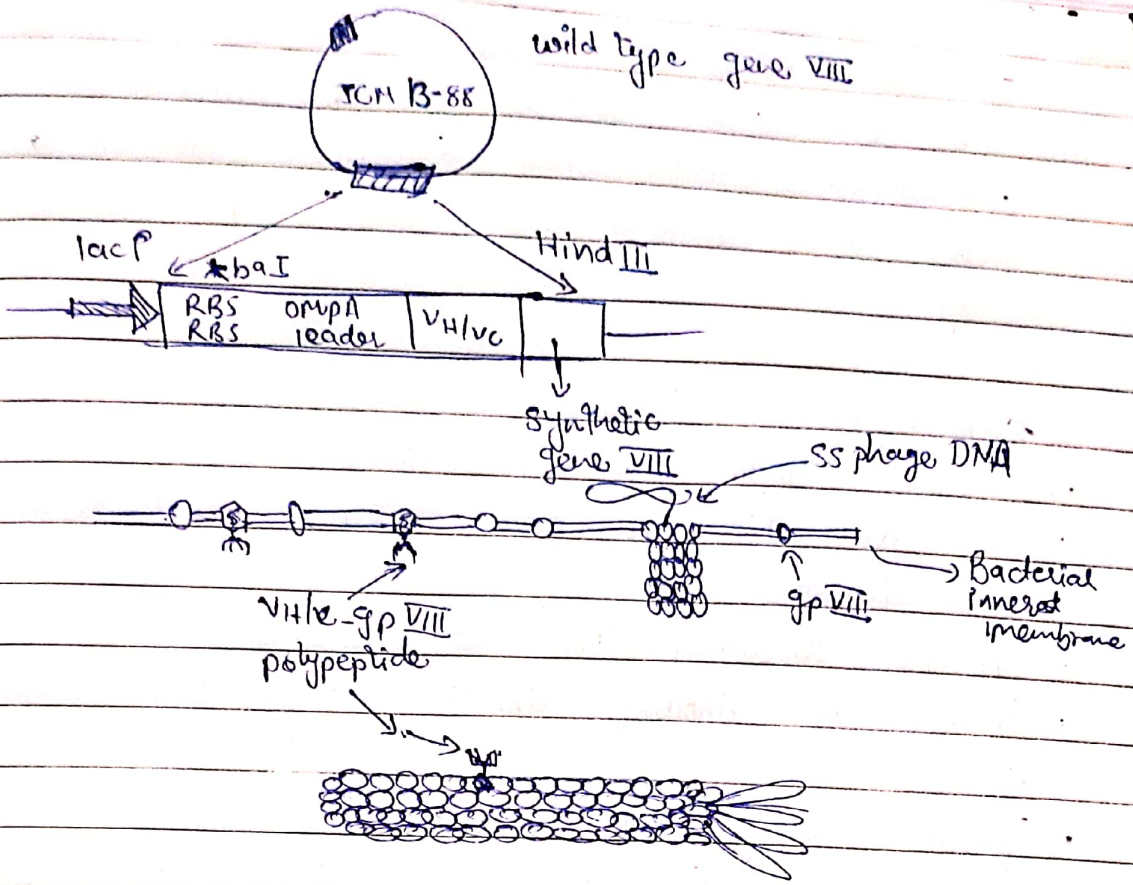
↓

Sequence the cDNA insert

- Steps involved in phage display →

(1) Creation of viral vectors :

- Recombinant DNA technology to insert foreign cDNA of interest into viral DNA.
- Spliced into gene for a coat protein so the protein will be displayed on outside of phage particles.



② Binding / selection or biopanning -

- Once the insertion of gene encoding any protein into the genome, it get displayed or expressed on surface of bacteriophage as a fusion product of protein of interest & coat protein. This then apply to standard affinity techniques to capture phage by showing protein which is displayed interact with its ligand.

③ Washing :- Unbound phages are then washed away leaving only those showing affinity for the receptor / ligand.

④ Elution :- Bound phages then eluted out by disrupting protein bounding interactions solution can be acidic buffer, alkaline buffer, urea, addition of soluble ligand for receptor or addition of soluble ligand for receptor or addition of host cells to infect in order to recover the bound phages.

⑤ Amplification - Eluted phages showing specificity are used to infect new host cells for amplification cycle repeated 2/3 times for stepwise selection of best binding

Sequence.

(6) Analyse - Final phages can be propagated then characterised with DNA sequencing. Common motif involved with binding may emerge for further study.

Applications -

- Epitope mapping and mimicking.
- For identification of new receptors and ligands.
- Drug discovery.
- Epitope discovery for new vaccines.
- Creation of Antibody libraries, and organ targeting.

(1) Epitope mapping and mimicking -
Use random libraries to determine if it is continuous then compare phage sequence motif to a.a. sequence of natural ligands. And then map critical binding sites of epitope / ligand.

(2) Drug discovery -
Test receptors as targets of drugs. Peptides can be acting as antagonists, agonists or modulators.
Large scale search but might not have good pharmacological properties.

(3) Epitope discovery - Use Antibodies as a receptor to select peptide that is antigen mimic. Then use this to mimic immunization and elicit antibodies.
This bypasses animal immunization by mimicking immune selection.

Advantages -

(1) Easy to screen large clones.
(2) Easy to amplify selected phages in E.coli.
(3) Selection process easy and already in use in various forms.

(4) Can create phage library variation by including mutation using error prone PCR.

Disadvantages-

(1) Could lose Phage variations if first bind / wash step is stringent.
(2) Might not have enough peptide insert so critical folding can be disrupted.
(3) Affinities / binding that results during selection might not work invivo.

Yeast - Two - Hybrid Method

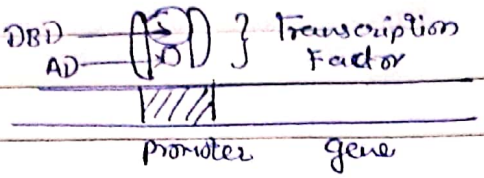
- Classical yeast - two - hybrid system -

Principle -

- Proteins are consisting of severally discrete domains. (Structurally and functionally independent)
- Domains can function through non-covalent interaction when brought together.
- For example -

Transcription Factor
- DBD (DNA Binding Domain) helps in locating promoter of target gene
- AD (Activation domain) → Recruits RNA pol III Complex.

- If DBD & AD expressed as separate peptide - Transcription factors doesn't work and no transcription occurs.
- It means DBD & AD works together and alone cannot activate / start transcription.
- On this basics, it is possible to use yeast - two hybrid system to confirm interactions between known proteins and to screen for unknown protein that interact with protein of interest.



Steps -

cDNA for protein of interest (x) is cloned into bait vector (fusion of DBD & protein of interest termed as bait)

↓

Bait is translocated to nucleus of yeast cell

↓

Binds to promoter located upstream of reporter gene

↓

No activation as bait lacks activation domain

↓

Second cDNA encoding interacting protein (Y) is cloned in prey vector (fusion of AD & Y)

↓

prey vector also cannot activate as it lacks DBD

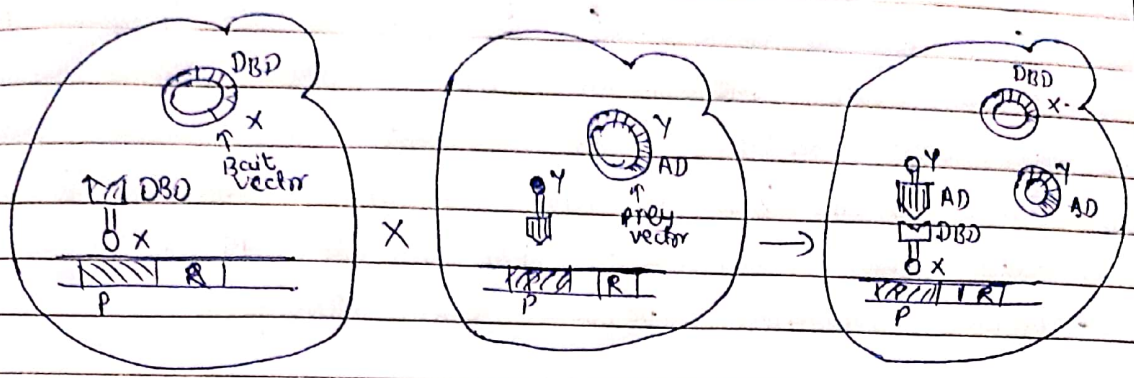
↓

But if protein X & Y have capability to interact the bait captures the prey.

↓

genes will be activated creating hybrid Transcription factor.

- Reporter genes used are - auxotrophic markers (HIS 3, ADE2, URA3)
  selectable marker lac z
- Success of yeast two hybrid depends on
  (i) Positive clones are selected on basis of auxotrophic markers.
  (ii) Selected clones are analysed further using convinient colour assay.



X → protein of interest, Y - interating protein
AD → Activating Domain, DBD - DNA Binding Domain
P - Promoter, R - Reporter gene.

- Classical YTH example -
  - Yeast two hybrid used to produce list of interactions of 6000 yeast proteins.
  - Each ORF of yeast genome was amplified using PCR.
  - cloned into two separate vectors → one having DBD other having AD.
  - Bait vector → MCS downstream to GAC-4 → GAL-4-DBD → 3' fusion of this to X → GAL-4-DBD-X
  - Prey vector → MCS upstream to GAL-4-AD → fusion (5') to Y → (Y-GAL-4-AD)
  - Both fusion constructs transformed into yeast cells of different mating types.
  - Two sets of 6000 transformants formed.
  - 6000 x 6000 matings.
  - diploid cells selected on minimal medium.
  - when X & Y interacted with each other → reporter gene would have been switched on.

* Variation of yeast two hybrid system -

- Limitation of classical methods is that it relies on protein interacting with nucleus → proteins having hydrophobic transmembrane or domains are not

Transported to nucleus. If transported, not properly folded.
- Some need modification by cytoplasmic or membrane associated enzymes for interaction with binding partner
- Some protein bind to RNA or small molecules to alter their confirmation for promoting PPI.
- Hence, modified versions are developed.

(a) Reverse two hybrid & Split-hybrid System-
Interaction of X & Y induces the transcription of a reporter gene that confers toxicity to yeast. For example - selection can be induced by addition of 5-fluoro-orotic acid (FOA) which is converted to toxic compound 5-fluorouracil by URA 3 gene product.

(b) Three hybrid System-
Third protein Z is expressed along with DBD & AD fusion proteins. Expression of reporter gene → only in presence of Z. RNA three hybrid is developed to analyze RNA-protein interactions → DBD & AD hybrid proteins are linked via binding of a bifunctional RNA molecule.

(c) Split ubiquitin System-
Ubiquitin, small protein (76 a.a) → act as tag for protein degradation. N-terminal ubiquitin (Nub) mutant that is unable to interact with C terminal ubiquitin (ub) of its own is fused to one protein. Cub reporter hybrid is fused to its prospective partner. Interaction between two proteins, reconstitution of ubiquitin, cleavage and release of reporter gene.

(d) SOS Recruitment System (SRS)

- It uses Ras pathway in yeast.

* RAS guanin nucleotide exchange factor RGNG cdc 25 → positioned at plasma membrane
↓
Stimulation of GDP/GTP exchange on Ras
↓
Downstream signalling
cell growth

→ Mutation in Yeast having temperature sensitive cdc-25
↓
grow at 25°C & not at 36°C

Human RGEF (hSOS) → Complements the mutation → hSOS mutant strain
↓
grow at 36°C also

- Translocation of hSOS is dependent on PPI
- Bait X is fused to C-terminally truncated hSOS, which is active but unable to target to plasma membrane
- prey Y - integral membrane protein or soluble protein anchored to membrane by myristoylation.
- interaction between X & Y recruits hSOS to the membrane cell growth.

Applications of YTH-

- Investigation of PPI, protein-nucleic acid interaction protein small molecule interaction.
- Identification of drug Targets → using ligands.
- Genome wide interaction mapping. -
  two methods are used.

(1) Matrix interaction screening -
construction of panels of defined bait & prey mated systemically in array format.

(2) Random library method -
Random clones from a highly complex expression library represents baits & prey.

Defined ORF's as bait → prey can be screened.
In matrix interaction, all constructs are predefined & interactors can be traced on the basis of grid position in array.
- In random library → interacting clones are characterised by sequencing and then compared to sequence databases for annotation.

Ques) Explain the importance of mass spectrometry in proteomic study?

Ans2) A mass spectrometer is an instrument that can measure the mass/charge ratio m/z of ions in a vacuum.
- From these data, molecular masses can be determined with a high degree of accuracy, allowing the molecular composition of a given sample or analyte to be determined.
- In proteomics, the analyte is usually a collection of peptides derived from a protein sample by digestion with Trypsin or a similar reagent.
- For example:- Using the data obtained from mass spectra for protein indentification.
- This protein identification method uses data from intact peptide masses which refers to as peptide mass fingerprinting (PMF)
- This method uses 2DGE and MALDI-TOF mass spectrometry, where proteins are separated before digestion into peptides.
- The principle of the technique is that each protein can be uniquely identified by the masses of its constituent peptides, this unique signature being known as the peptide mass fingerprint.

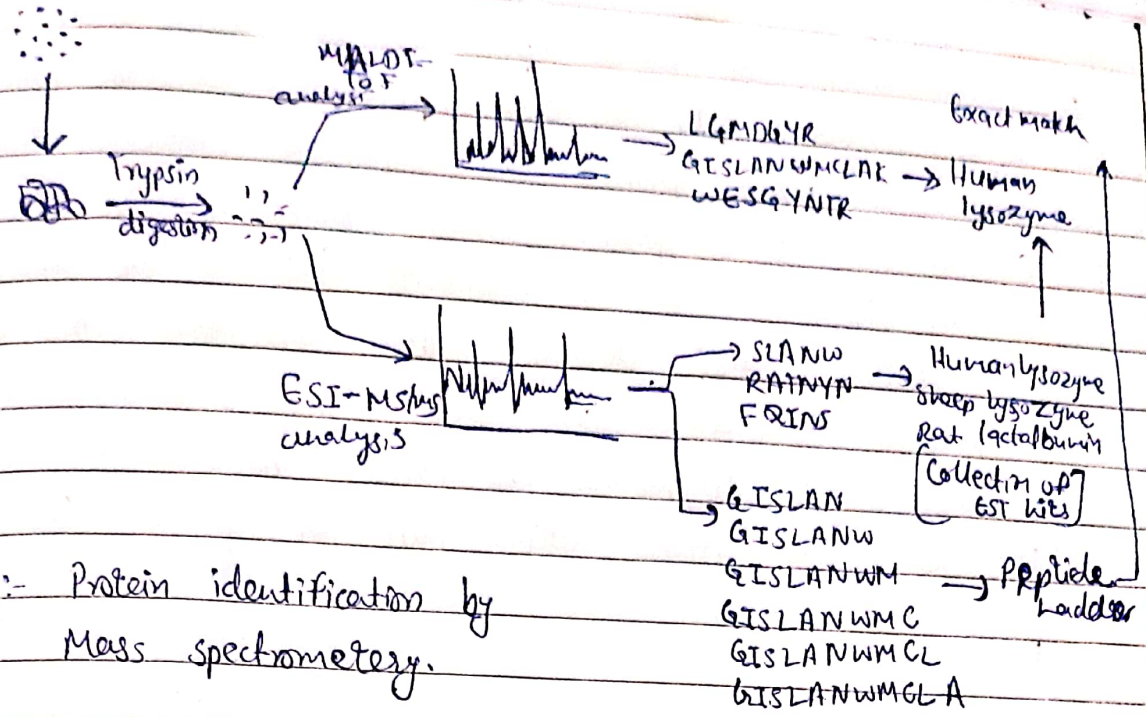


Fig:- Protein identification by Mass spectrometry.

- In a typical strategy, digested peptides are analysed by MALDI-TOF mass spectrometry in order to determine the masses of intact peptides. These masses can be used in correlative database searches to identify exact matches.
- If this approach fails, ESI-MS/MS analysis can be used to generate peptides fragment ions. These can be used to search less robust data sources (for example, EST databases) and to produce de novo peptide sequences.

- Mass spectrometry has arguably become the core technology in proteomics.

- Also through mass spectrometer, three types of analysis can be carried out -

(a) For intact peptide ions - This allows the masses of intact peptides to be calculated, and these masses can be used to identify proteins in a sample by correlative database searching.

(b) For fragmented peptide ions - This allows the masses of fragments to be determined, and these can be used in correlative database searching, or to derive de novo sequences, or in hybrid approaches.

(c) Fragmentation of whole proteins-
This is known as top-down proteomics because one begins with an intact protein. The approach became more challenging with proteins whose mass is greater than 50 kDa approx. 500 a.a. residues. This is the upperlimit for reliable mass detection using standard MS due to inefficient ionisation and poor protein stability. The analysis of peptides or fragments is called bottom of proteomics. Because one starts with part of protein than whole protein.

- MS/MS spectra can be used to derive protein sequences de novo. That is uninterpreted MS/MS spectra can only be used to search full protein sequences in the major databases, partially interpreted spectra can be used to search lower-quality sources such as expressed sequence tags (ESTs).
- The alternative denovo sequencing approaches is a good example for how Edman degradation and mass spectrometery can be used together to generate sequence information.

- MS is most importantly used to study biomarkers of disease states in humans. Variants of mass spectrometry like SELDI-TOF, MALDI-TOF/TOF 2D LC-MS/MS, triple quadrupole are used for identification of serum biomarkers.

Que 3) Explain with the help of appropriate example any two applications of proteomics?

Ans 3) —

(P-T-O)

Vs 3) Proteomics in Drug Discovery-

- Drug discovery is a lengthy and highly expensive process that uses variety of tools from diverse fields.
- Proteins are the principle target for drug discovery, so evolution of proteomics techiques is of major importance.
- Proteomic technology and computational methods have been advanced owns other techniques.
- This enables scientists to screen large number of proteins within clinically distinct samples that help to discover disease biomarkers, identify and ~~help to discover disease~~ validate drug targets, design more effective drugs, assessment of drug efficacy and patient response i.e to interfere with almost every step in drug discovery process.
- Proteomics approach in drug discovery include finding an unstable protein that is causing an undesirable affect and then usage of molecule to modify its effect.

Identification and assignment of candidate target-

- Drug targets are protein or Signal transductin pathway. in which proteins are involved.
- Recombinant protein microarray and computatinal drug design are the two unique technique serve the purpose of identifying drug targets, target validation and 3D structure elucidation.

* Recombinant protein Microarray-

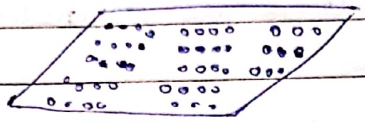
- It consist of purified active recombinant protein.
- They enable the investigation of multiple proteins simultaneously.
- This help to find interaction formed between protein and other molecules, thus offering unique assay system for studying protein function.
- This involves high throughput screening platform to identify and to validate protein targeted molecules as

potential drug candidates.

Protein | or | DNA | or | RNA | or | ligand

↓ Incubate with recombinant protein array

↓ wash away unbound probe



Detection

Strength of signal correspond to amount of bound probe

- Computational drug design-

- Proteomic technique with the help of computer software figure out target for disease which will ultimately help to discover the right drug.
- All the computer modelling involve few general objectives.
- Use of computer power to simplify drug discovery and development process.
- Utilization of chemical and biological information about ligands / targets to identify and optimize new drugs.
- Design of insilico filters to eliminate compounds with undesirable properties and select most promising candidate.

Use of proteomics in drug's mode of action and toxicity studies-

- Mode of action and toxic studies are important to obtain a new safe and effective drug molecule.
- Following is the work flow of mode of action study in the field of drug discovery. -

(P.T.O)

Target identification

↓

Target validation

↓

Lead molecule screening

↓

Target - lead pair

↓

Mode of action studies of new drug molecules

↓ ↓

Therapeutic effects Toxic effects.

- Proteins whose levels are modified in response to drug administration could provide vital clues with respect to drug effectiveness and toxicity.
- These proteins will serve as efficacy or toxicity biomarkers to guide clinical trial studies.
- Similarly analysis of protein profiles before and after pharmacological treatments could also confirm the mechanism of drug actions and provides insight for new drug discovery.

(2.) Proteomics in identification of biomarkers in disease diagnosis -

- Biomarkers are biological parameters used to detect or confirm presence of a disease or condition of interest to identify individuals with subtype of the disease.
- These molecular signatures can also be used for follow up of disease response, survival of patients as well as various parameters.
- An ideal biomarker should enable unbiased diagnosis determination, particularly in patients without specific symptoms.

- Proteomics on human samples has mainly focused on available fluids such as blood, CSF (cerebrospinal Fluid), urine and saliva.
- For biomarker application, a single biomarker is probably insufficient for accurate representation of disease.
- Therefore, multiple biomarker profiles need to be identified in different types of DNA, RNA, microRNA (miRNA) and protein including modification from DNA, gene and post-translational ~~pro~~ modification proteins.

- Protein microarray based system used to identify biomarkers which could be use for early detection as well as accurate detection.
- So, protein microarrays have greatly enhanced the biomarkers discovery process because they allow a high throughput platform for simultaneous and rapid screening of thousands of protein.
- Other techniques allows to test for proteins produced during a particular disease help to diagnose the disease quickly.
- Techniques include - western blot, immunohistochemical staining ELISA or MS.
- ~~Secretonin~~ Secretomics, a subfield of proteomics that studies secreted proteins and secretion pathways using proteomics approach, has recently emerged as an important tool for the discovery of biomarkers of disease.

Example - Biomarkers in Alzheimer's Disease.

Blood based biomarkers - Aβ, ApoE (Apolipoprotein E)
CSF based biomarkers - visinin like 1 (VLP-1), tau