Parallel analysis of translated ORF (PLATO)

Technical Journal Club

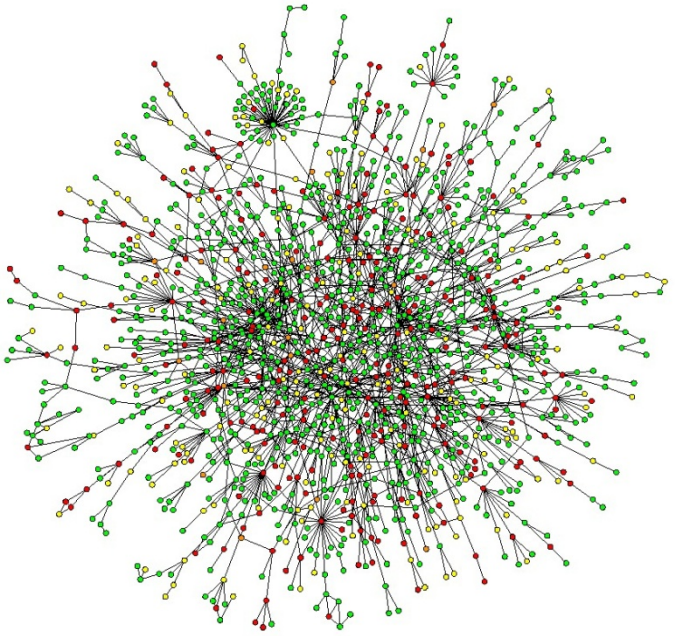
16.04.2013

Mary-Aude Rochat, PhD student
Speck group

Proteomics

Rationale: Increase in the DNA sequence information
Understand biological processes

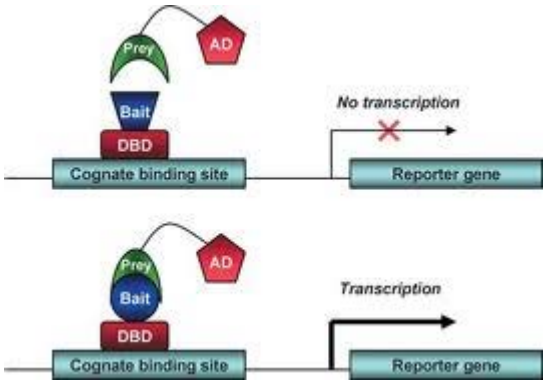
>Development of Large scale analysis of protein:
*Characterization of gene function
*Building functional linkage
*Insight into biological mechanisms
>Protein-Protein interaction map



Methods: *Mass spectrometry
*2-hybrid system
*Phage display technology
*Protein microarray

Previous methods:
Two-hybrid and split reporter

Procedure: Bait > Target protein fused to a DNA-BD
Prey> Protein fused to a transcription AD
Physical interaction
>transcription of reporter gene

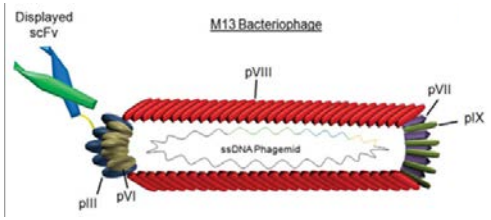


Advantages:*Powerful, quick and easy
*High-throughput

Limitations: *Not comprehensive
*False positive (50%)
*Membrane protein
(>split ubiquitin system)
*Only within a cell
>Not suitable for drug or ab target identification

Previous methods:
Phage/Phagemid display of cDNA

Procedure: *Exogenous peptide expression (fusion: pIII or pVIII)
*Selection of phage (affinity purification: specific ligand)
*Elution and amplification (E. *coli*)
*Sequencing



Advantage: Rapid generation of large librairies

Advantage Phagemid:*Larger foreign DNA fragment
*Efficient in transformation: High diversity
*MCS
*Genetically more stable than recombinant phages

Limitations: *Small portion of the protein
*Small fraction of in-frame polypeptide

>Low target cDNA in the initial library
>Highly biaised clonal abundancies

Previous methods:
Phage/Phagemid display of cDNA

Improvement:*cDNA fragmentation

>Increase the probability of expressing functional domain

>Limited by the lack of post translational modification and folding capacity of the host

*Directional cloning

>Priming of mRNA instead of cDNA (interference stop codons)

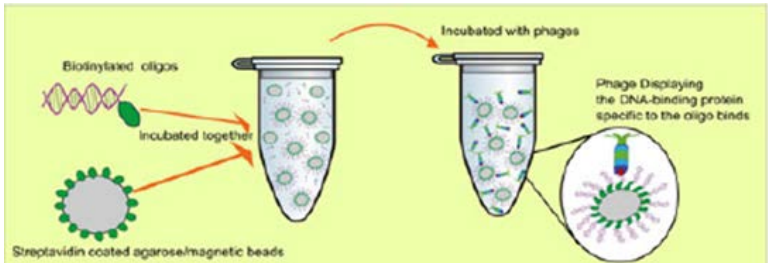
>Maintain the native orientation of the fragment

>Double the probability to obtain inserts with continuous ORFs

*ORF selection (ampicillin)

>Improve the yield of full cDNA expression

>Avoid premature stop codons



Previous methods:
Protein microarray

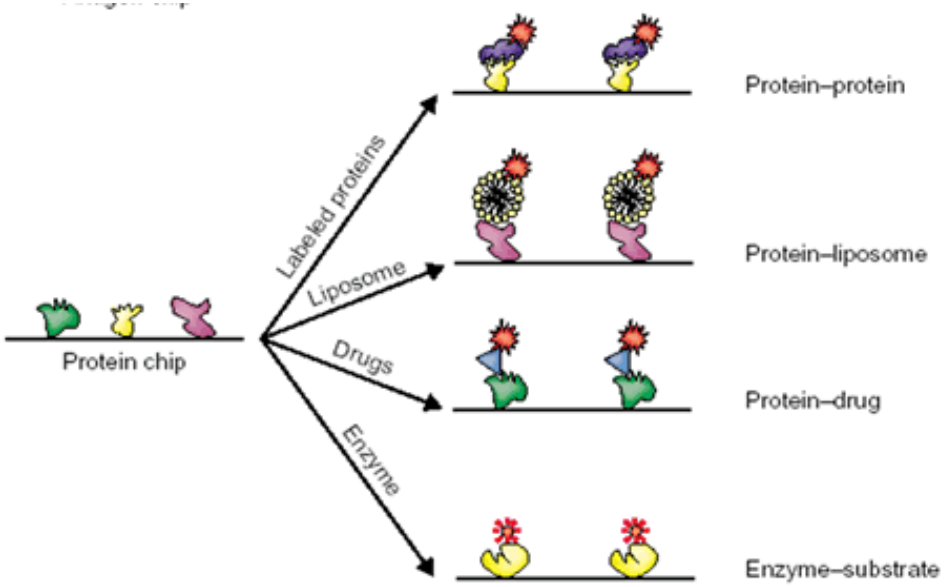
Detection: Interaction between protein-lipid, drugs, enzyme-substrate and disease biomarkers

Procedure :-Baits bound on a support
-Molecule of interest is tagged with a fluorescent dye
-Detection by fluorescence

>Pattern of +/-spots
>Signal intensity is proportional

3 types of microarrays:

-Analytical (capture array)
-Functional protein
-Reverse phase microarray
(post-translational modification altered in disease)



Previous methods:
Protein microarray

Advantages: -High number of interacting partners
-Quantitative
-Rapid, automated, and highly sensitive
-Post translational modification

Limitations: -In vitro assays
-Cross-reactive contaminants
-Denaturation
-Conjugaison with Tag

Protein Interaction discovery using parallel analysis of translated ORFs (PLATO)

Jian Zhu[1–3,10], H Benjamin Larman[1–5,10], Geng Gao[1,3], Romel Somwar[6], Zijuan Zhang[7], Uri Laserson[2,4,8], Alberto Ciccia[1–3], Natalya Pavlova[1–3], George Church[2,4], Wei Zhang[7], Santosh Kesari[9] & Stephen J Elledge[1–3]

nature biotechnology

AIM

- Identify physical interaction between proteins and others molecules

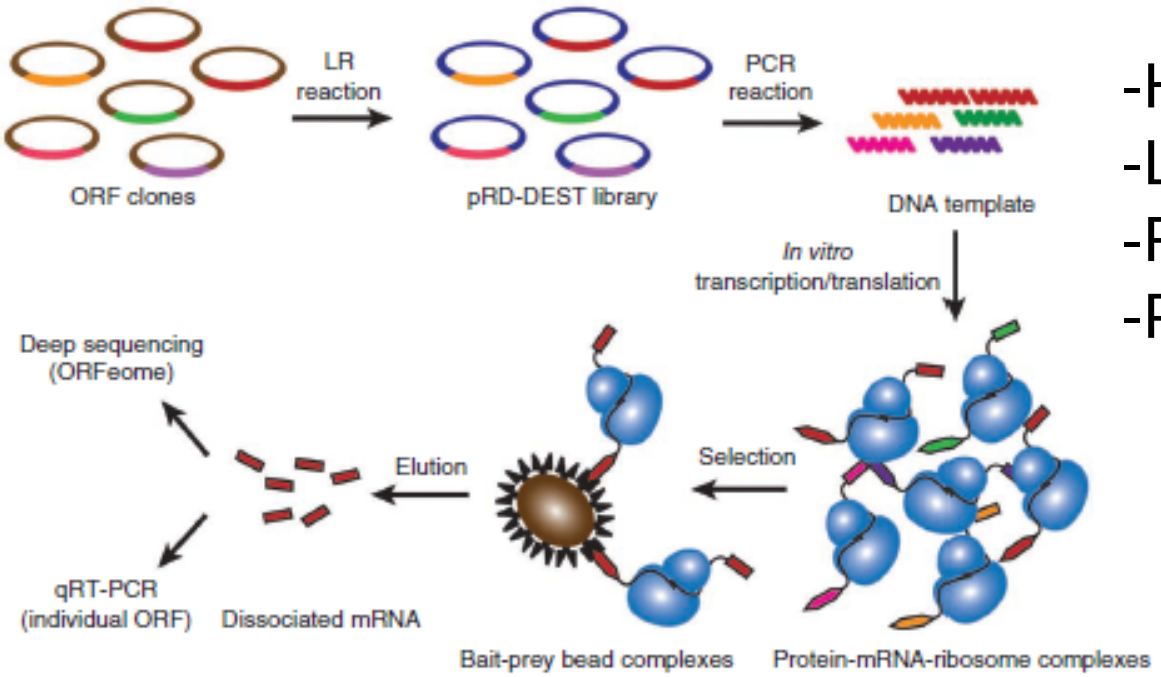
PLATO

- Combine:
  - In vitro display of full length protein
  - High-throughput DNA sequencing

Confirmation on:

- LYN kinase
- Patients auto-antibodies
- Small molecules: Gefitinib

Method

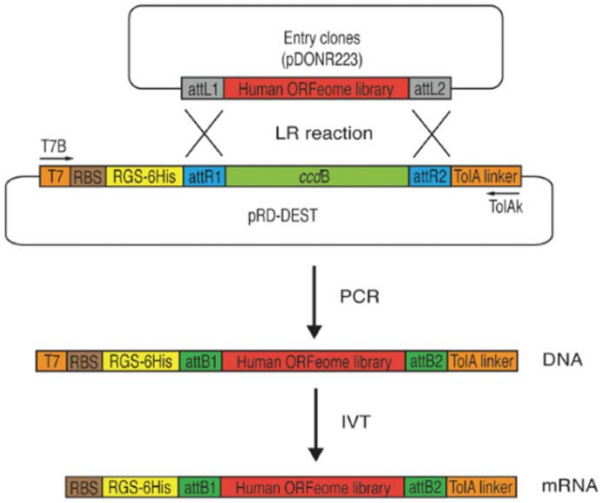


-Human ORFeome library
-LR reaction/superpools
-Recombination (attL-attR)
-PCR amplification

-In vitro txp (T7)
-In vitro translation (RTS)
-Affinity purification
-Elution/mRNA purification

Ribosome display vector

Gateway vector



Gateway cloning system

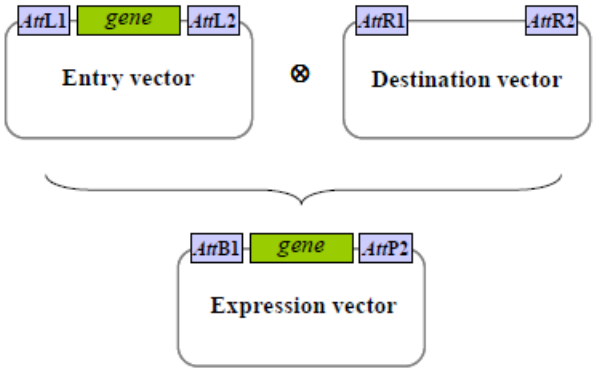
-Commercialized cloning strategy

-Based on DNA recombination mechanisms utilized by the phage Lambda for integration

-Efficient transfer of DNA fragment: recombination sequence and clonase enzymes (BP and LR reaction)

-Maintain the reading frame: specific sites

>Allow functional analysis



-Low background rates

-No requirements on the sequence to be cloned

-Large libraries of Gateway-adapted ORFs have been created by both academic and commercial entities

Method: ribosome display

- In vitro translation: Physical coupling of polypeptide with mRNA

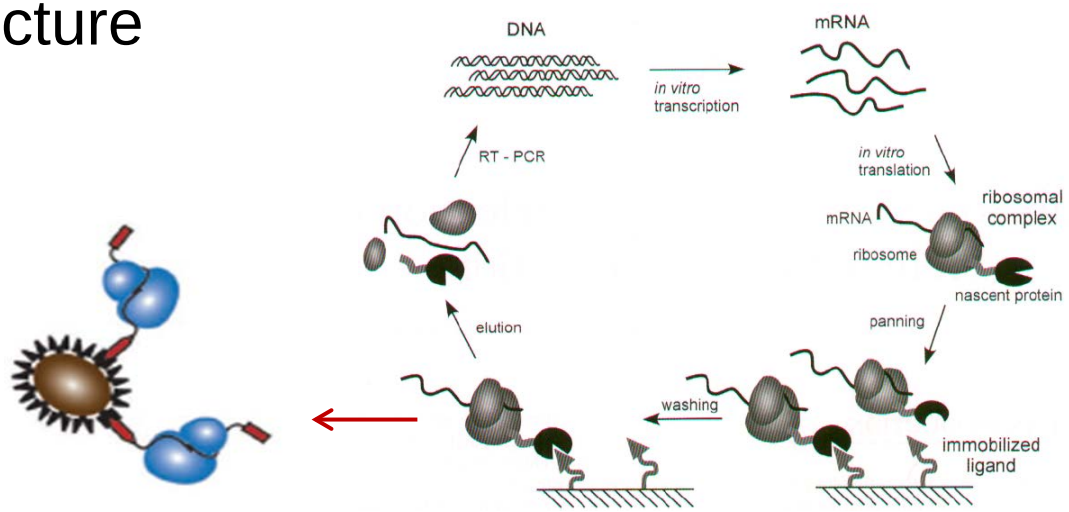
- Procedure:

>DNA of interest fused in frame with a spacer lacking stop codon

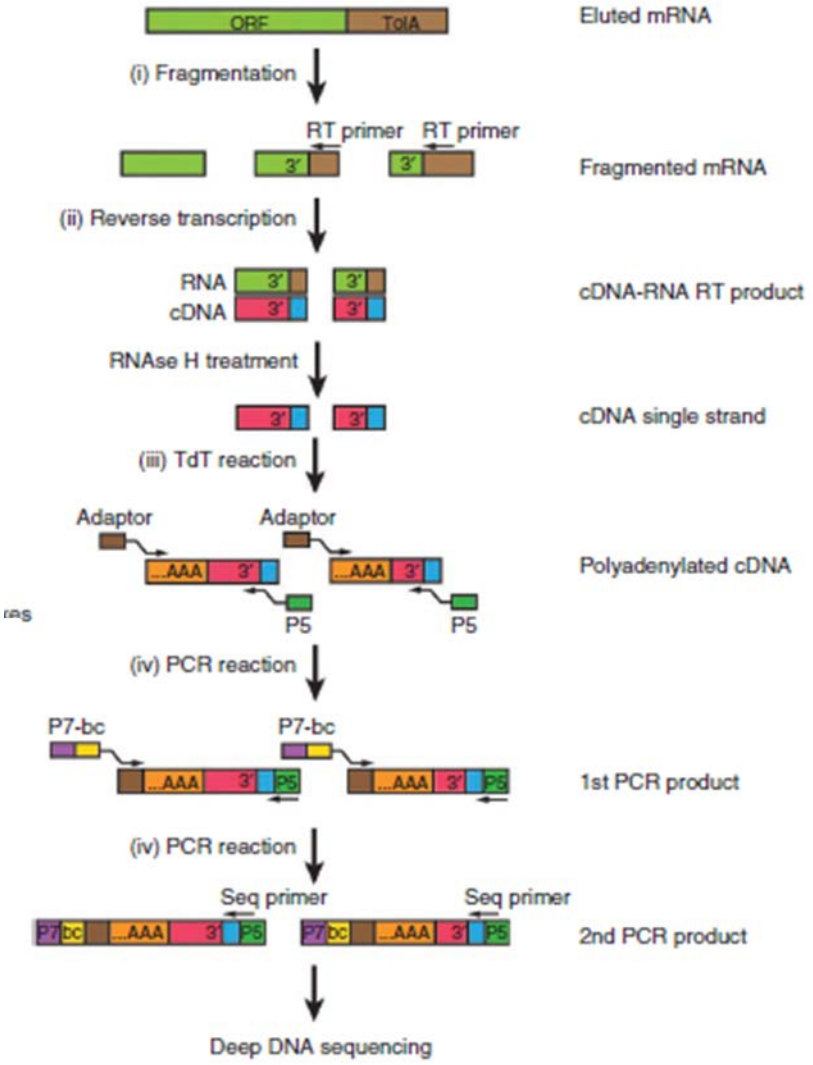
>Chaperones

>Stabilizing of the third structure

>Elution



Method



-Strategy for deep sequencing of enriched libraries

-Recovery of the 3'termini of the ORF:

>minimize RNA degradation
>allow stochiometric correlation between barcode counts and transcripts abundance

Deep sequencing-Illumina

- <u>3 steps</u>: 1) Library preparation

>Fragmentation of the DNA

>Ligation adaptors P5, P7

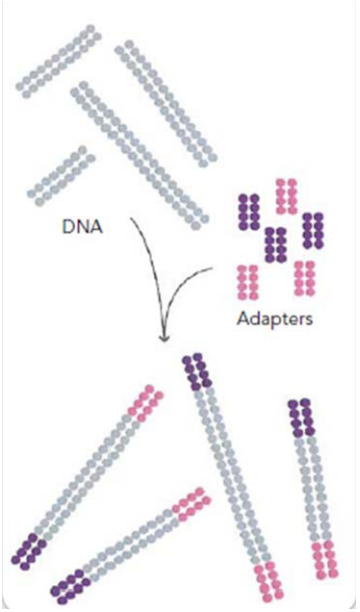
>DNA attach to oligont

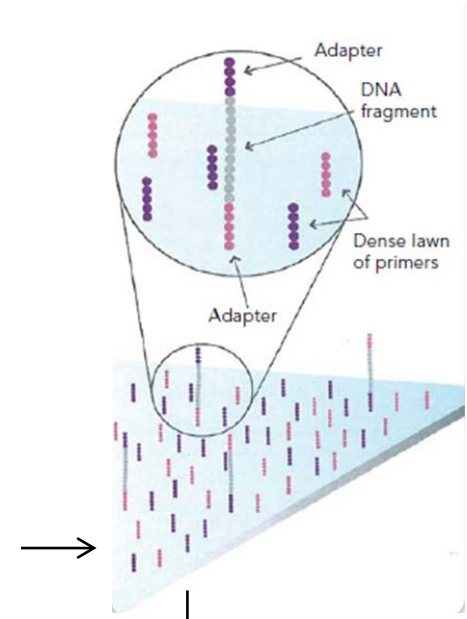
2) Bridge amplification

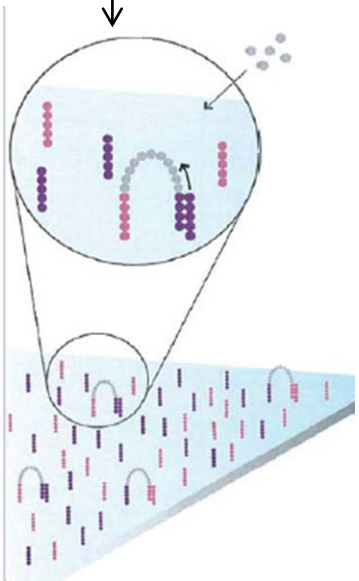
>PCR using P5, P7 primers

>Cluster of unique sequence

>Reverse strand cleaved and wash away







Deep sequencing-Illumina

3)Sequencing

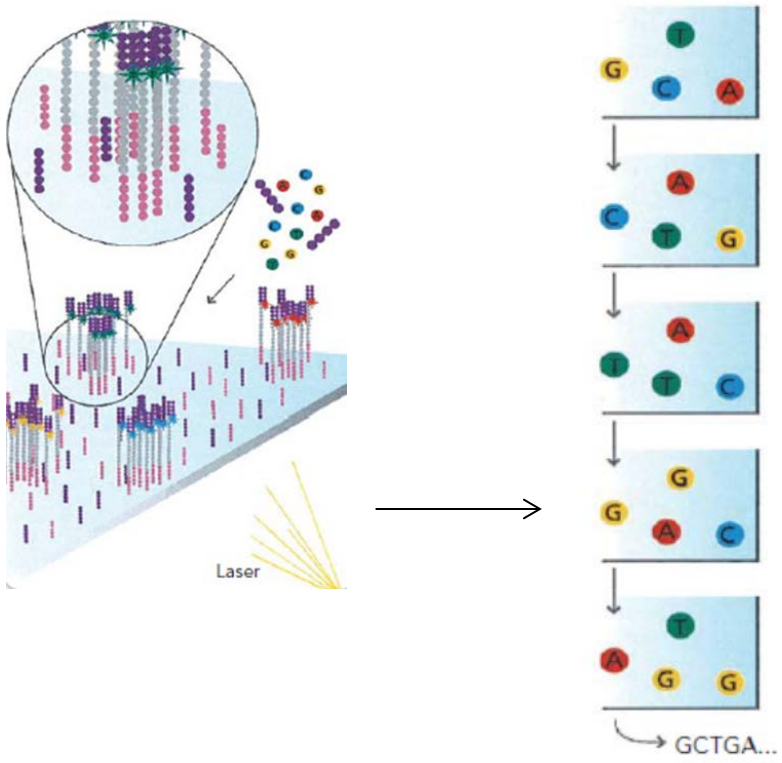
>Fluorescent bases

>Annealing

>Fluorescence detection

>Wash

Mapping of newly add nuclotide



- <u>Advantages</u>:

Accessibility

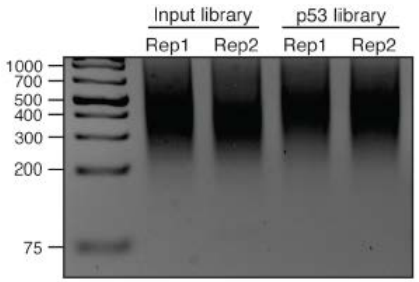
Accuracy

Ease of use

Problematic sequences

Deep sequencing

Preservation of mRNA from RT to deep sequencing



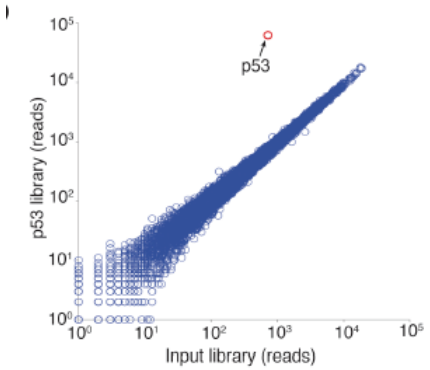
Analysis of unenriched library:

>Libraries spiked with p53 mRNA (100x)

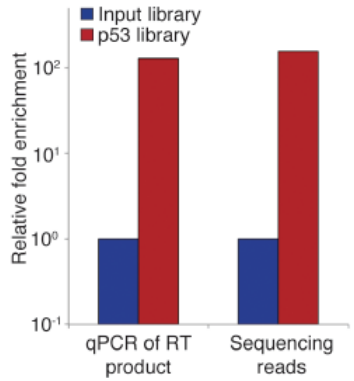
>Agarose gel of the second PCR

>Scatter plot of clone sequencing counts:

No p53 ORF are well correlated



>RT-qPCR of p53



Results:

-Reproducible

-Quantitative

Deep Sequencing

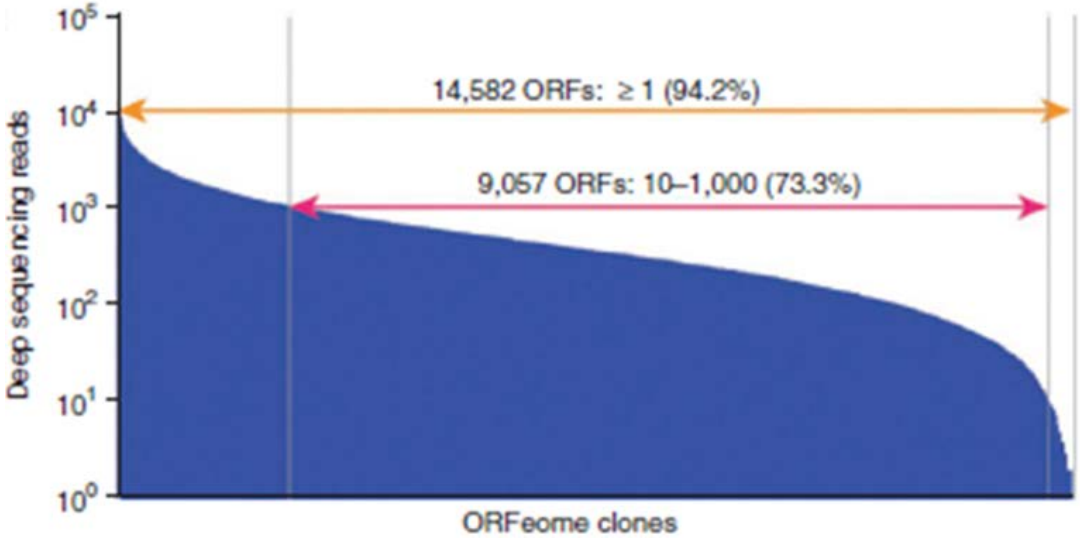
Multiplex: Pool of multiple samples
Identification by barcode
Fast and cheap

Single-end with custom primers (P5-attB2)

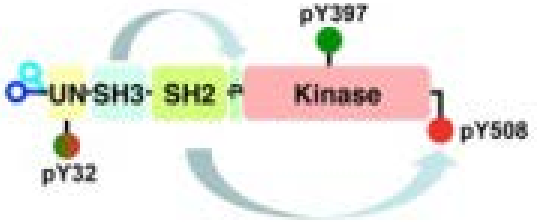
Sequencing reads of the unenriched human mRNA library

>Most ORF were sequenced at least once

>14’582/15’483 (=94%)



LYN kinase



SRC family: >SH3, SH2, kinase domain

>2 isoforms

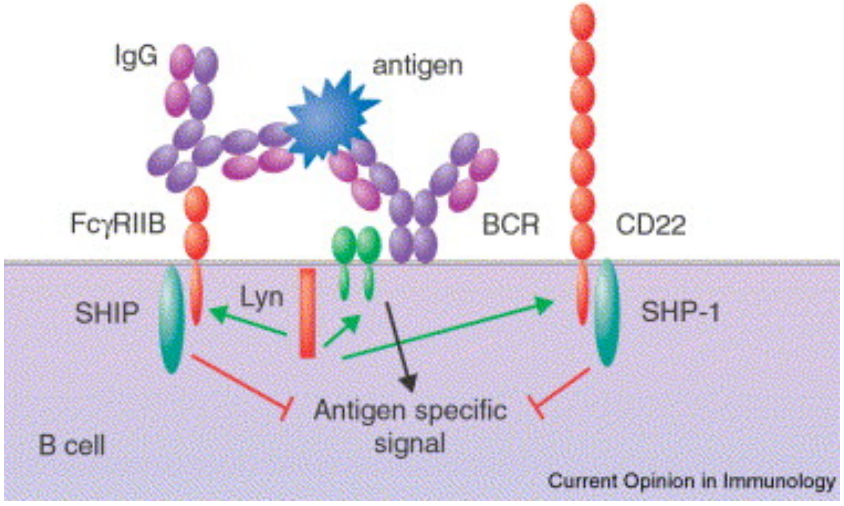
Expressed :myeloid, Bcells (hepatocytes, adipocytes)

Fonction: >Key signaling modulator of immune cell response

>Dual function

>Activating> P of ITAM: activation of PLCγ2 and PI3kinase

>Inhibition > P of ITIM: activation of phosphatase (SHIP-1, SHP-1)



LYN kinase

Pathology: Autoimmunity, Leukemia

Loss of inhibitory function,usually dominant

>$LYN^{-/-}$: splenomegaly, myeloproliferation

hyperactivity BCR > autoimmunity

> amplification loop

>Bafetinib: inhibitor of LYN, clinical trials

apoptosis in glioblastoma

>Dasatinib: BCR-ABL/LYN kinase inhibitor

>Need better understanding on the dual role of LYN kinase

*Interacting partners

*Inhibitors

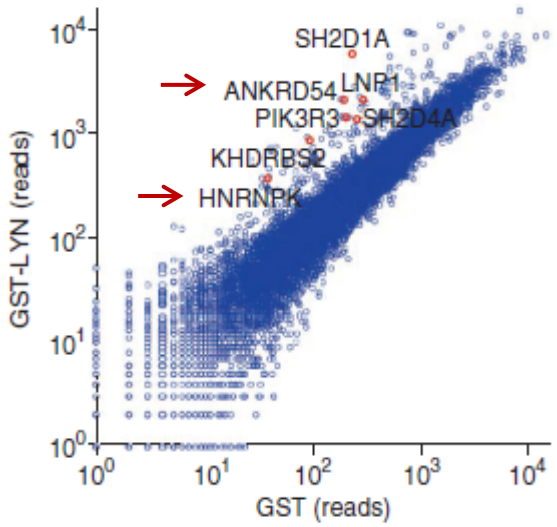
LYN kinase-Interacting partners

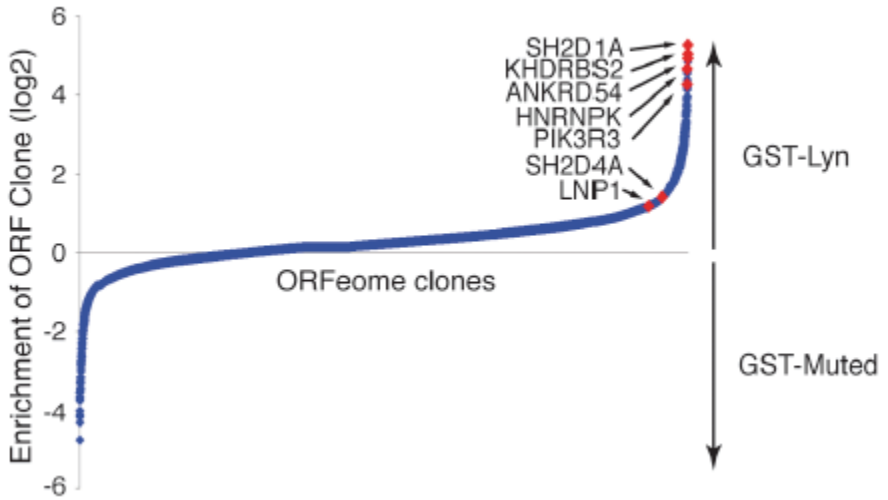
>Ability of PLATO to identify interacting partners of LYN kinase

>Affinity enrichment of ORFeome using GST-Lyn or GST/mutated (=bait)

>Illumina sequencing

>Results: number of known/new LYN kinase partners identified





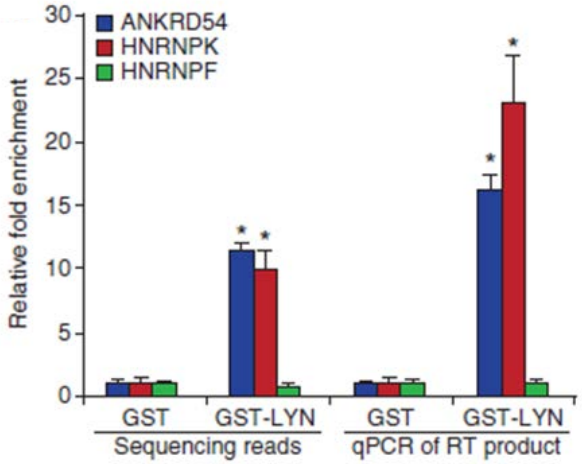
LYN kinase-Interacting partners

- Validation

>qPCR:

-ORF specific primers

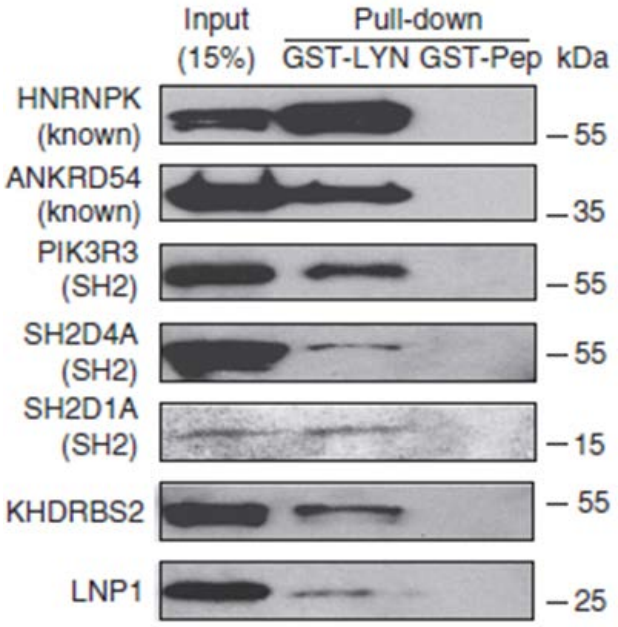
-2 known LYN binding partners



>Western Blot

-v5-His tagged candidate protein

-anti-V5 antibody detection



LYN kinase-Interacting partners

- Ranking of the LYN interacting candidate
- Enrichment of SH2-domain containing proteins

(p-value: 0.0.098)

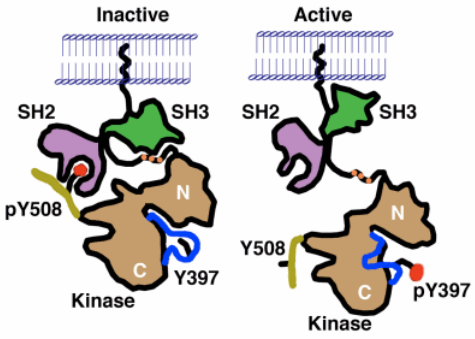
>Role of LYN autophosphorylation in interacting with SH2 domains

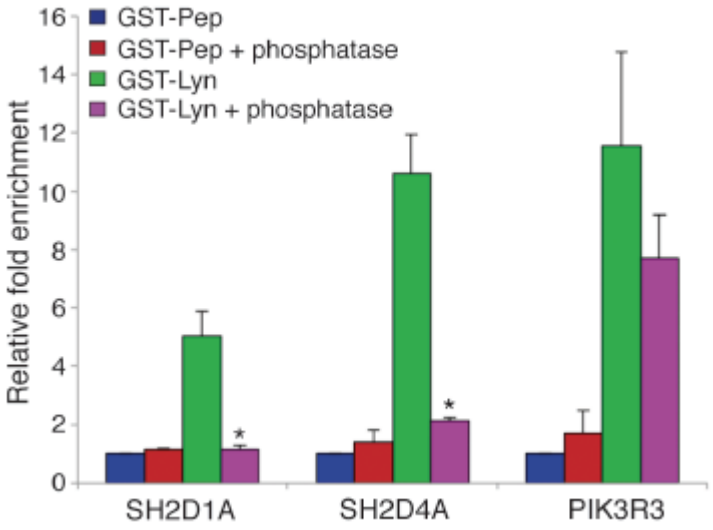
>Phosphatase treatment of immobilized GST-Lyn

>Abolish the binding of:

SH2D1A and SH2D4

>Evidence for an additional binding domain for PIK3R3





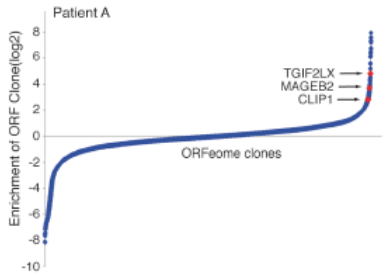
LYN kinase-Autoantibodies

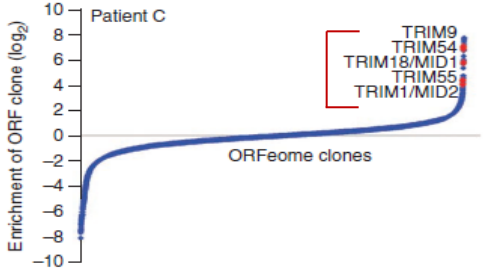
Ability of PLATO to identify protein target of antibodies from patients with autoimmune disease

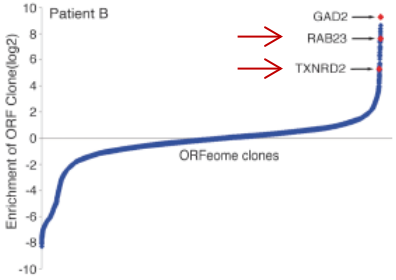
>cerebrospinal fluid of patients with paraneoplastic neurological disorder

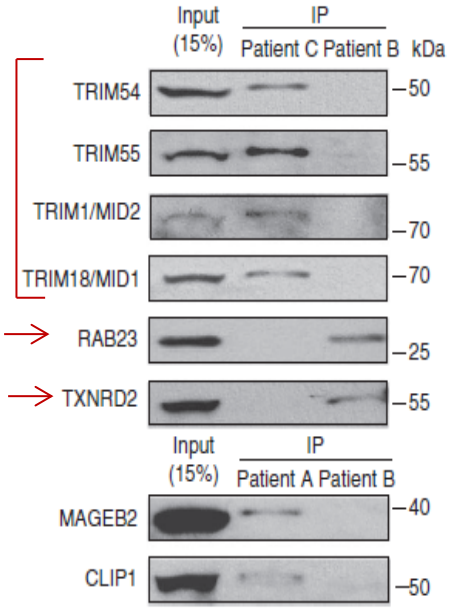
>Results: -Detection/confirmation of reactive antigens by IP and WB

-Heat map: Patients specificity









PND / ORF	A	B	C
TGIF2LX	26.3	6.0	1.1
MAGEB2	13.2	0.4	0.9
CLIP1	6.8	0.1	1.3
GAD2	40.3	613.3	7.0
RAB23	17.9	170.6	56.3
TXNRD2	1.3	38.1	2.8
TRIM9	3.0	2.8	126.0
TRIM54	0.2	3.4	112.3
TRIM55	1.3	0.9	20.7
TRIM1/MID2	1.1	1.5	20.4
TRIM18/MID1	25.5	0.3	53.2

Enrichment

>100 | 75 to 100 | 50 to 75 | 20 to 50 | 10 to 20 | 5 to 10 | 0 to 5

LYN kinase-small molecule

High limitation of small molecule interaction with protein:

>cell extract (abundancy of protein)

>biaised of MS analysis towards the highly produced proteins

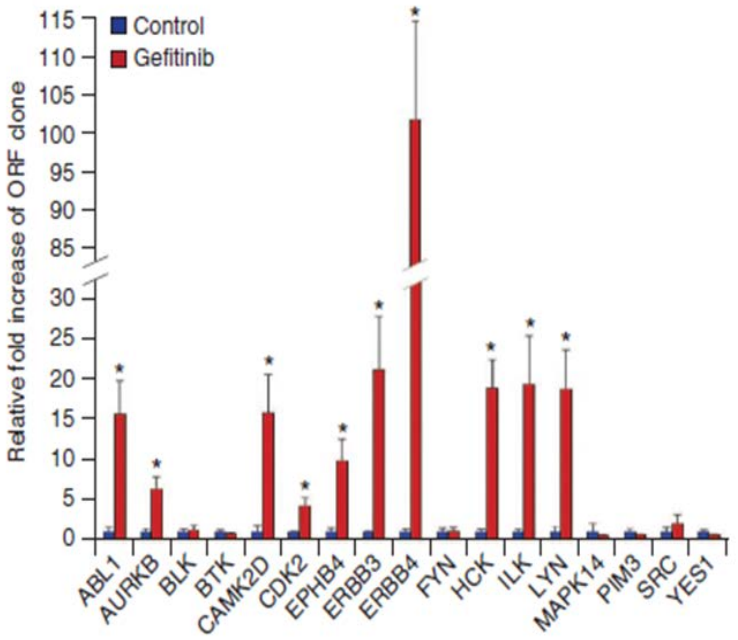
Gefitinib:

>Inhibitor of EGFR tyrosine kinase domain

>Used in breast, lung cancer

>Interact with the ATP binding pocket EGFR

Results: 10/17 predicted targets



Summary

- Advantages: -Minimal constraints of ribosome display on

>Lengths and composition of protein

-Multiplexed sequencing reduce cost

-Compatible with large number of samples and with automation

-Broad utility

- Limitations: -Incomplete ORFeome collection

>constant improvement: quality, completeness, availability

-Lack of protein post-translational modifications

-Low display efficiency of large proteins ORF

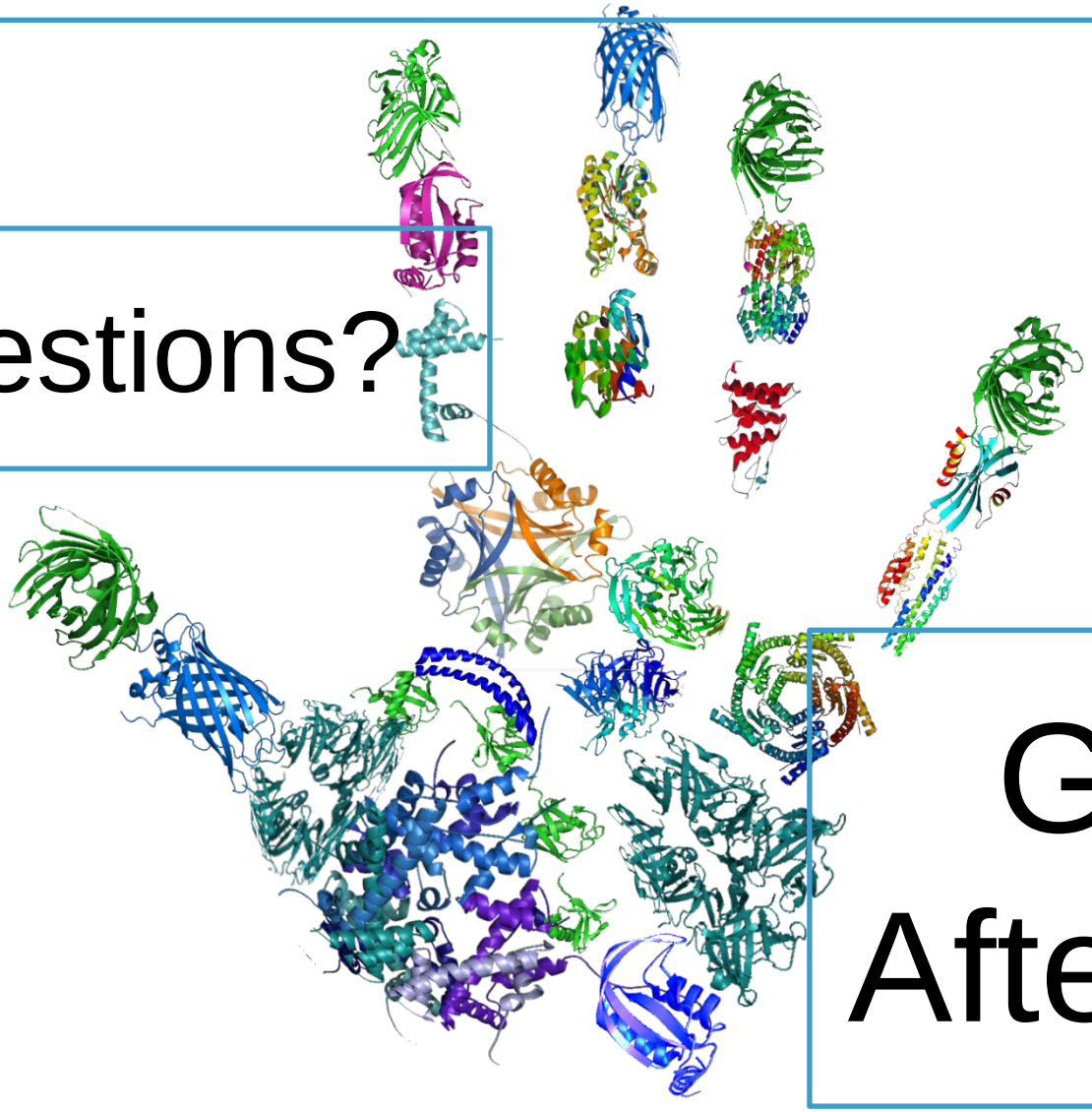
-Aggregation of proteins requiring host cellular machinery for proper folding

-Affinity purification sensitive

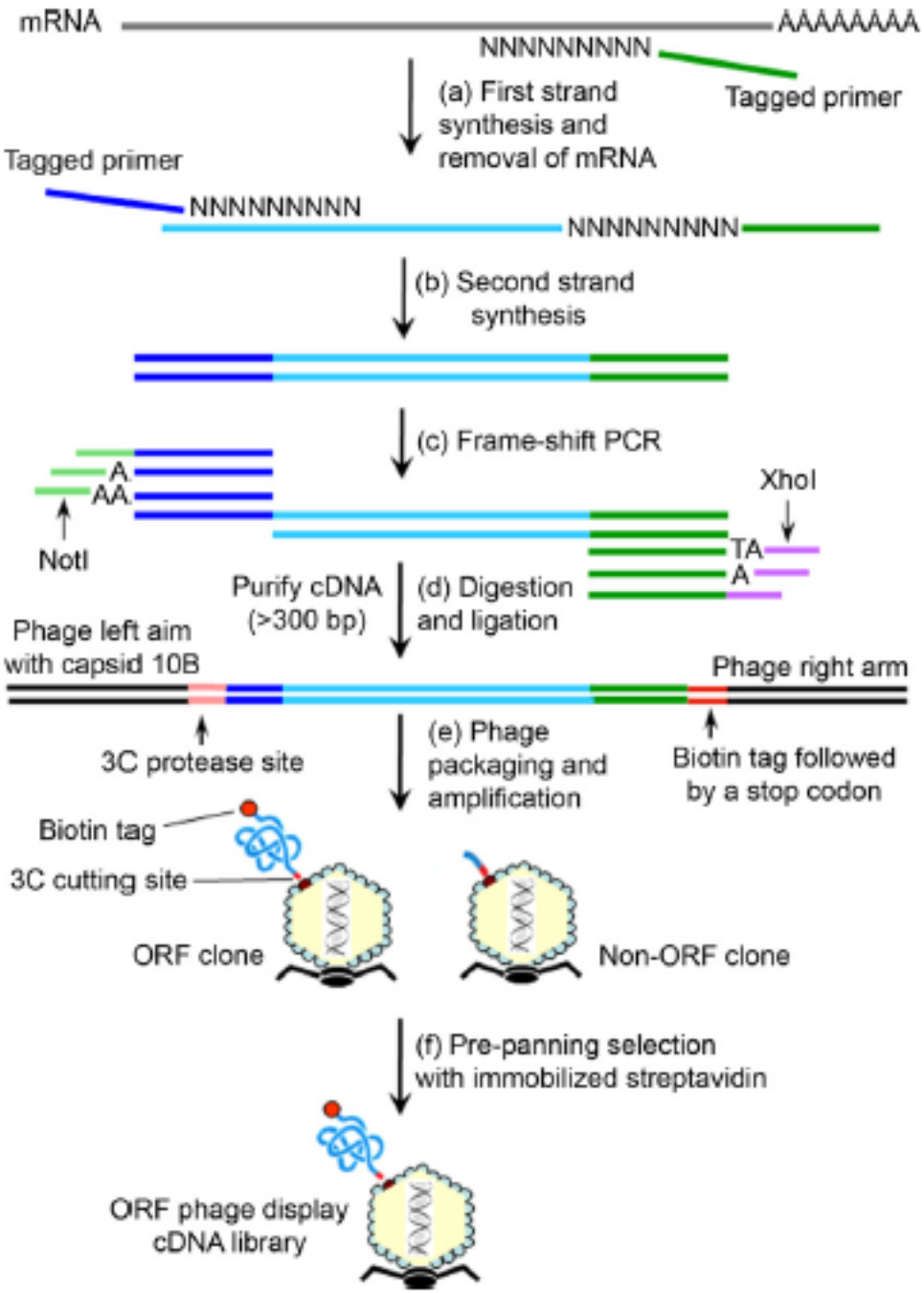
-non specific binding of protein containing nucleic acid domain as bait

Thanks for your attention

Questions?



Good Afternoon


mRNA
AAAAAAAA
NNNNNNNNN
Tagged primer
(a) First strand synthesis and removal of mRNA
Tagged primer
NNNNNNNNN
NNNNNNNNN
(b) Second strand synthesis
(c) Frame-shift PCR
A
AA
NotI
XhoI
TA
A
Purify cDNA (>300 bp)
(d) Digestion and ligation
Phage left aim with capsid 10B
Phage right arm
3C protease site
(e) Phage packaging and amplification
Biotin tag followed by a stop codon
Biotin tag
3C cutting site
ORF clone
Non-ORF clone
(f) Pre-panning selection with immobilized streptavidin
ORF phage display cDNA library

Deep sequencing-Illumina