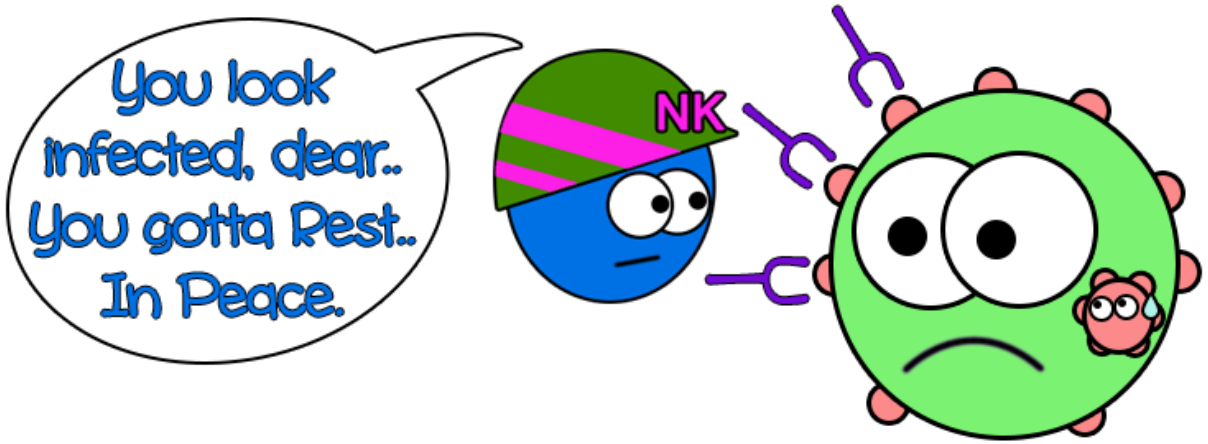
High throughput screening of phage display libraries for production of fully human antibodies challanged to cells expressing native Claudin-1

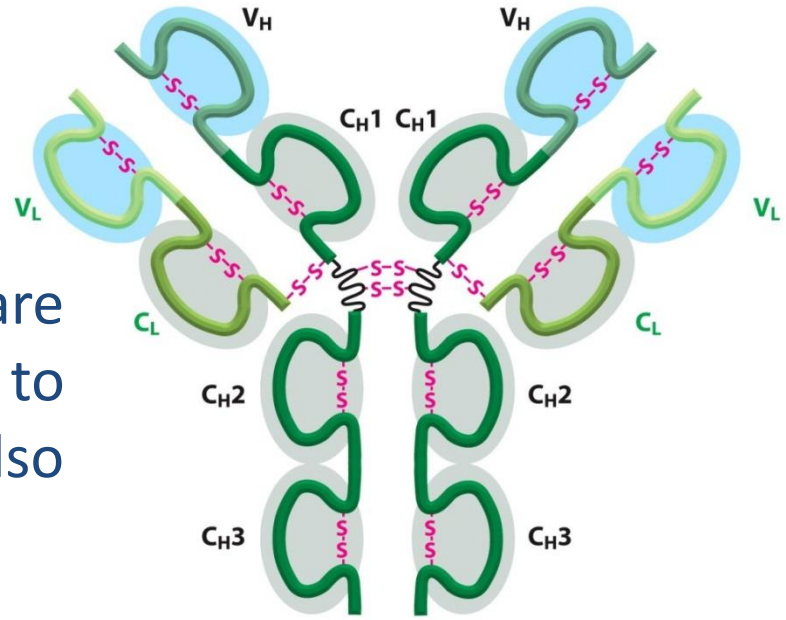
Emanuele Sasso

Genomics and Bioinformatics Workshop 11.2015



The Structure of Antibodies

Immunoglobulins of class G (IgG) are constituted by two heavy chains H equal to each other, and two light chains L also equal to each other.

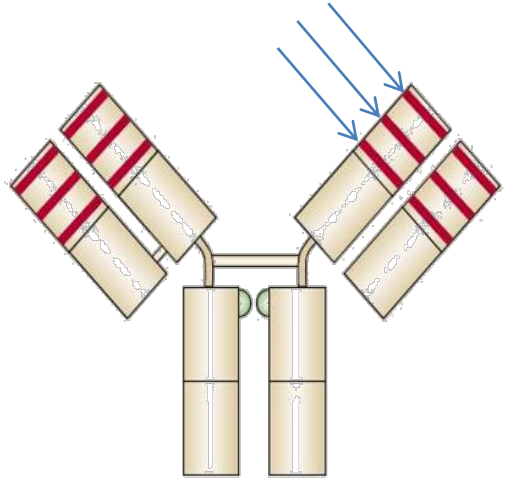


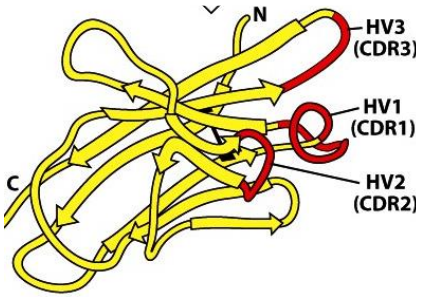
The L and H chains are composed of variable domains (VL and VH) and constant domains, one for the L chain (CL), and three for the H chain (CH1 CH2 CH3)

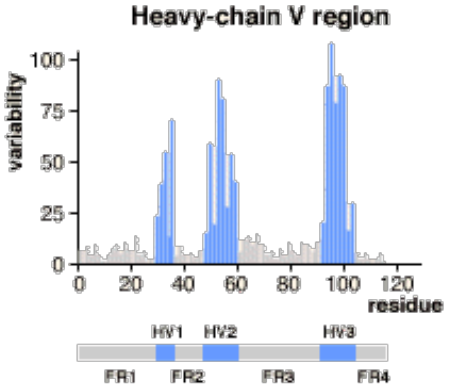
The protein domains responsible for antigen recognition are in the variable regions VH and VL

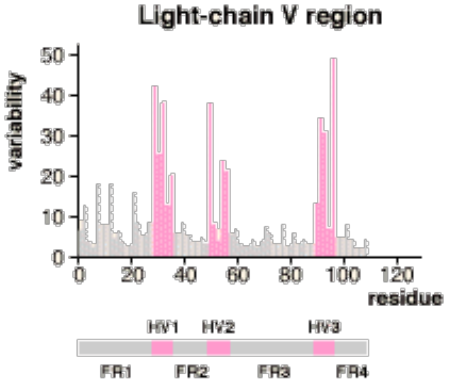
The Structure of Antibodies

Hypervariable domains
(complementarity determining regions CDR)



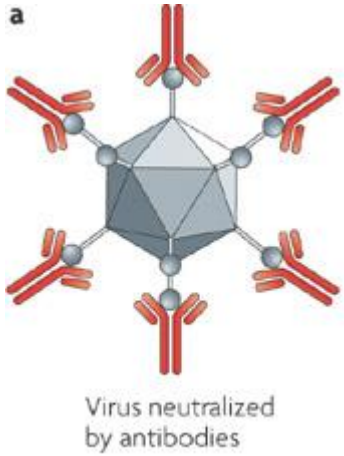


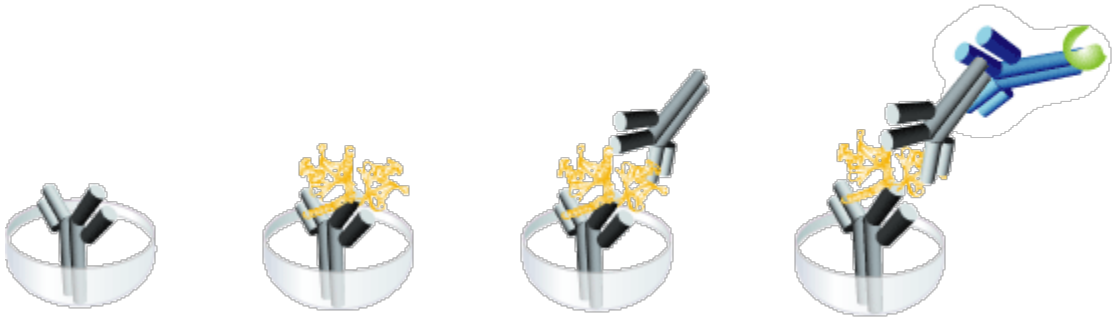


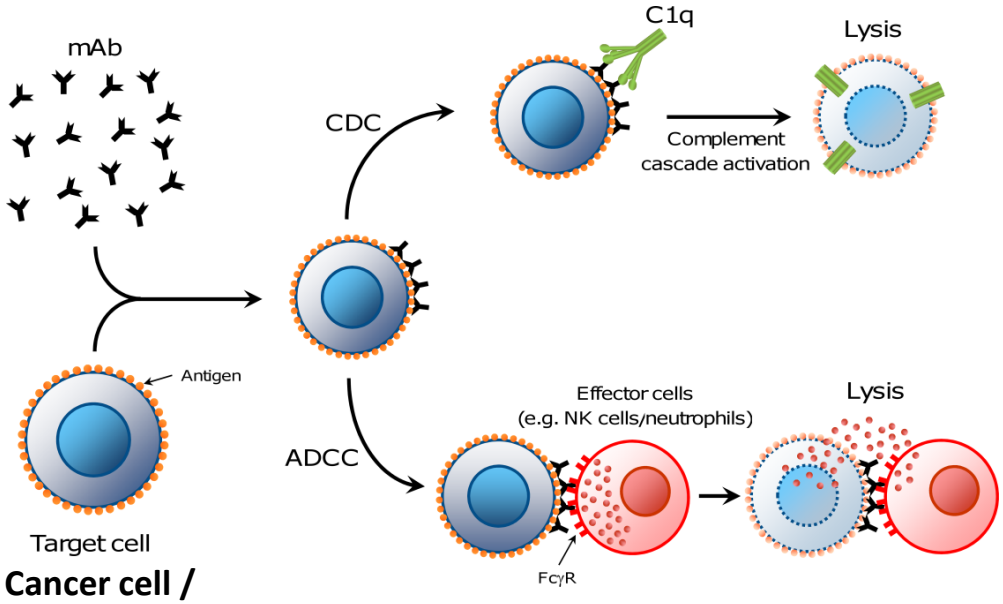


The variable region of an antibody is responsible of its binding to an epitope. The variable region is further subdivided into hypervariable (HV) and framework (FR) regions. Within light and heavy chains, exist three hypervariable regions and four FR regions. The hypervariable regions are definitely involved in epitope binding.

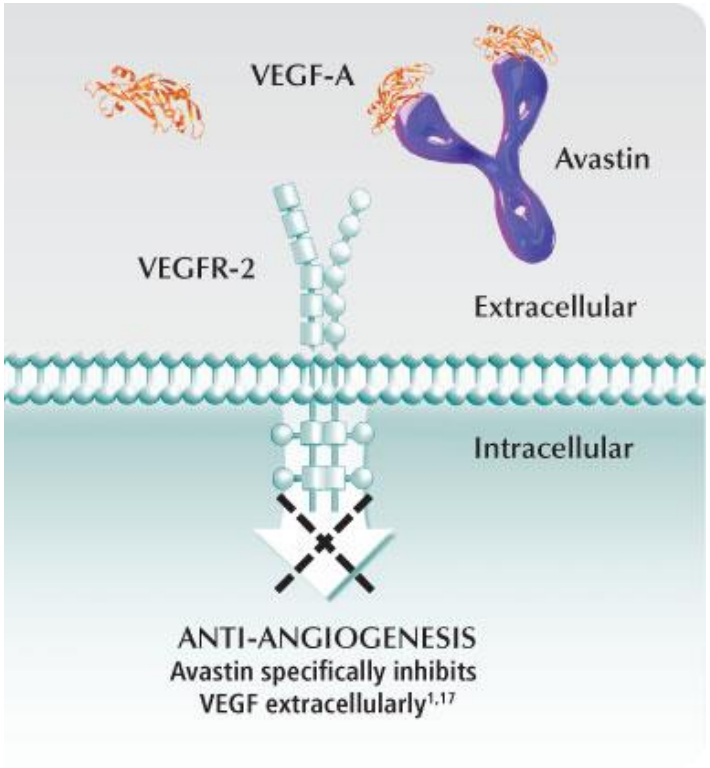
Application of recombinant monoclonal antibodies



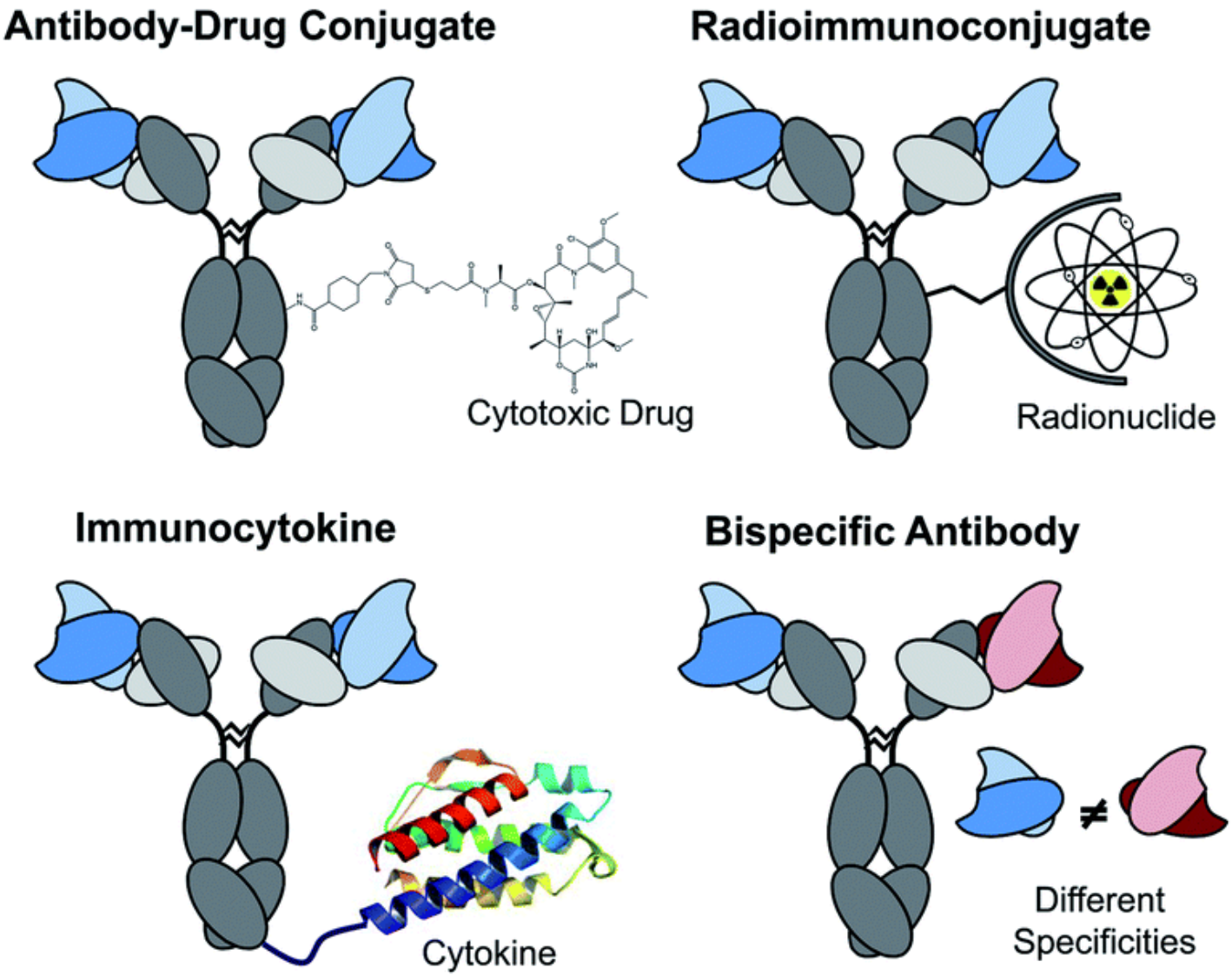




Cancer cell / infected cell



Advanced applications



Murine antibodies by hybridoma technique

1) Immunize animal (mouse or rabbit)

2) Isolate spleen cells (containing antibody-producing B cells)

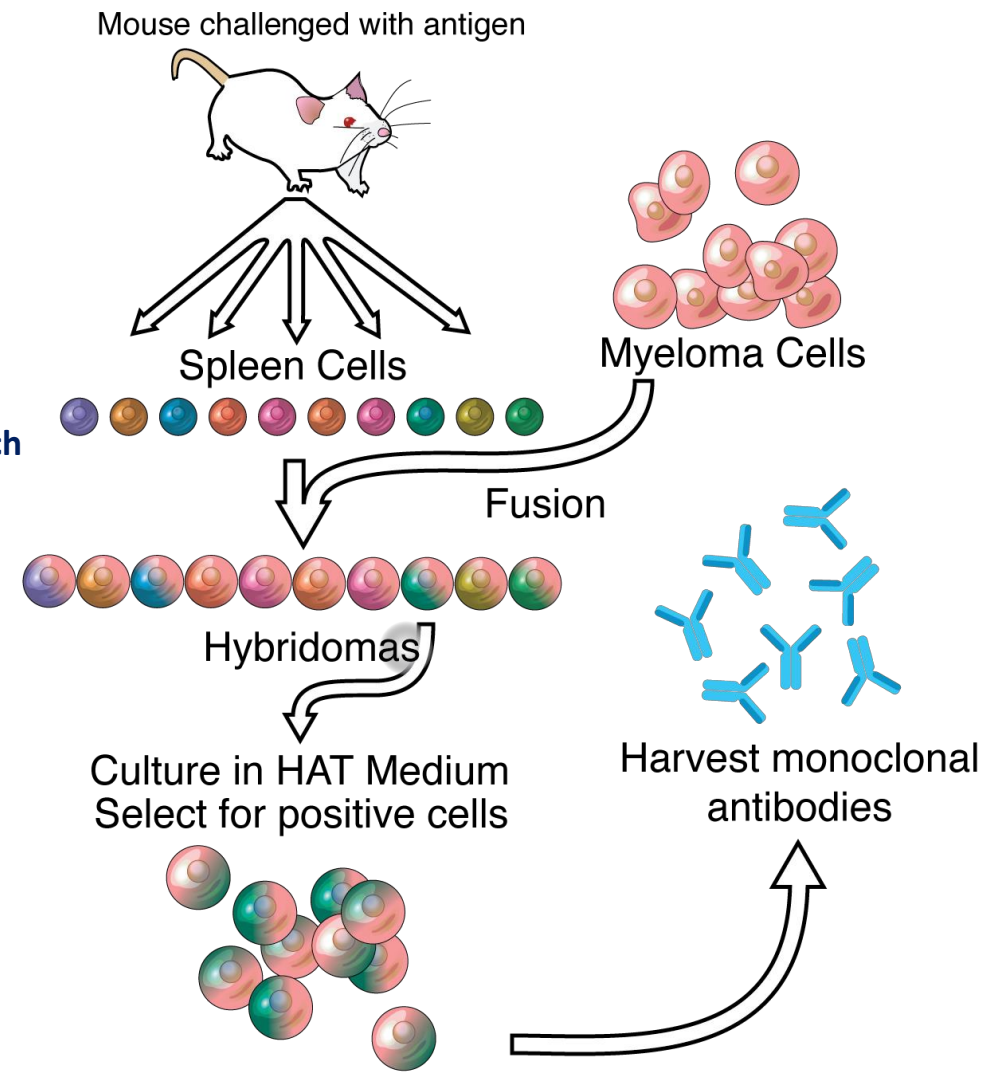
3) Fuse spleen cells with myeloma cells (e.g. using PEG - polyethylene glycol)

4) Allow unfused B cells to die

5) Add HAT culture to kill unfused myeloma cells

6) Clone remaining cells (place 1 cell per well and allow each cell to grow into a clone of cells)

7) Screen supernatant of each clone for presence of the desired antibody (ELISA)

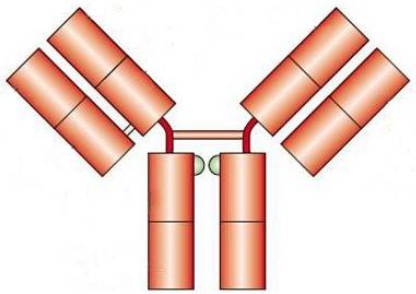


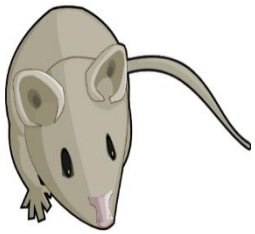
Problems with using mouse mAb

•The therapeutic use of rodent monoclonal antibodies in humans is limited by their immunogenic, short circulating half-life, and inability to efficiently trigger human effectors mechanisms:

- This is due to differences between the mouse and humans.
- Also severe allergic response in human when mouse mAb are introduced to a patients.
- Also constant region of marine mAb are not effective in interacting with human effectors molecules.

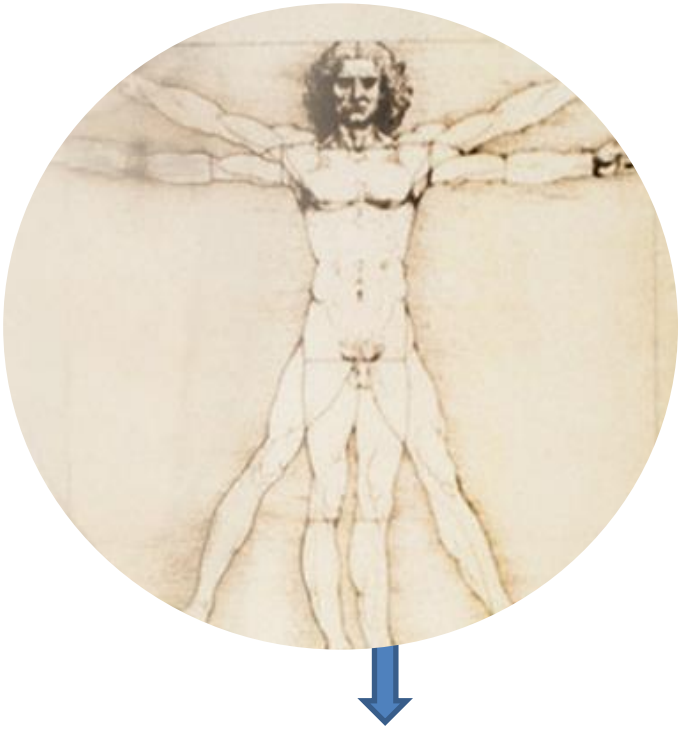
Induction of human anti-murine mAb : HAMA (Human Anti-Mouse Antibodies)





Injection for therapeutic purposes in human

different species

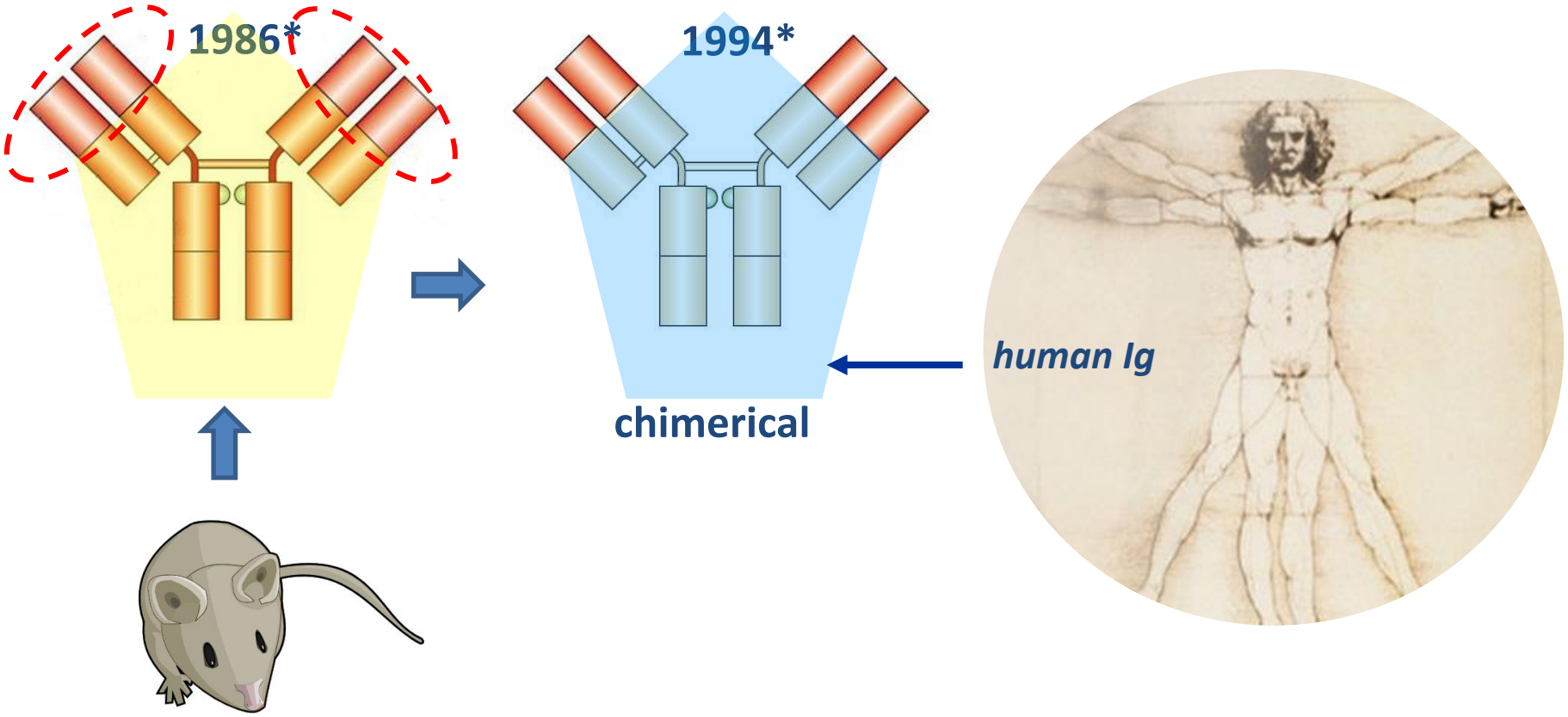


- Murine antibody is recognized by the human immune system as' foreign'
- We produce human antibodies against murine mAb that is inactivated
- Therapy becomes ineffective

Generation of chimeric monoclonal antibodies

The solution:

- **Replace the constant part (CH) of the mouse antibody with that of a human one.**

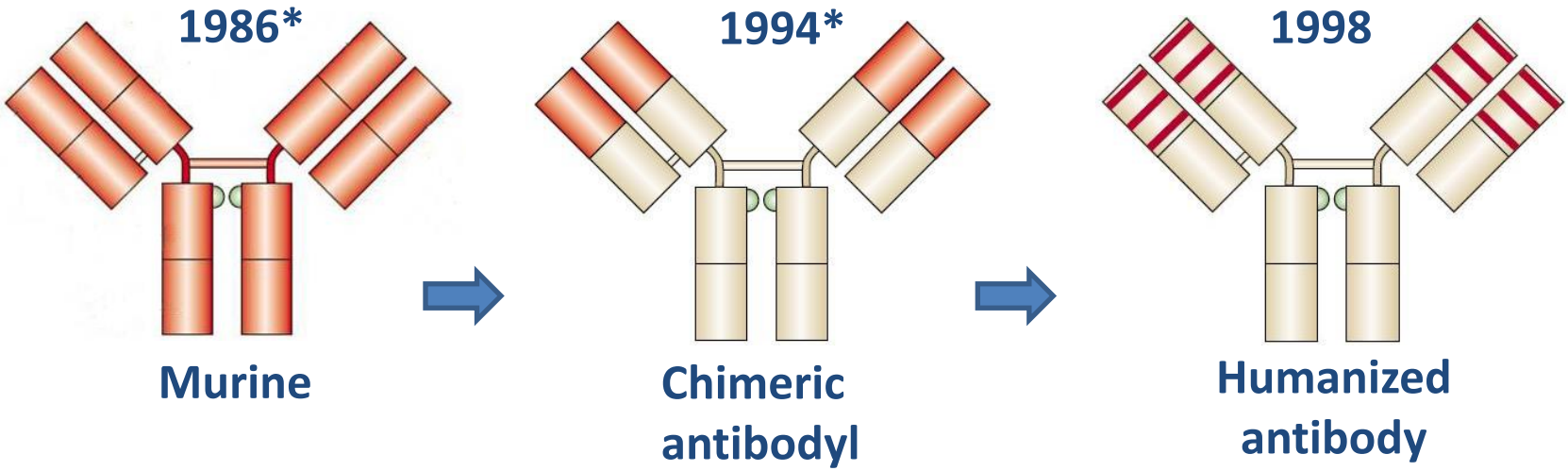


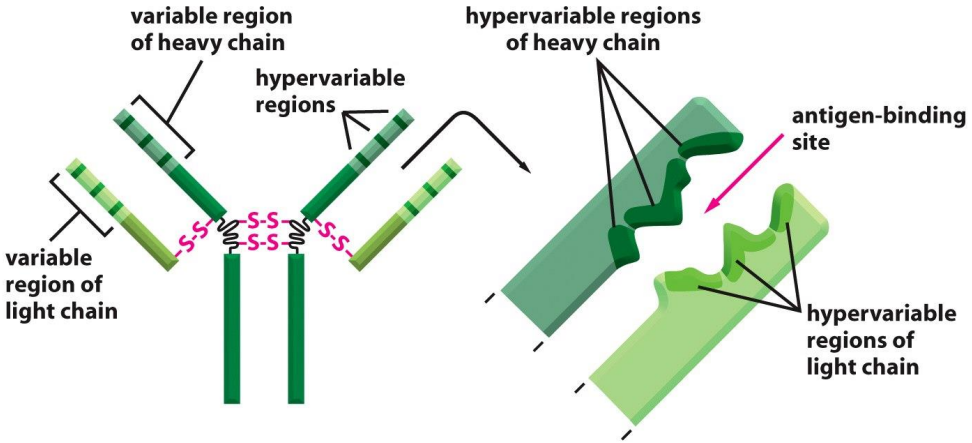
From chimeric to humanized antibody

The solution:

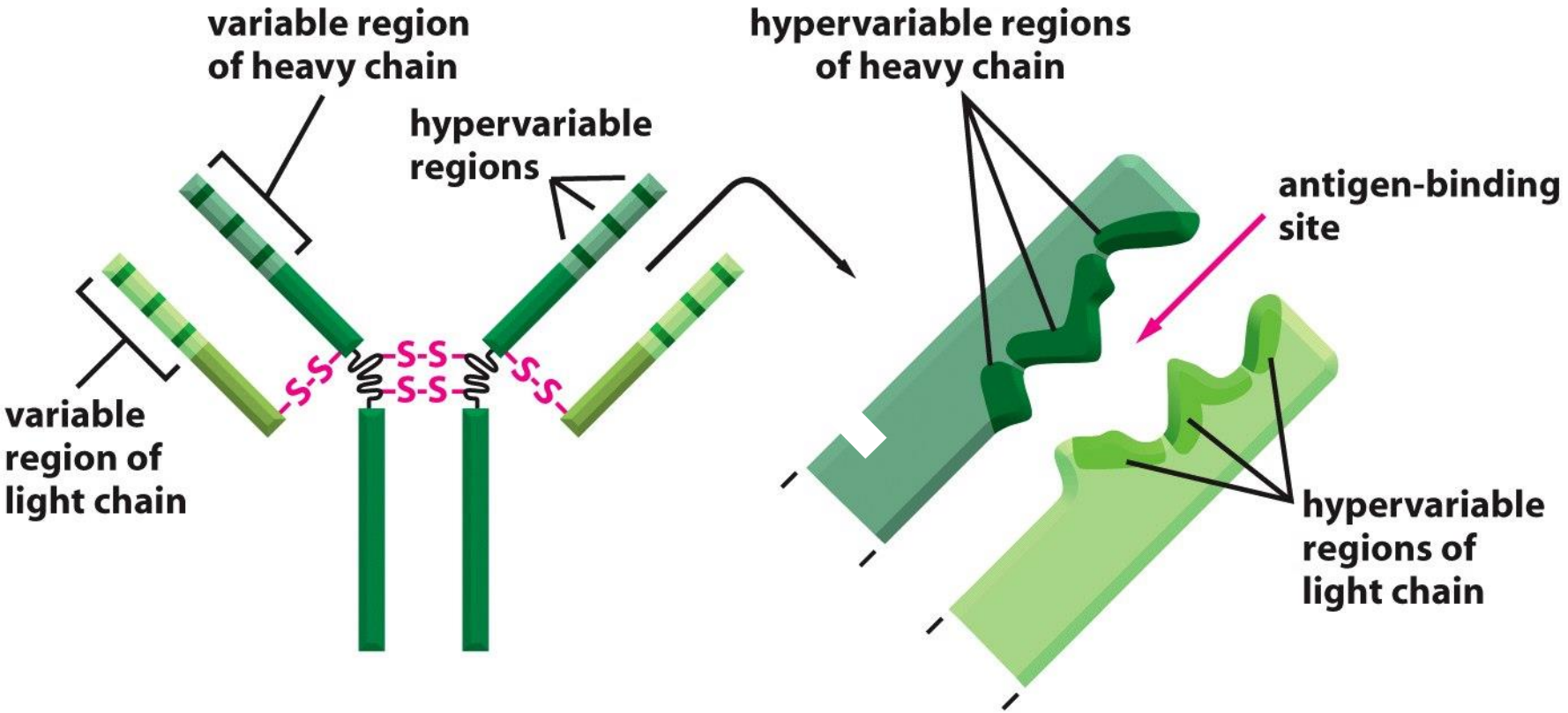
- **'Transplant' only the hypervariable murine antibody in a human antibody**

  = Humanized antibody



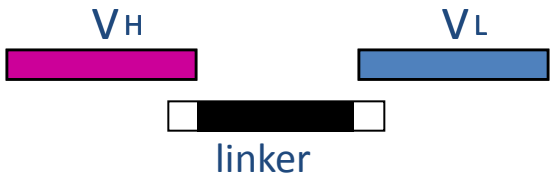


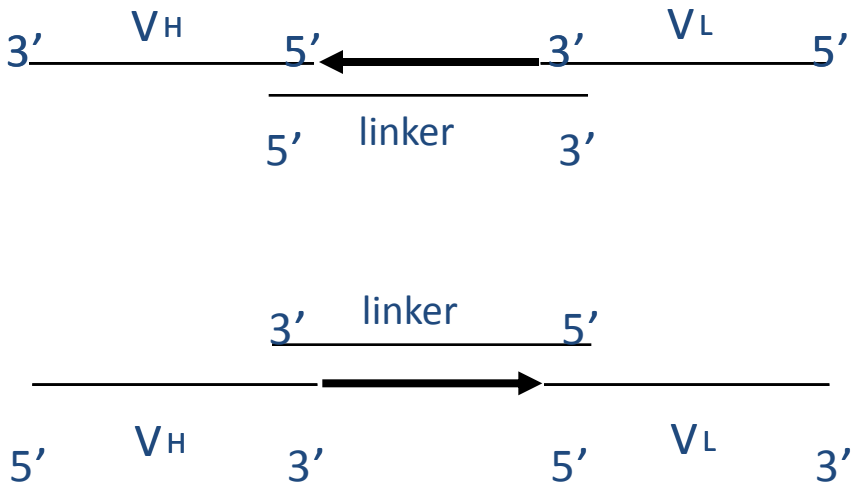
The only variable regions are needed for epitope binding

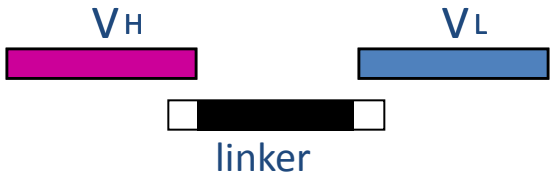


Random combination of V_H-V_L by flexible linker

- VH and VL are amplified and combined through a poli-peptidic linker(Gly 4, Ser 3)

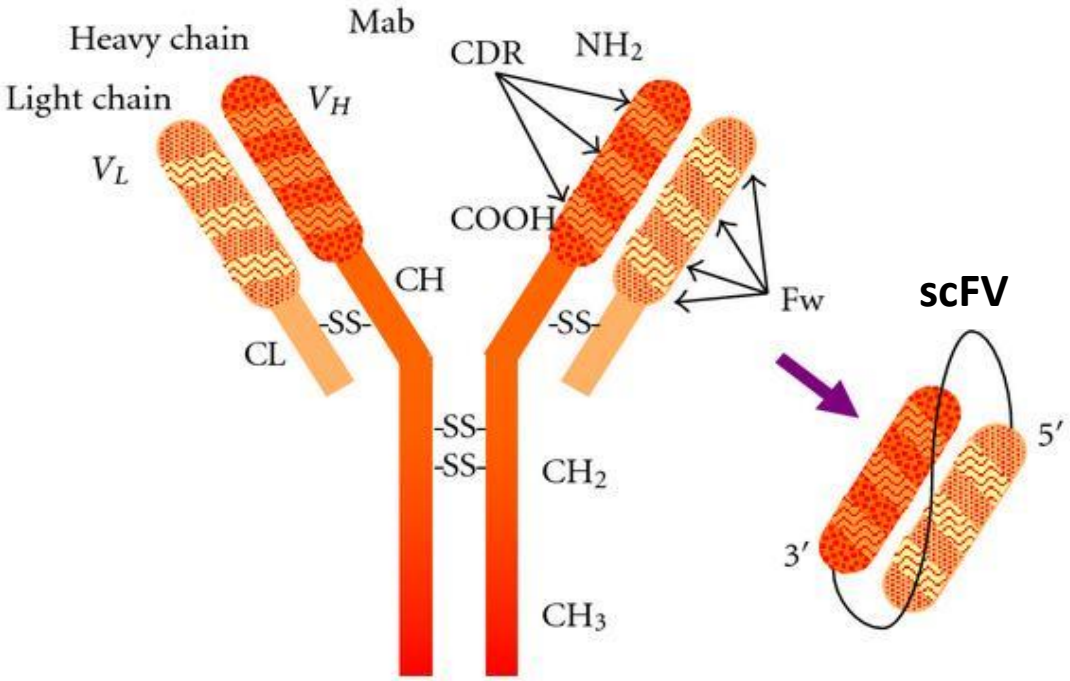




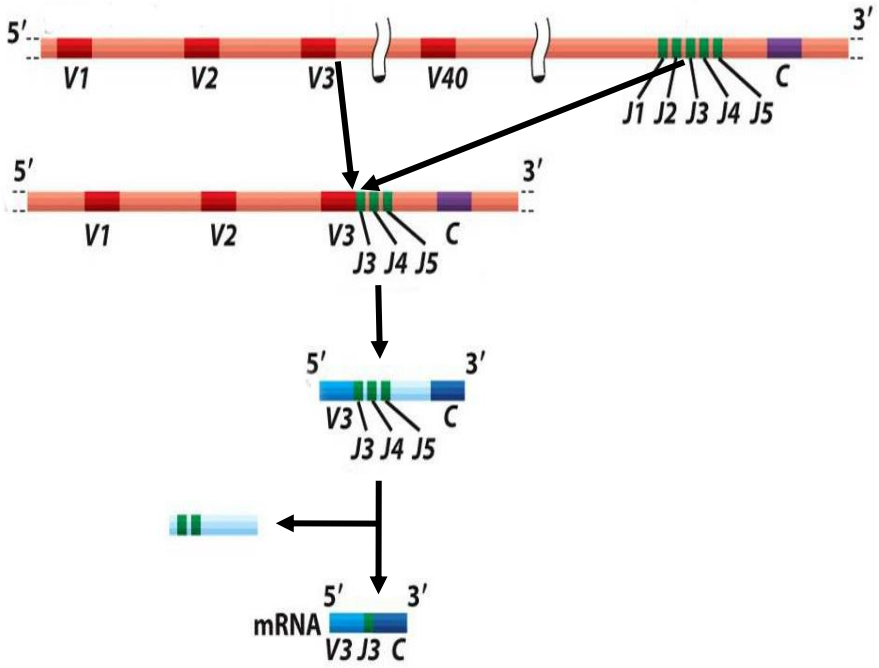

VH
VL
linker


3’
VH
5’
3’
VL
5’


5’
VH
3’
5’
VL
3’


Heavy chain
Mab
CDR
NH2
Light chain
VH
VL
COOH
CH
-SS-
CL
-SS-
Fw
scFV
-SS-
-SS-
CH2
CH3
5′
3′

In vivo vs in vitro diversity



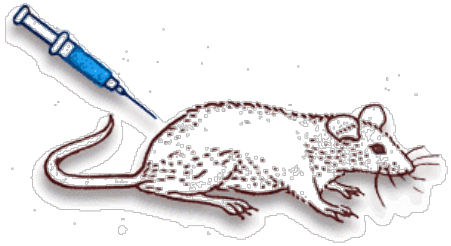
10^8 in vivo immune sistem different antibodies

10^9-10^{11} ScFv

Where from to isolate VH and VL?

Immunised library:

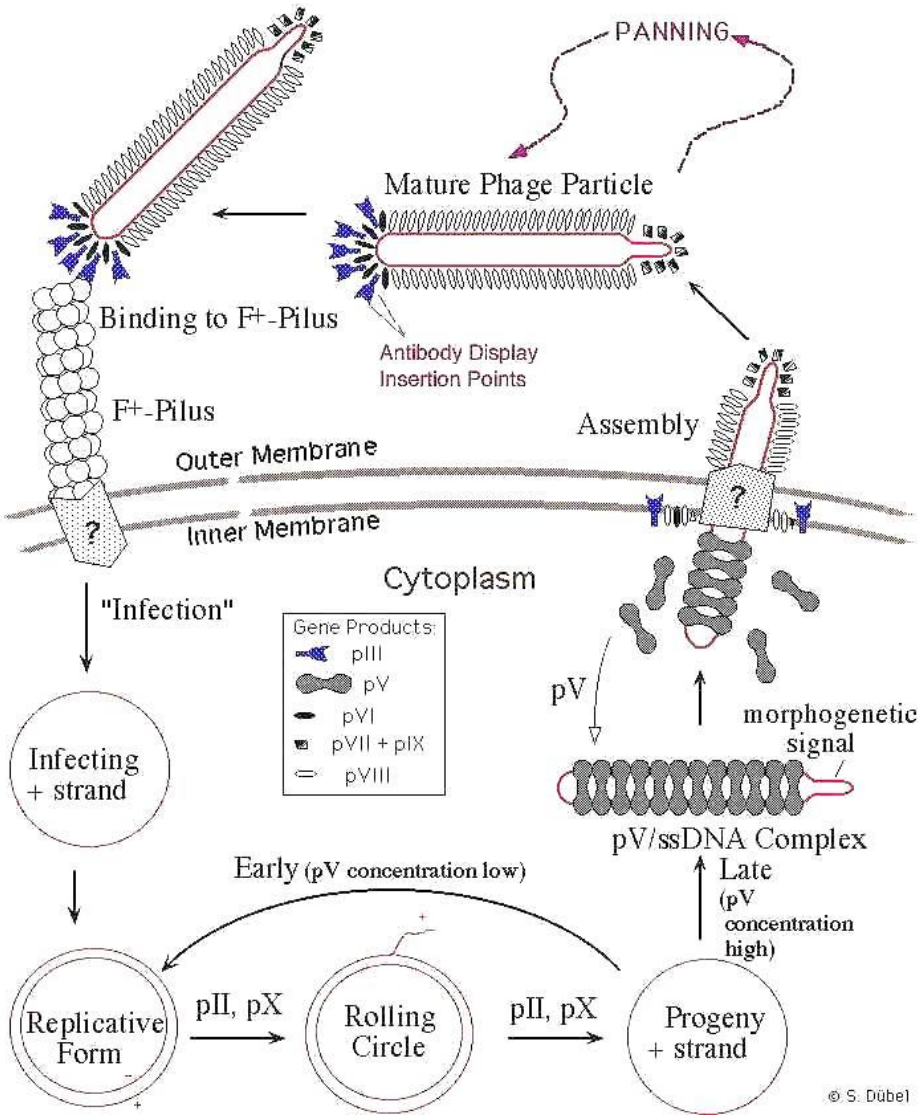
- VH and VL are isolated from B cells of an immunised animal or humans.
- Suitable for difficult antigens



Naive library/germline library:

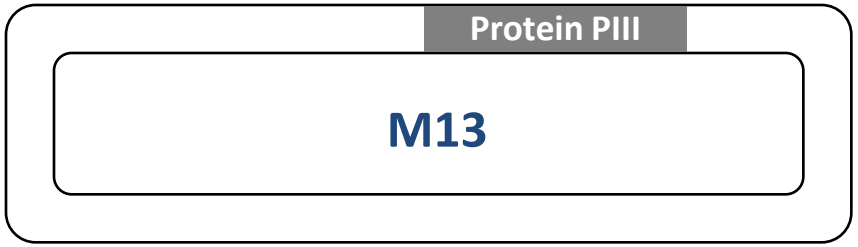
- VH and VL are isolated from healty **human** donors and randomly combined.
- Suitable for almost all antigens

Life cycle of filamentous phage

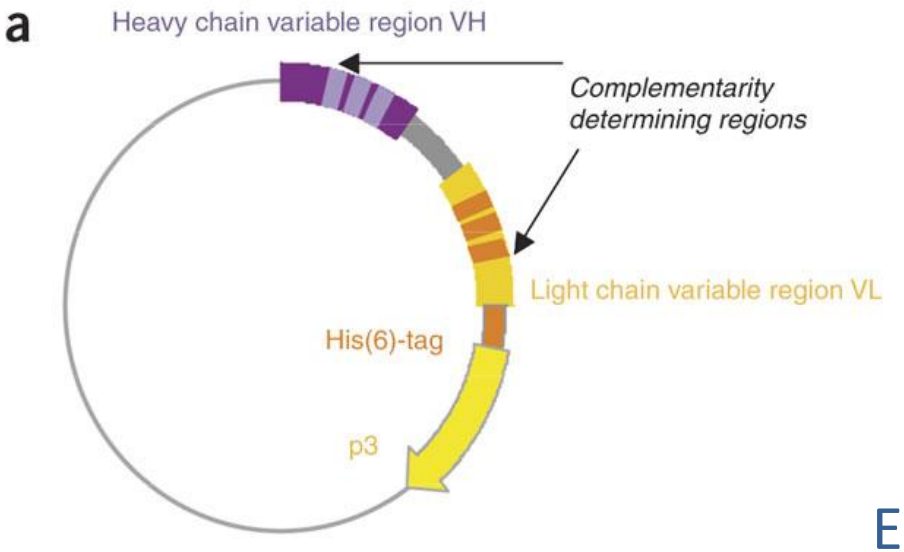


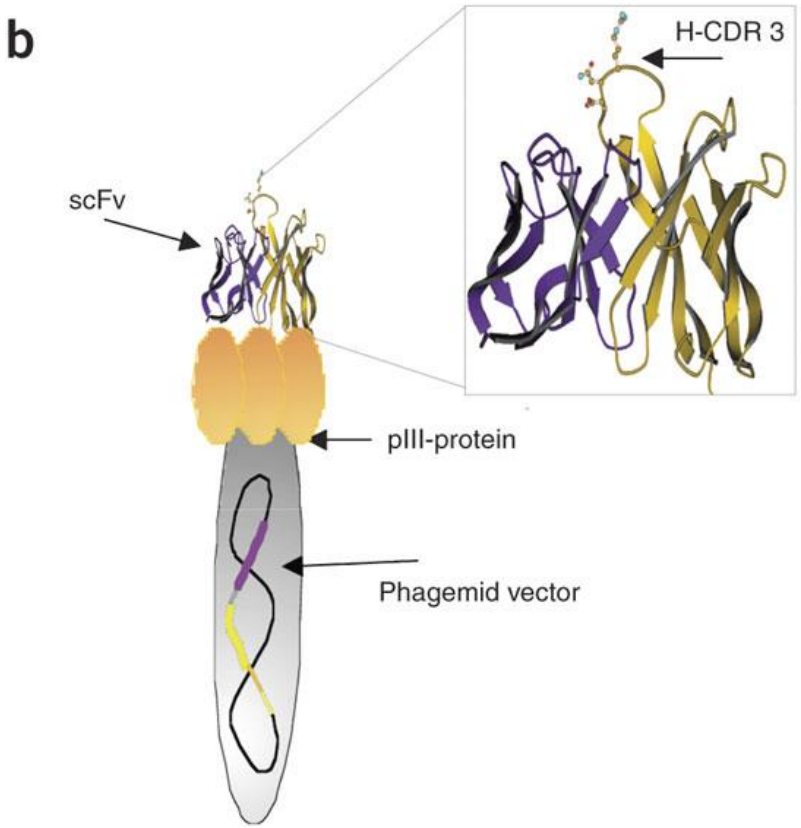
How to screen the scfvs?

Phage Display



Scfv are “exposed” on M13’s capsid in frame fused to PIII protein

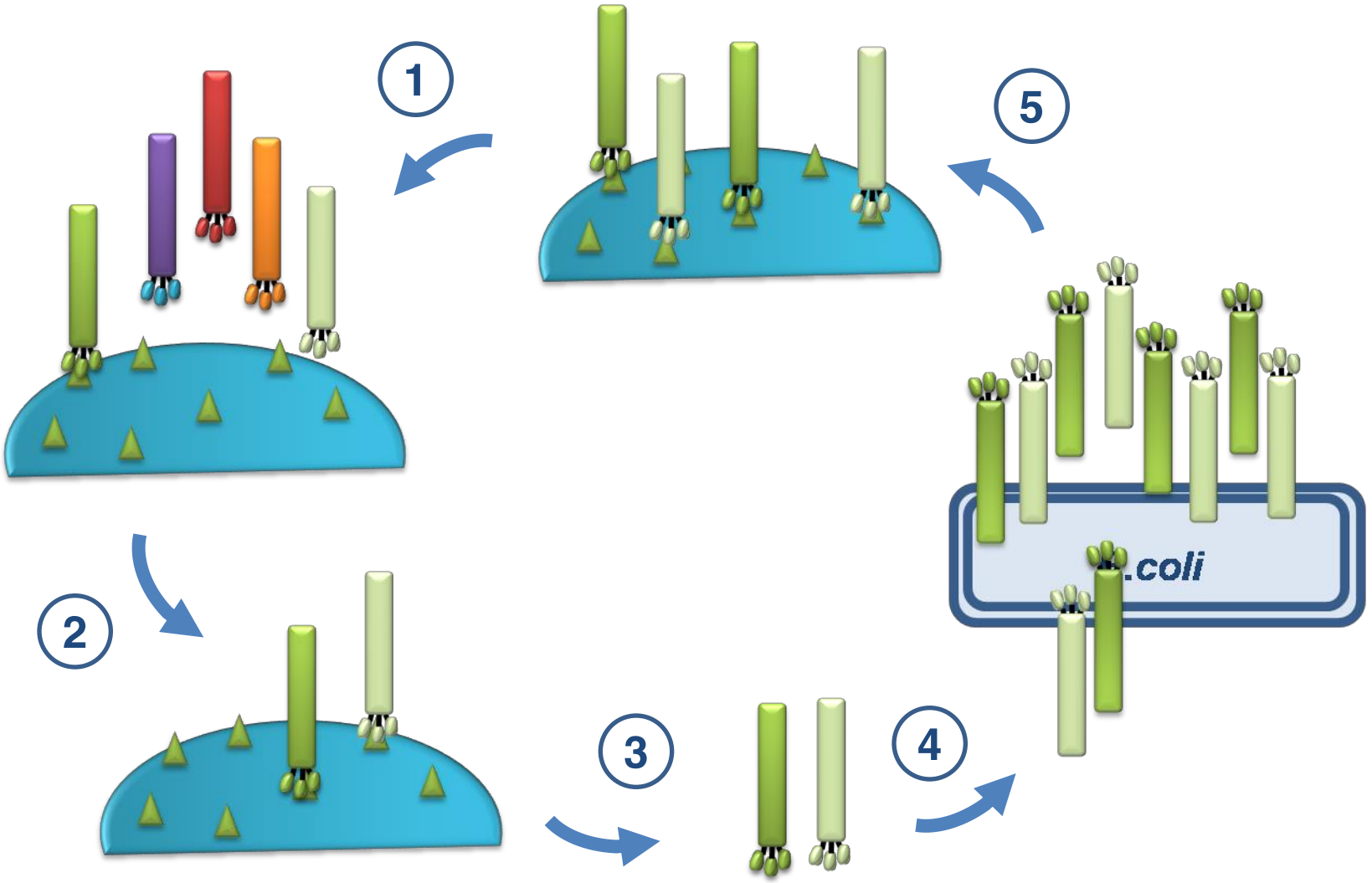


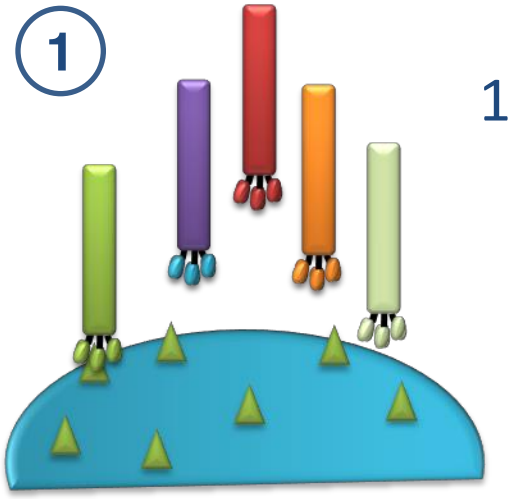


Each phage bring together the scfv protein and its «genetical information»

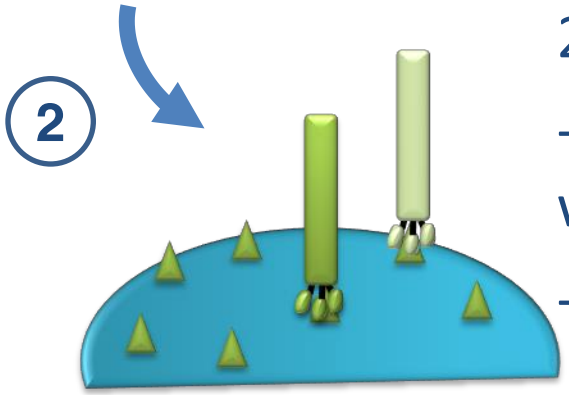
1phage = 1ScFv

Phage display





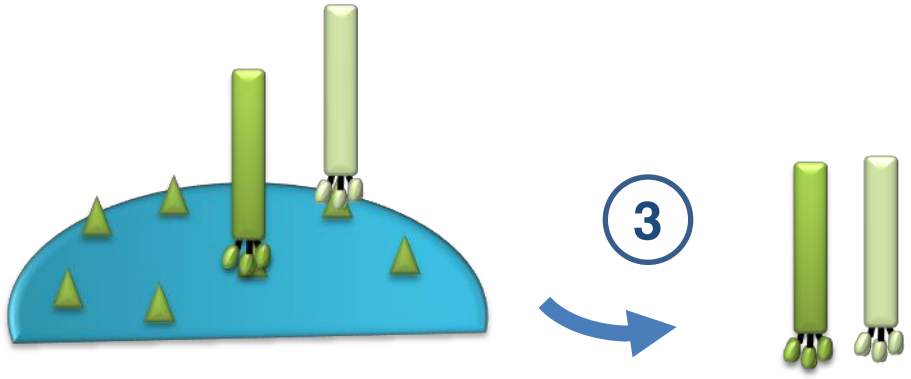
1. The scfv/phage library is incubated with epitope



2.

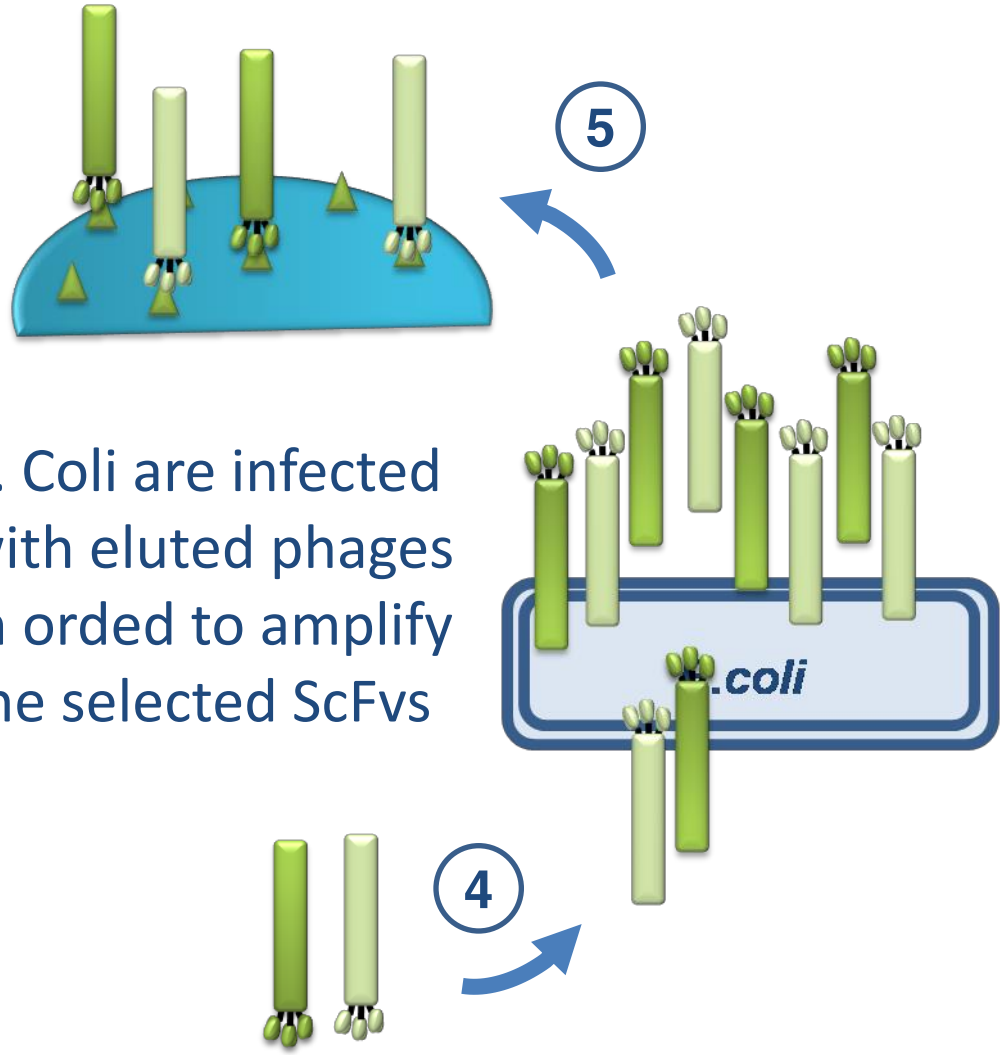
-The non-specific ScFv/phages are washed up.

-Specific ScFv/phage stay on epitope

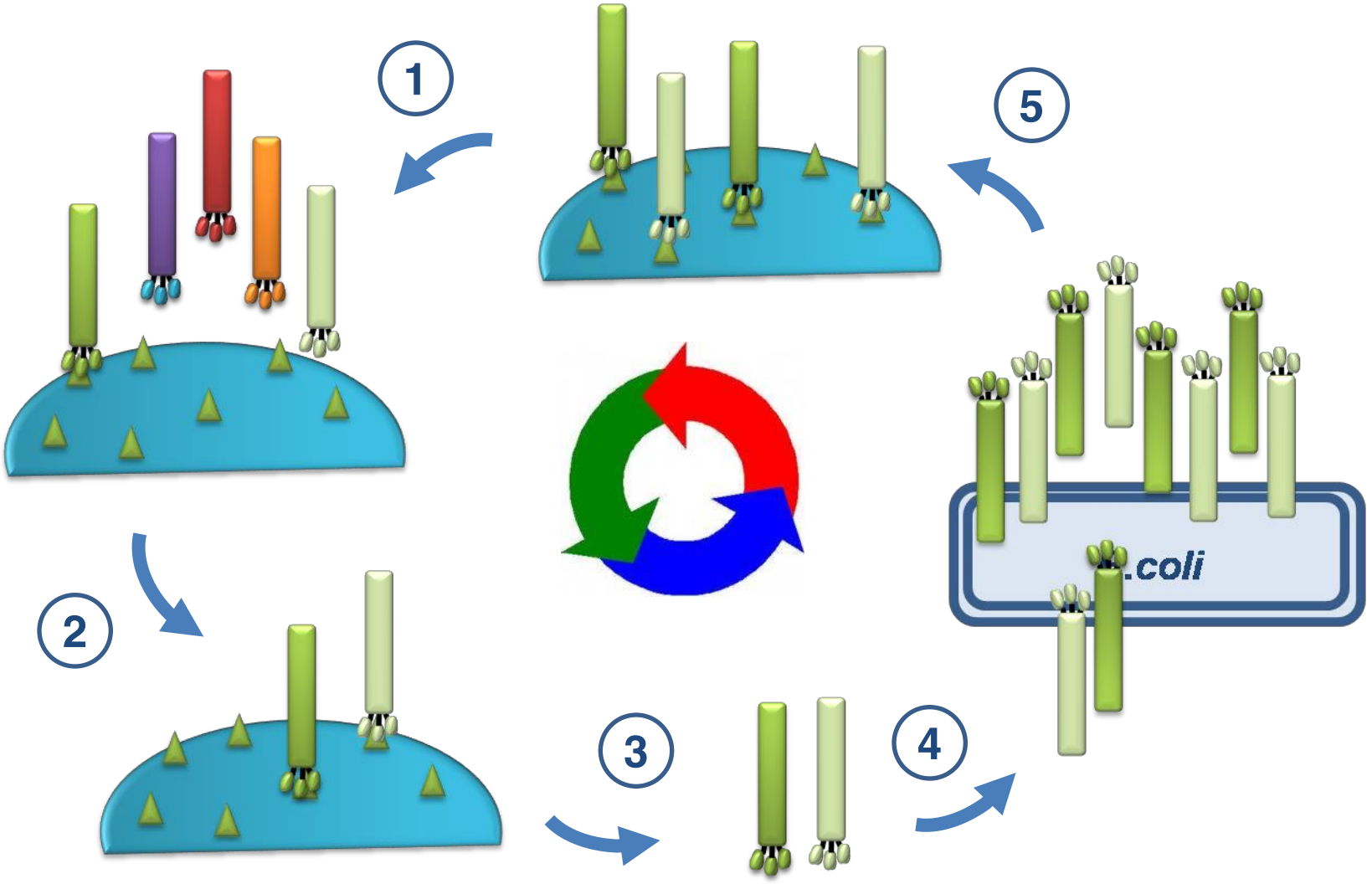


3. The phage bound are eluted (pH, trypsin)

5. Amplified phages are incubated with epitope

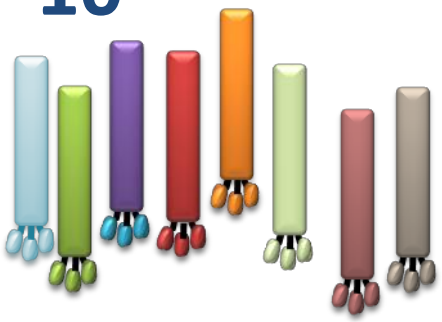


4. E. Coli are infected with eluted phages in orded to amplify the selected ScFvs

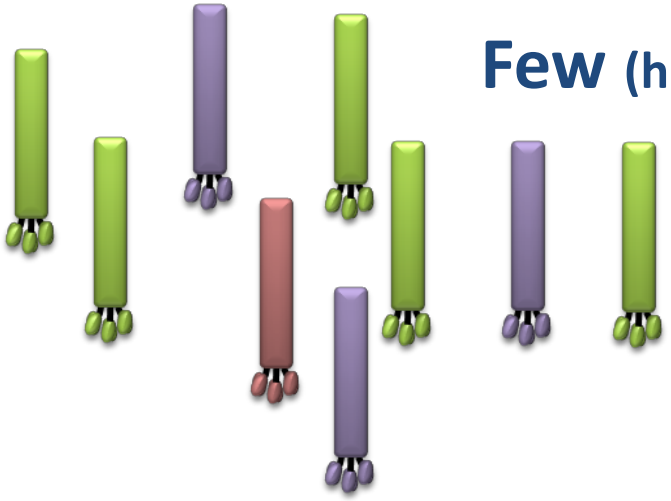

1
2
3
4
5
.coli

From protein to DNA

10^9



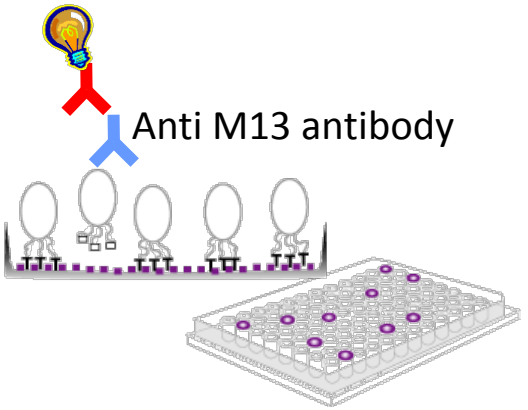
Selection cycles



Few (hundred)

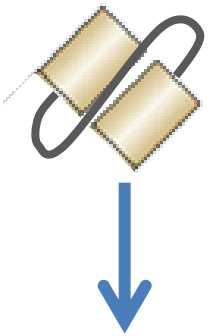
Random phages are tested by ELISA

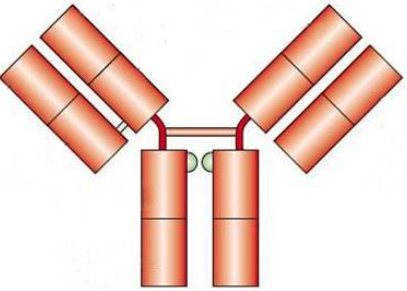
Positive phages from ELISA are sequenced



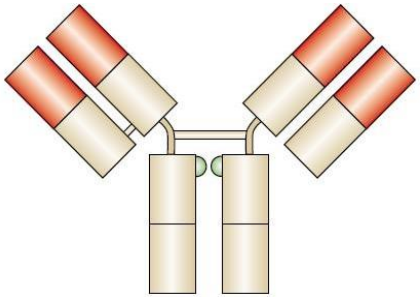
TTAGTTGTTCCTTTCTATGCGGCCCAGCCGGCCATGGCCGAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAG
CCTGGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTCACCTTTAGCAGCTATGCCATGAGCTGGGTCCGCCAG
GCTCCAGGGAAGGGGCTGGAGTGGGTCTCAGCTATTAGTGGTAGTGGTGGTAGCACATACTACGCAGACTCCGTGAAG
GGCCGGTTCACCATCTCCGGAGACAATTCCAAGAACACGCTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACG
GCCGTGTATTACTGTGCAAGGTCGAACTGGTACTTCGATCTCTGGGGCAAAGGGACAATGGTCACCGTCTCGAGTGGA
GGCGGCGGTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGAAGTGCACTTTCTTCTGAGCTGACTCAGGACCCTGCTGTG
TCTGTTGCCTTGGGACAGACAGTCACGCTCACATGCCGAGGAGACAACCTCAGAAGTTATTATGCCAGCTGGTACCAG
CAGCGGCCAGGACAGGCCCCTAGACTTGTCATGTATGGGAAAAACATCCGGCCCTCAGGGATCCCAGACCGATTCTCT
GGCTCCAGCTCAGGGGACACAGGTTCGTTGACCATCACTGGGGCTCAGGCGGAGGATGAGGCTGACTATTACTGCAAC
TCCCGGGACAGCCTTACTAACCGTCCCTTTTTCGGCGGAGGGACCAAGGTCACCGTCCTAGGTGCGGCCGCACATCAT
CATCACCATCACGGGGCCGCAGAACAAAACTCATCTCAGAGAGGATCTGAATGGGGCCGC

ScFvs allow to produce fully human monoclonal antibodies

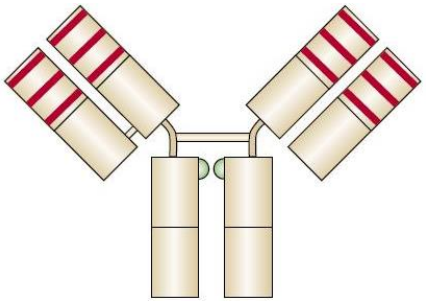




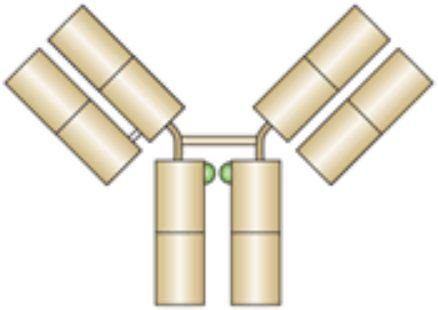
mouse



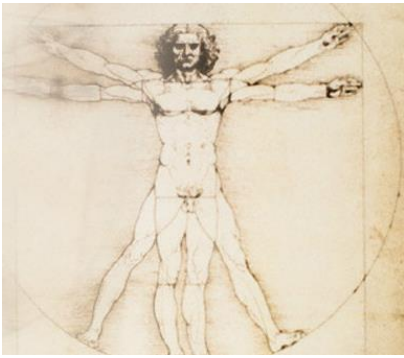
chimeric



humanized

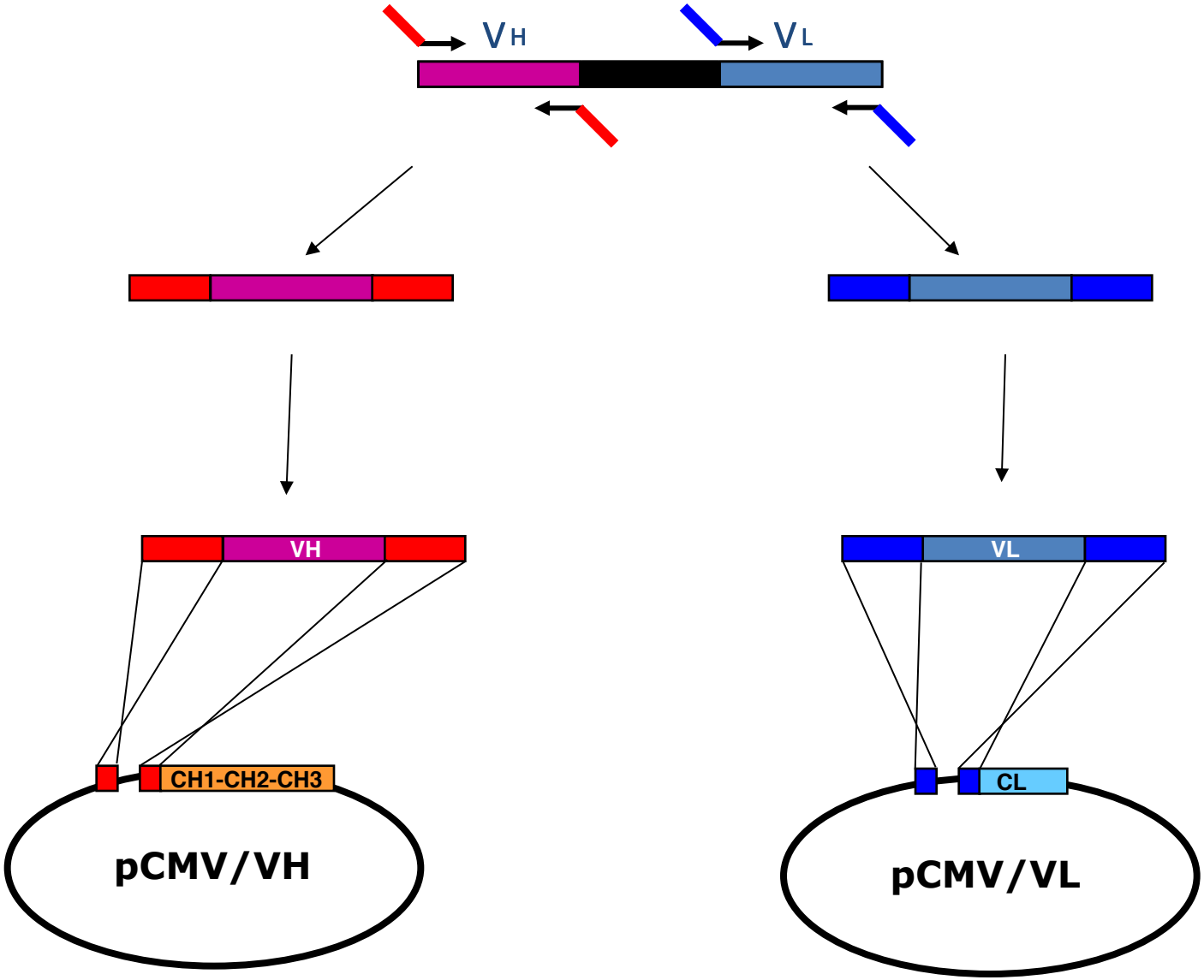


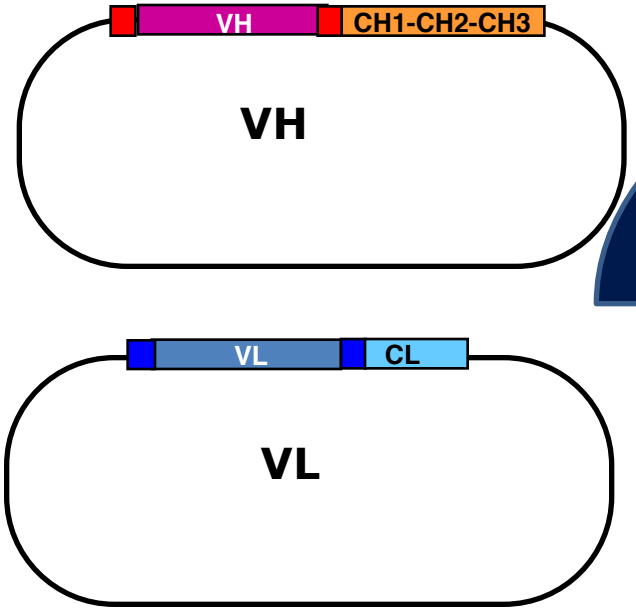
human



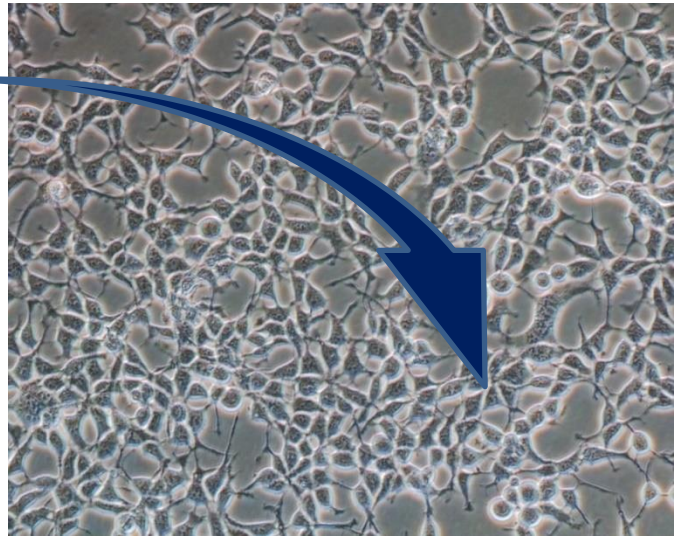
Nature Reviews | Cancer

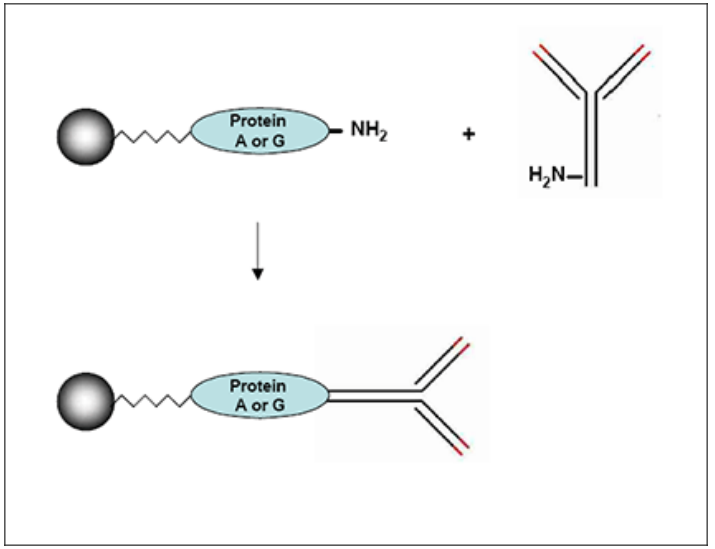
From ScFv to human antibody

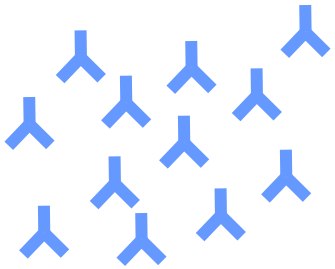




Co- transfection

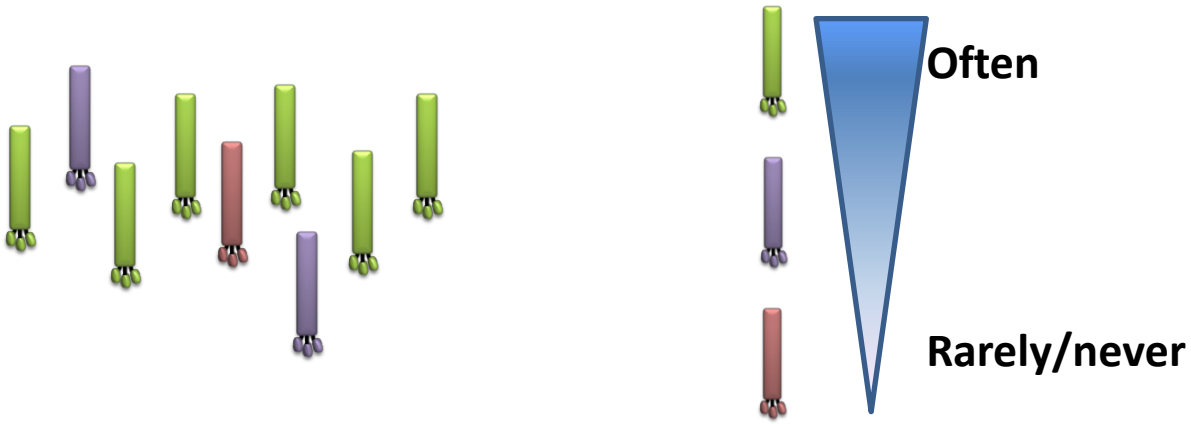






Problems related to classical ELISA screening of phages libraries:

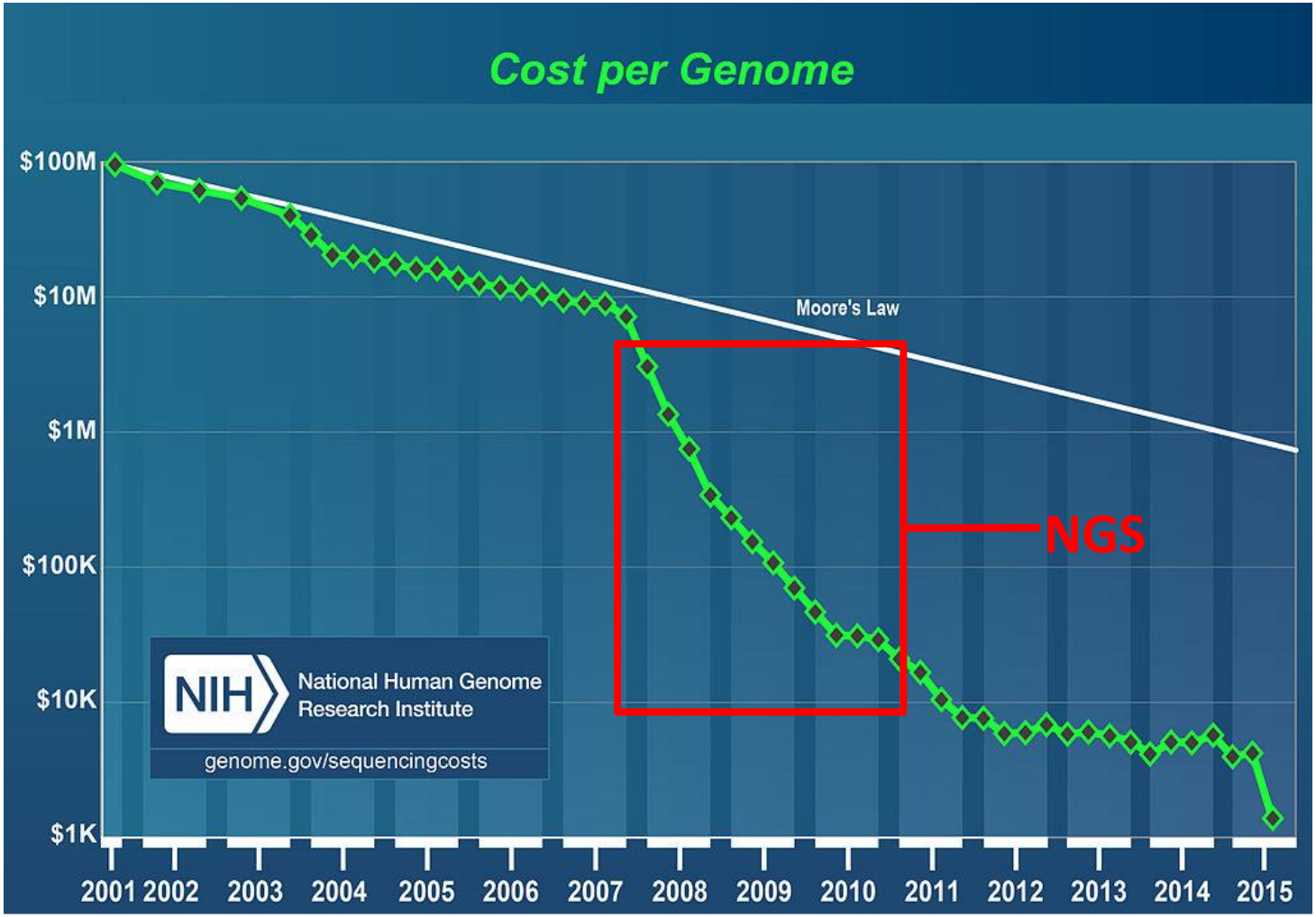
- Repetitive isolation of the same clones within and between selection steps



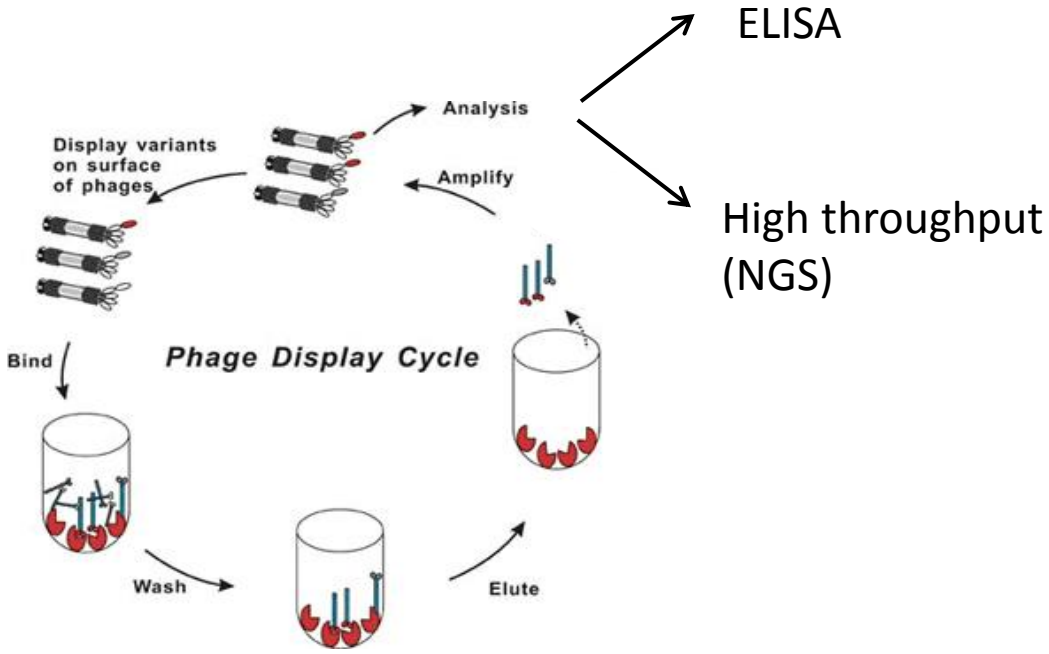
- Information in phage clones enrichment needs redundant sequencing after each step

```
1       CGACCTGGAAAGCGGGCAGTGGAGCGCAACGCAATTAATGTGAGTTTAGCTCACTCATTA
2       CGACCTGGAAAGCGGGCAGTGGAGCGCAACGCAATTAATGTGAGTTTAGCTCACTCATTA
3       CGACCTGGAAAGCGGGCAGTGGAGCGCAACGCAATTAATGTGAGTTTAGCTCACTCATTA
4       CGACCTGGAAAGCGGGCAGTGGAGCGCAACGCAATTAATGTGAGTTTAGCTCACTCATTA
5       CGACCTGGAAAGCGGGCAGTGGAGCGCAACGCAATTAATGTGAGTTTAGCTCACTCATTA
6       CGACCTGGAAAGCGGGCAGTGGAGCGCAACGCAATTAATGTGAGTTTAGCTCACTCATTA
7       CGACCTGGAATATAGGCAGTGGAGCGGGACGCAATTAATGCACGTTTAGCTCACTCATGT
8       CGACCTGGAATATAGGCAGTGGAGCGGGACGCAATTAATGCACGTTTAGCTCACTCATGT
9       CGACCTGGAAAGCGGGCATACTCUGGCAACGCAATTATTCGGTGTTTAGCTCACTCATAC
        **********    ****       *  ********* *    ***************
```

- Continuous handling of clones during selection steps (isolation, growth, ELISA)


Cost per Genome
$100M
$10M
$1M
$100K
$10K
$1K
Moore's Law
NGS
NIH
National Human Genome Research Institute
genome.gov/sequencingcosts
2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015

Selecting scFvs from phage libraries by next generation sequencing



The **deep sequencing approach** could i**dentify under-represented clones** in phage screens. It could also preserve diversity of phage clones and identify ligands previously lost in phage display screens.

How do we sequence scfv?

```
ATGAAATACC TATTGCCTAC GGCAGCCGCT GGATTGTTAT TACTCGCGGC CCAGCCGGCC  60
ATGGCCGAGG TGCAGCTGGT GGAGTCTGGG GGAGGCGTGG TCCAGCCTGG GAGGTCCCTG 120
AGACTCTCCT GTGCAGCCTC TGGATTCACC TTCAGTAGCT ATGCTATGCA CTGGGTCCGC 180
CAGGCTCCAG GCAAGGGGCT GGAGTGGGTG GCAGTTATAT CATATGATGG AAGCAATAAA 240
TACTACGCAG ACTCCGTGAA GGGCCGATTC ACCATCTCCA GAGACAATTC CAAGAACACG 300
CTGTATCTGC AAATGAACAG CCTGAGAGCT GAGGACACGG CCGTGTATTA CTGTGCAAGA 360
GGGTTTCCTA TGCCGTGGGG CCAAGGTACC CTGGTCACCG TGTCGAGAGG TGGAGGCGGT 420
TCAGGCGGAG GTGGCTCTGG CGGTGGCGGA TCGTCTGAGC TGACTCAGGA CCCTGCTGTG 480
TCTGTGGCCT TGGGACAGAC AGTCAGGATC ACATGCCAAG GAGACAGCCT CAGAAGCTAT 540
TATGCAAGCT GGTACCAGCA GAAGCCAGGA CAGGCCCCTG TACTTGTCAT CTATGGTAAA 600
AACAACCGGC CCTCAGGGAT CCCAGACCGA TTCTCTGGCT CCAGCTCAGG AAACACAGCT 660
TCCTTGACCA TCACTGGGGC TCAGGCGGAA GATGAGGCTG ACTATTACTG TAACTCCCGG 720
GACAGCAGTG GTAACCATGT GGTATTCGGC GGAGGGACCA AGCTGACCGT CCTAGGTGCG 780
GCCGCAGAAC AAAAACTCAT CTCAGAAGAG GATCTGAAT
```

-The whole scfv consists of about 800 bp. It is too long for NGS sequencing

-The separated sequencing of VH and VL would be too expensive and doesn't allow to associate the VH with the correspondent VL

Which technology?

HiSeq 2500/2000

HiSeq 2500

MiSeq

Roche 454

Mode	High Output	Rapid Run		
Output (maximum)	600Gb	180Gb	8.5Gb	1 Gb
Run Time	2 - 11 Days	7 - 40 Hr.	4 - 39 Hr.	10-20Hr.
Paired-end Reads (maximum)	6 Billion 6x10^12	1.2 Billion 10^12	15 Million 15x10^6	1 million 10^6
Single-end Read (maximum)	3 Billion	600 Million	7.5 Million	
Max Read Length	2x100 bp	2x150 bp	2x250 bp	700 bp

MiSeq platform(Illumina) allows to cover the whole Heavy Variable region with a coverage al least of 100x

Advantages	Disadvantages
•Cheap •Good coverage **10^7**	•NGS doesn't allow to sequence the whole fragment of VH and VL

The VH is usually more involved than VL in epitope binding.

	MiSeq
Mode	
Output (maximum)	8.5Gb
Run Time	4 - 39 Hr.
Paired-end Reads (maximum)	15 Million
Single-end Read (maximum)	7.5 Million
Max Read Length	2x250 bp
Bases above Q30 (2x100bp)	>85% (2x100 bp)
Required Input	1ng with NexteraXT 50 ng with Nextera 100 ng – 1 µg with TruSeq

Workflow of high throughput phage display selection

1. Perform 4 selection cycles by phage display

2. Excise VHs from the 4 sub-libraries of phages by restriction endonucleases

VH sequencing of sub-libraries

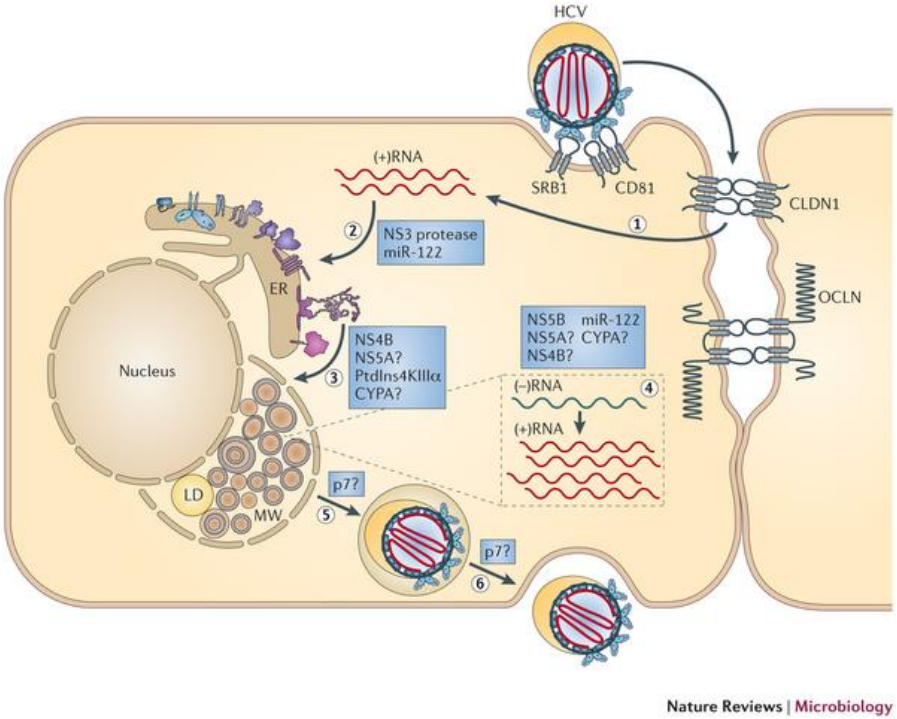
4. Bioinformatic analysis

5. Recovery of the whole ScFvs (VH+VL)

ELISA of the recovered clones to confirm binding

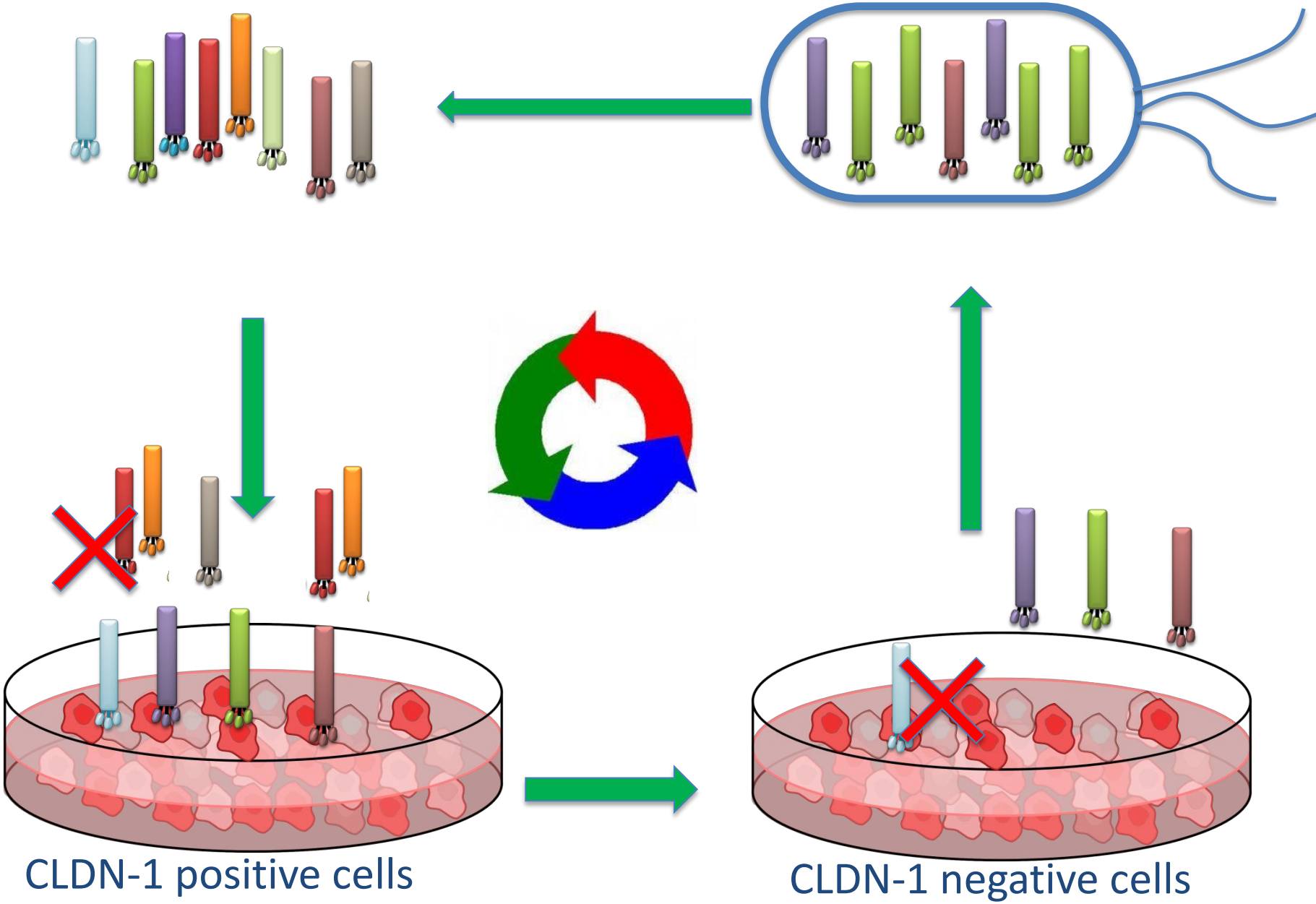
Proof of concept: NGS on 4 cycles of selection on Claudin 1 expressing cells

Hepatitis C virus (HCV) is a major cause of chronic hepatitis and liver carcinoma and new therapies based on novel targets are needed.

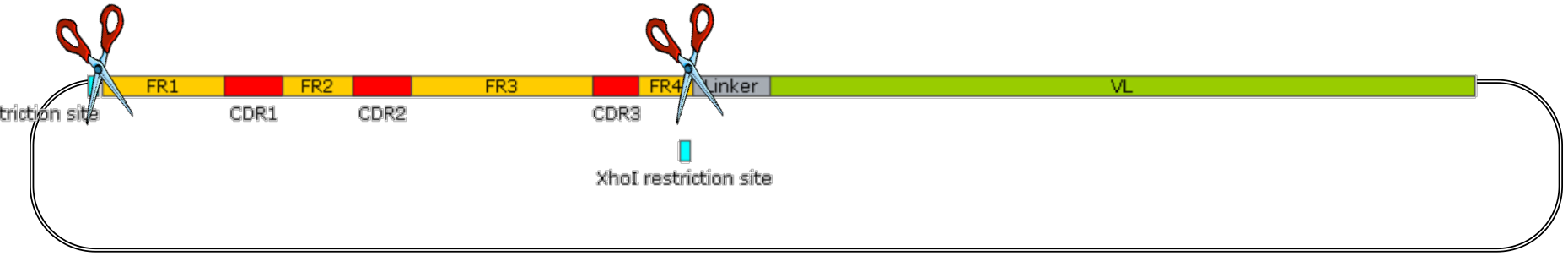


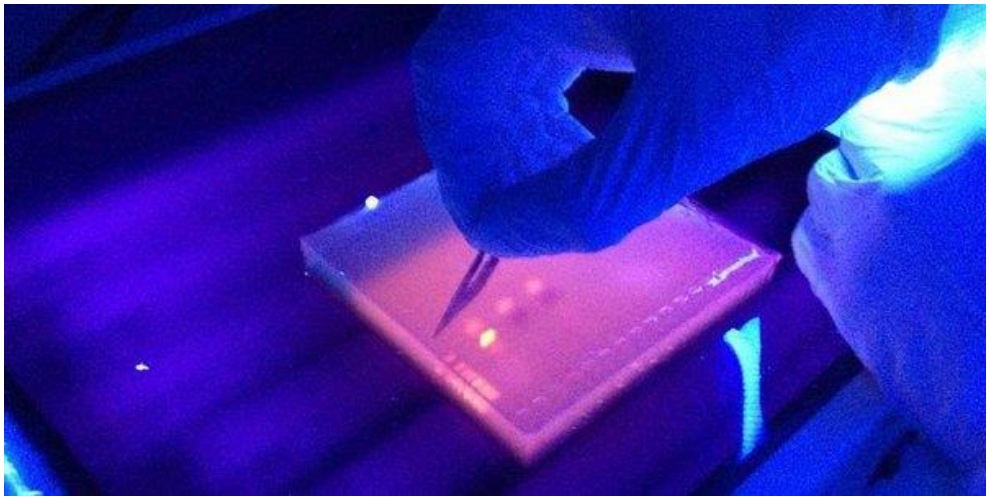
- The tight junction protein Claudin-1 (CLDN-1) is essential for hepatitis C virus (HCV) cell entry and spread.
- Human antibodies are not available
- We screened a phage display library of human single chain antibody fragments (scFv) by using a panel of CLDN-1-positive and –negative cell lines and identified phage specifically binding to CLDN-1.

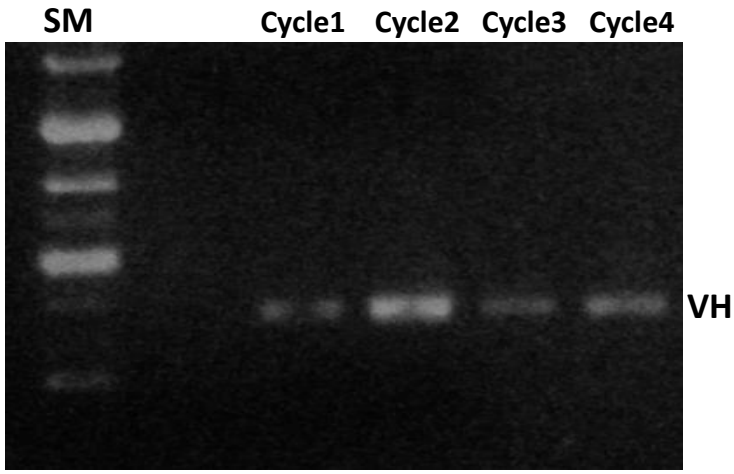
1. Perform 4 selection cycles by phage display

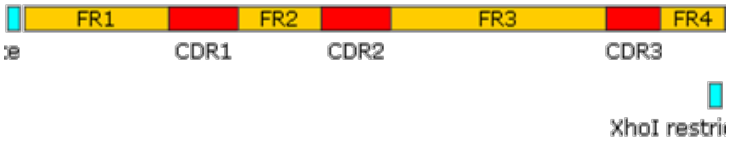


2. Isolation and purification of VHs from scFv plasmid

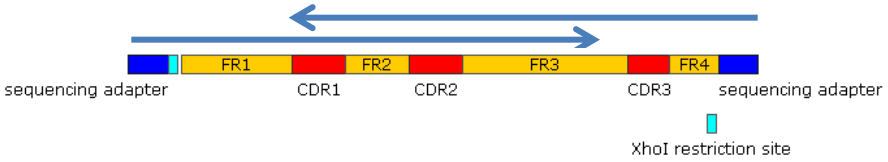


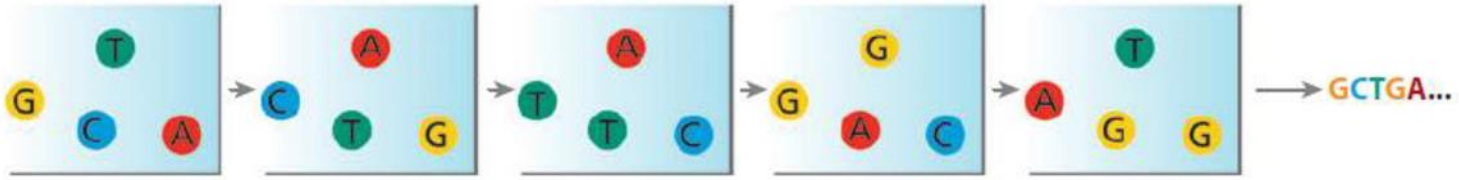




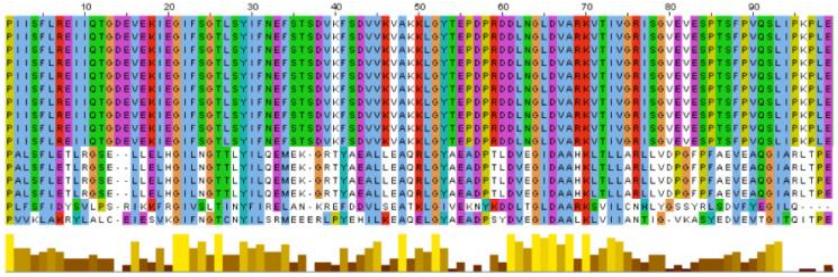


3. Construction of library and sequencing paired-end reads (2x300 v3)

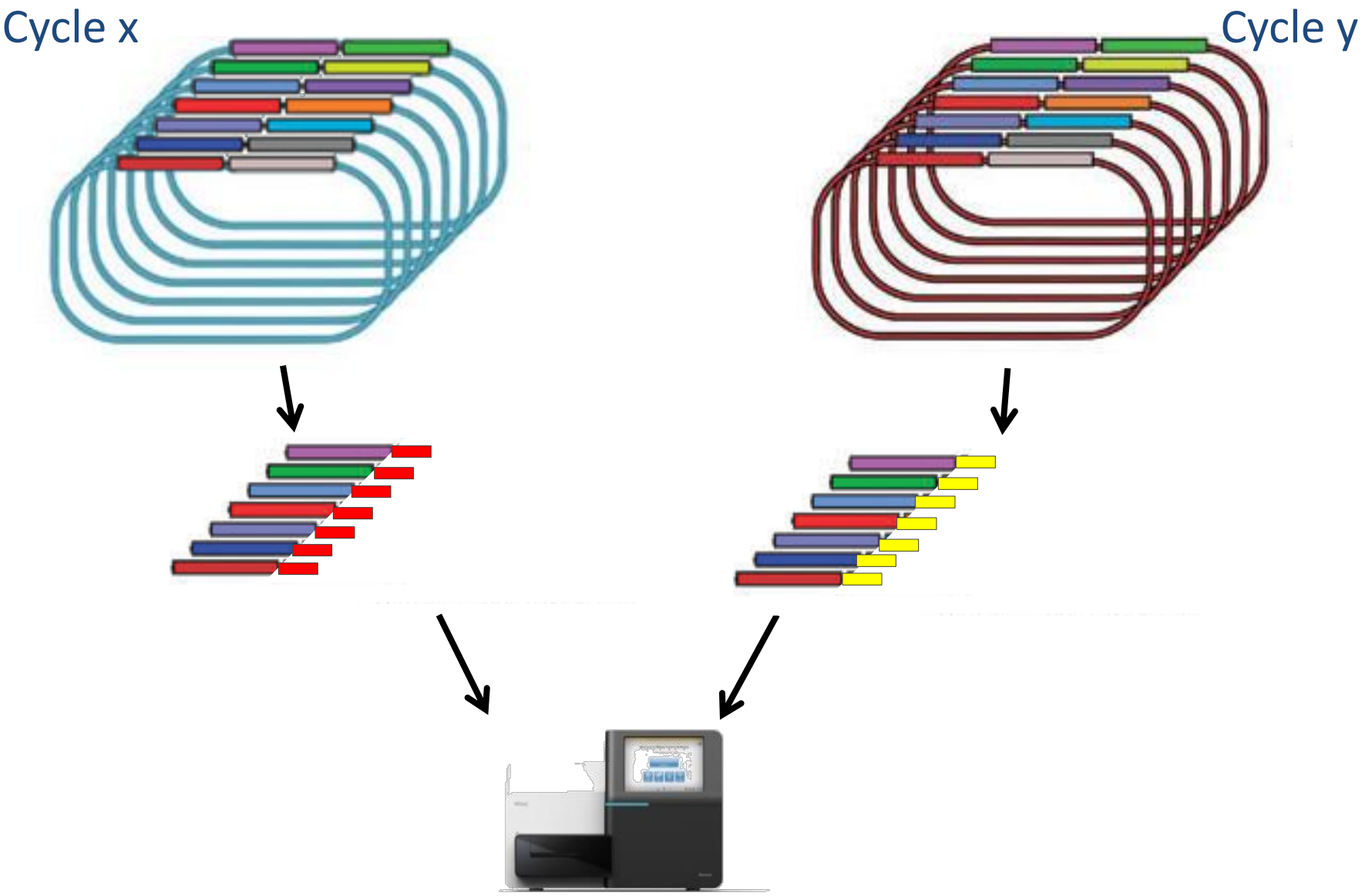




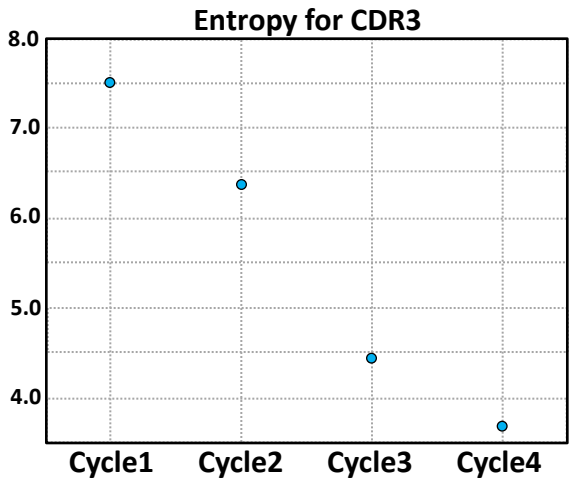
4. Bioinformatics: clustering sequences based on counts

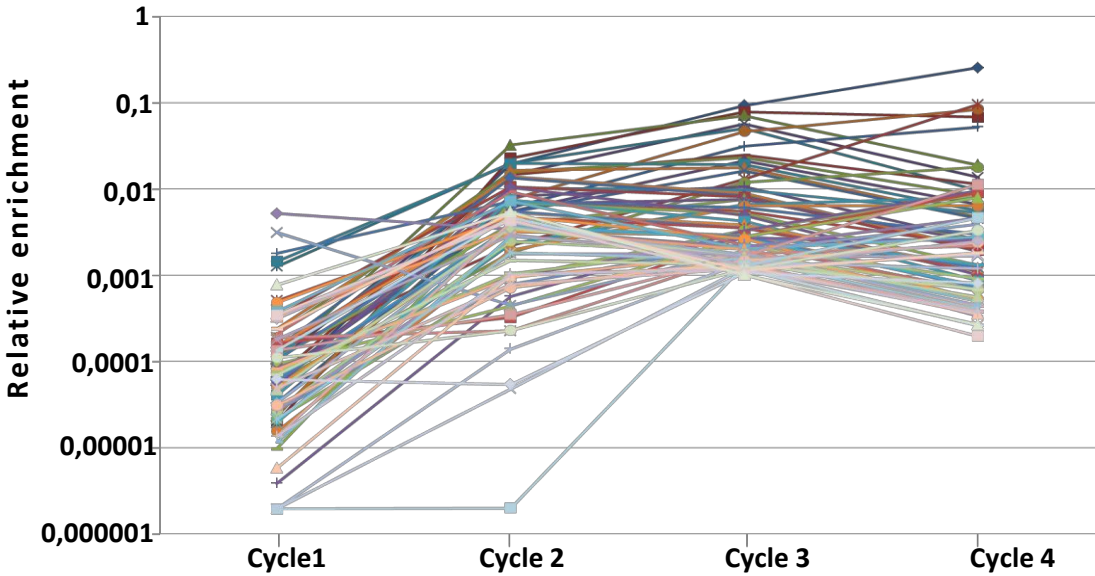


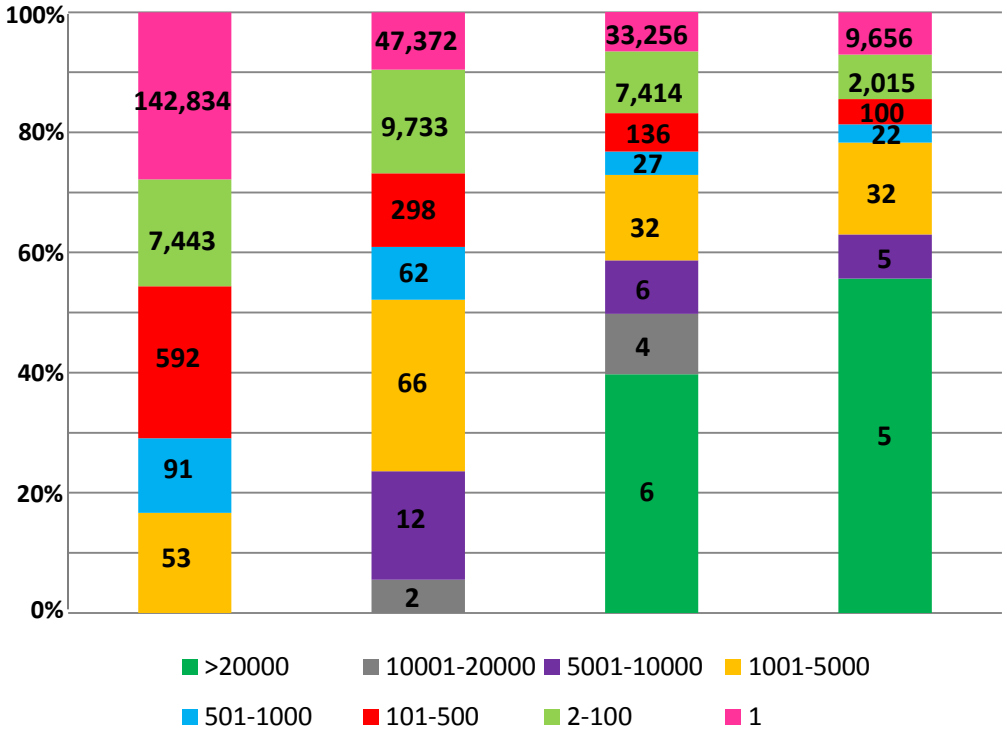
Using different barcodes it's possible to sequence VHs extracted at least from two selection cycles in the same run of sequencing



Enrichment of sequences within and between the cycles







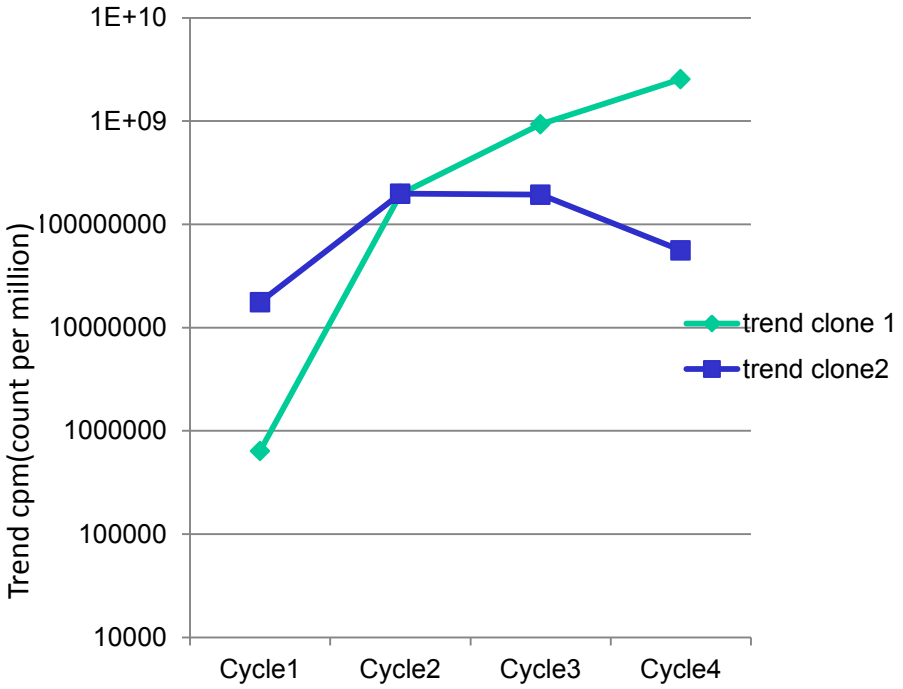
Cycle	1	2	3	4
Maximal relative enrichment	0.76%	3.24%	9,30%	25,49%
Total number of clones	151,013	57,545	40,881	11,835

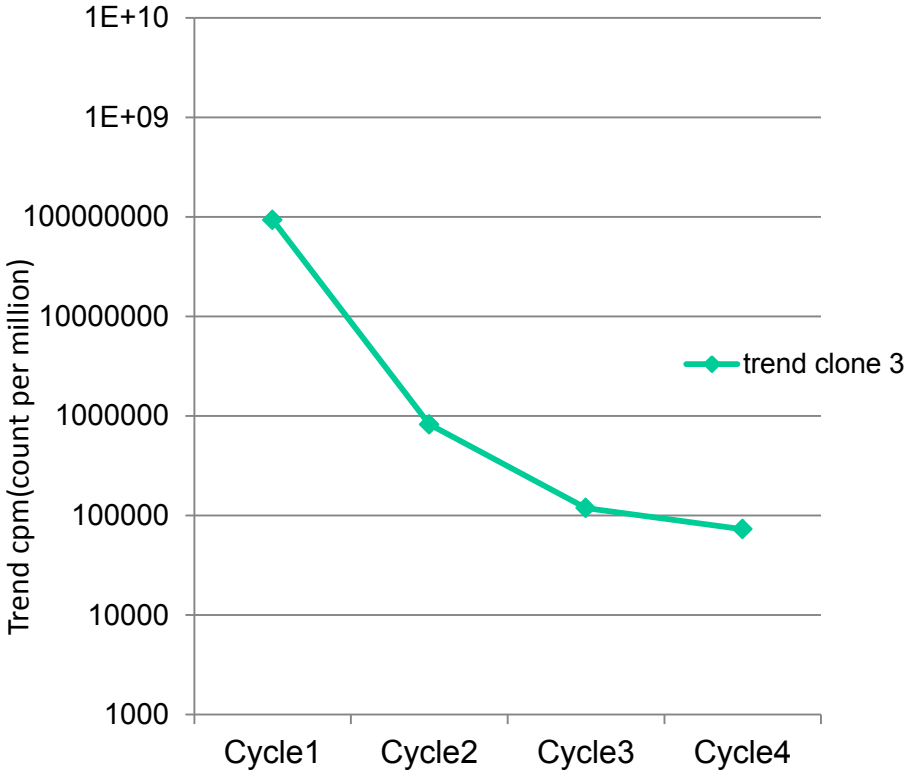
Output of bioinformatic analysis

NGS allows us to rapidly identify the potential binders, based on the counts of the corresponding VH fragment, within a cycle, and on the kinetic of their enrichments, within consecutive cycles

	Relative abundance			
VH sequence	Cycle1	Cycle2	Cycle3	Cycle4
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	638280	197490000	930730000	2549800000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	1583900	21999000	131300000	960500000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	307320	76484000	465810000	847170000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	236400	229510000	790000000	684690000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	401880	61235000	314890000	523510000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	1276600	324720000	714790000	190820000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	709200	18403000	119640000	180040000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	1205600	147480000	570190000	137140000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	1631200	3495700	16746000	113760000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	803760	146700000	247810000	113110000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	47280	5705700	33097000	96501000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	15461000	200660000	502160000	96064000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	1985800	3254600	28162000	91330000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	260040	9402300	33235000	87689000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	2505800	156120000	230490000	83683000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	330960	23044000	66550000	80333000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	1134700	4440000	27924000	79532000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	803760	55007000	213300000	66859000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	189120	19628000	64112000	63654000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	17612000	198210000	193660000	55934000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	1418400	102960000	97209000	53094000
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	283680	164720000	176010000	47486000
...				
GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTG	2788	3	0	0

Examples mode of enrichment or trend of several clones according to NGS analysis

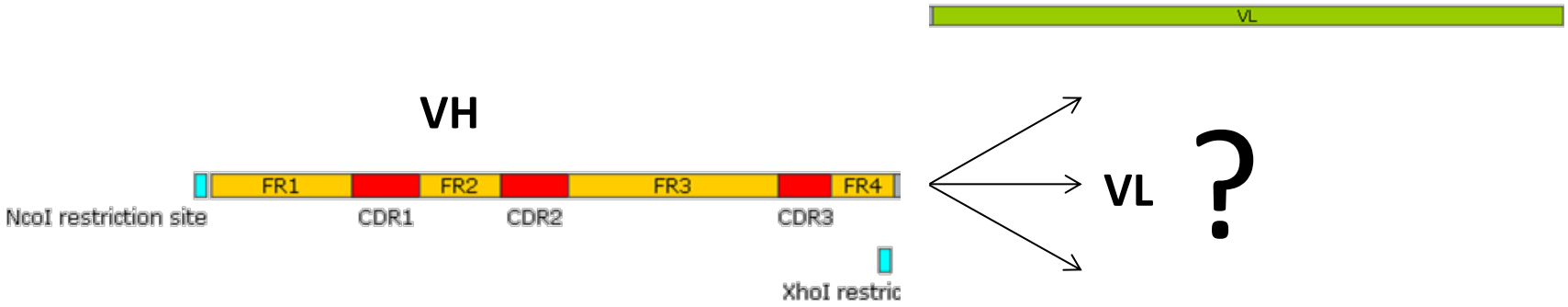




- The analysis of NGS data shows that the mode of enrichment or trend could be different between several potential good binders.
- Non-specific binders could be counter-selected during selection steps.

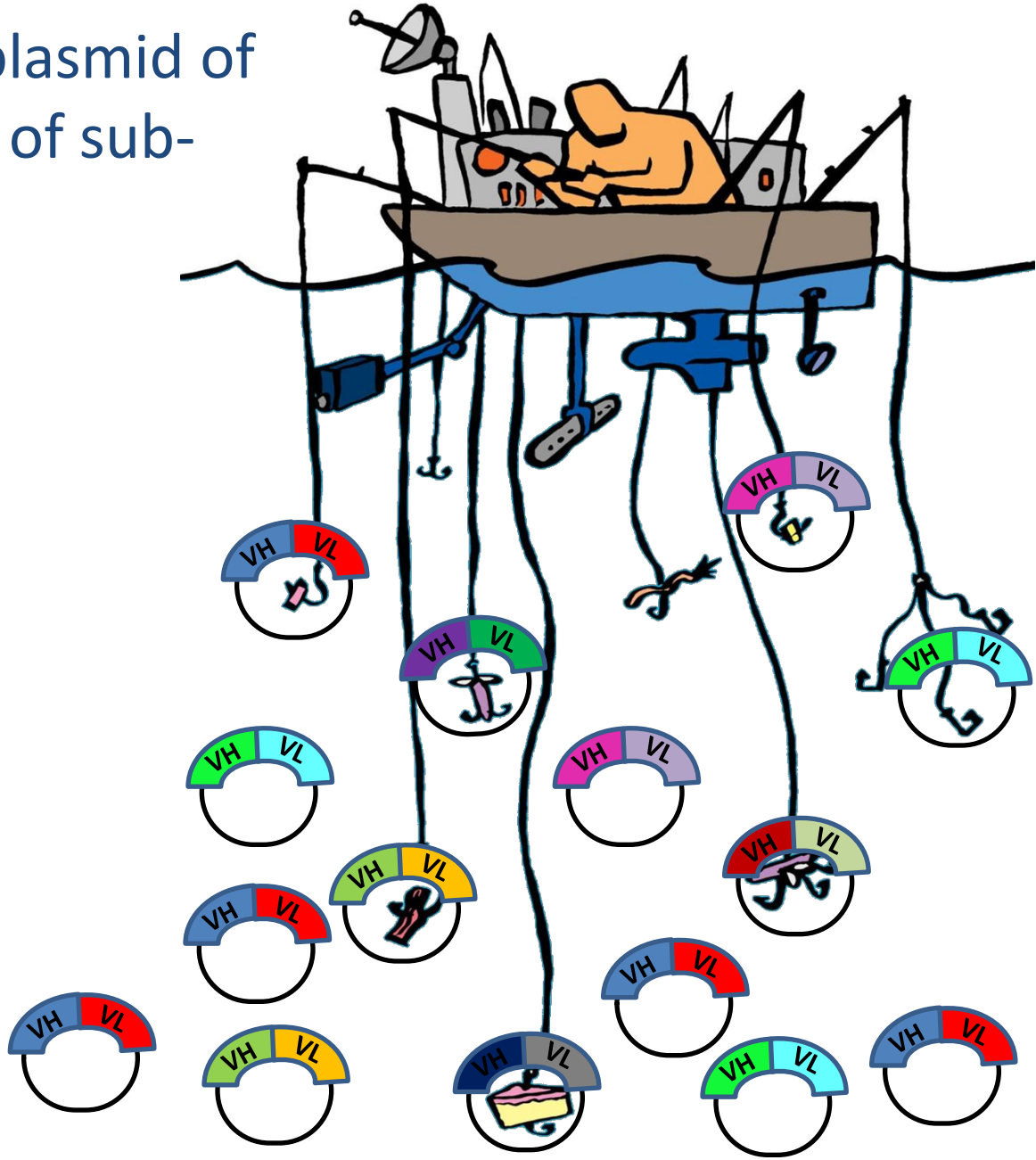
Problems:

- Sequencing only the extracted VH, we have no information about VL sequence
- How to identify the VL associated to the enriched VH of interest?

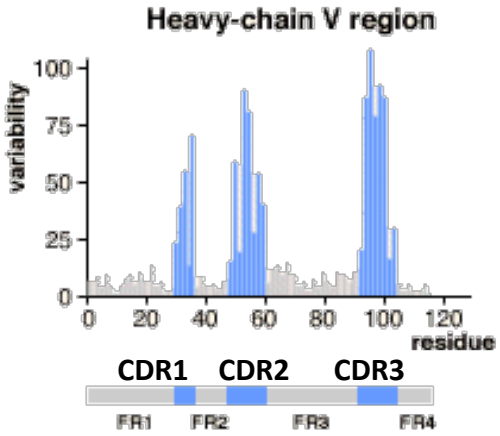


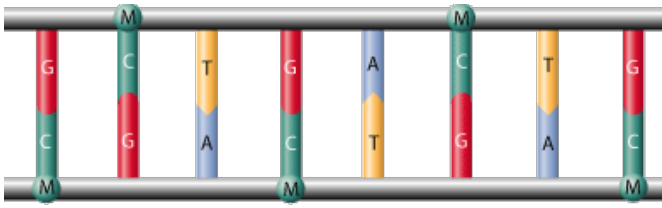
Problems:

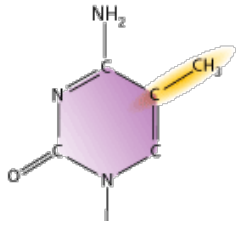
- How to rescue the plasmid of interest from the pool of sub-libraries?



5. Rescue strategy of the whole ScFv (VH+VL)



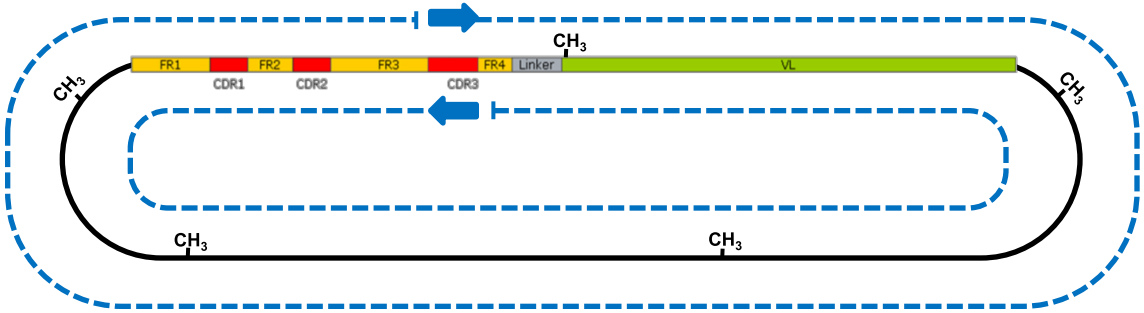




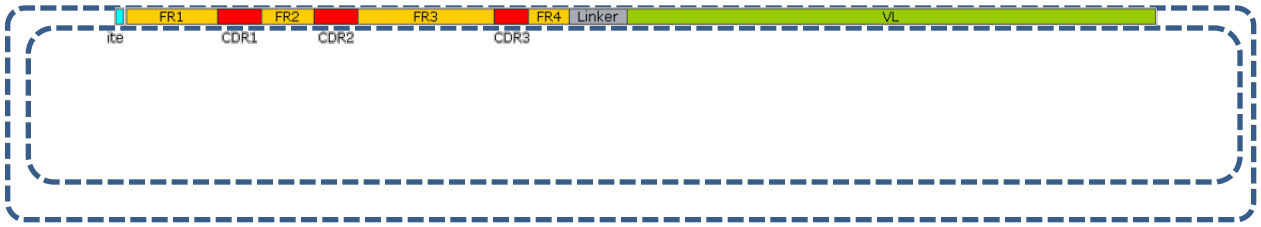
DNA methylation is the addition of a methyl group (M) to the DNA base cytosine (C).

The **CDR3** of VH is the **most variable** region of an scFv, and it's **almost unic**

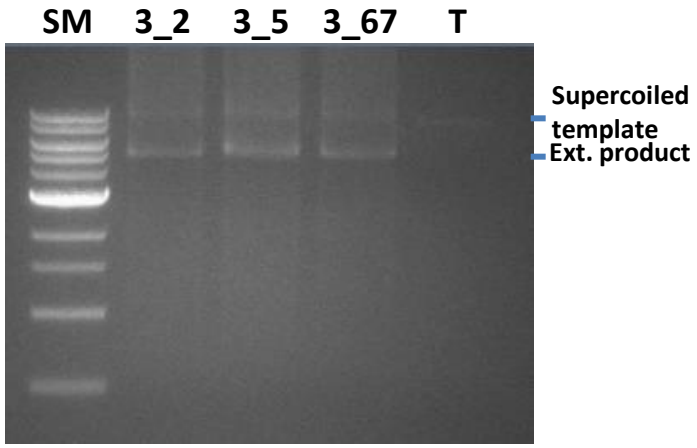
The oligonucleotide sequences were designed inside the HCDR3 regions



Dpn1 allows to degradate the methylated and hemimethylated template DNA



5. Rescue strategy of the whole ScFv (VH+VL)



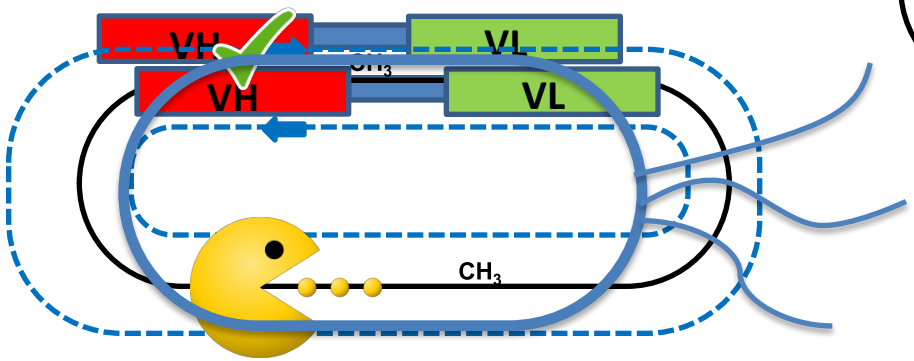
We decided to rescue 3 clones according to:

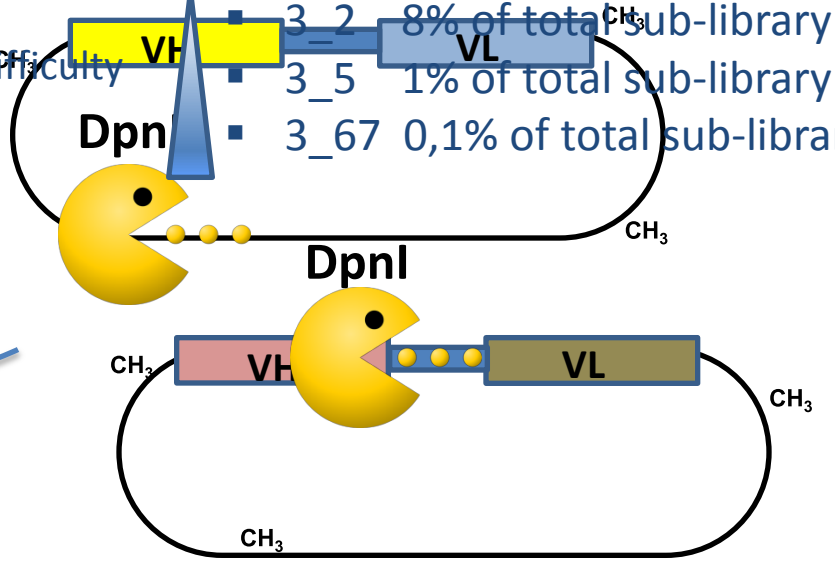
- CDR3 lenght:
  - Long CDR3 (17 a.a., clone 3_2)
  - Medium CDR3 (13 a.a., clone 3_67)
  - Short CDR3 (10 a.a., clone 3_5)

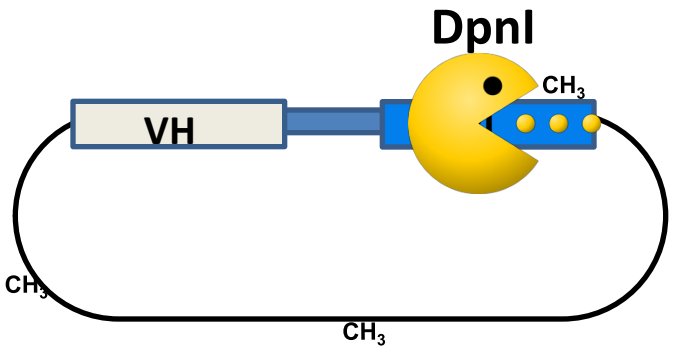
Growing constraint in design of specific primers

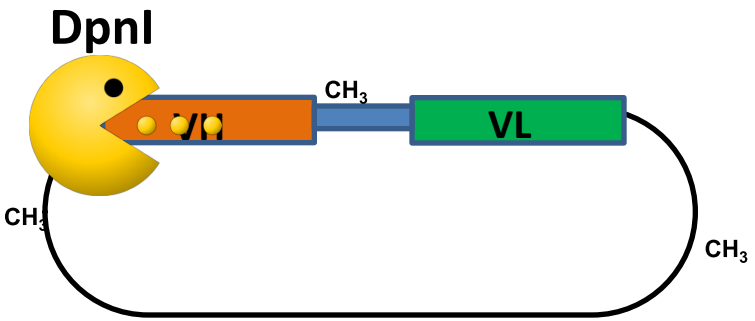
- Relative enrichment
  - 3_2 8% of total sub-library
  - 3_5 1% of total sub-library
  - 3_67 0,1% of total sub-library

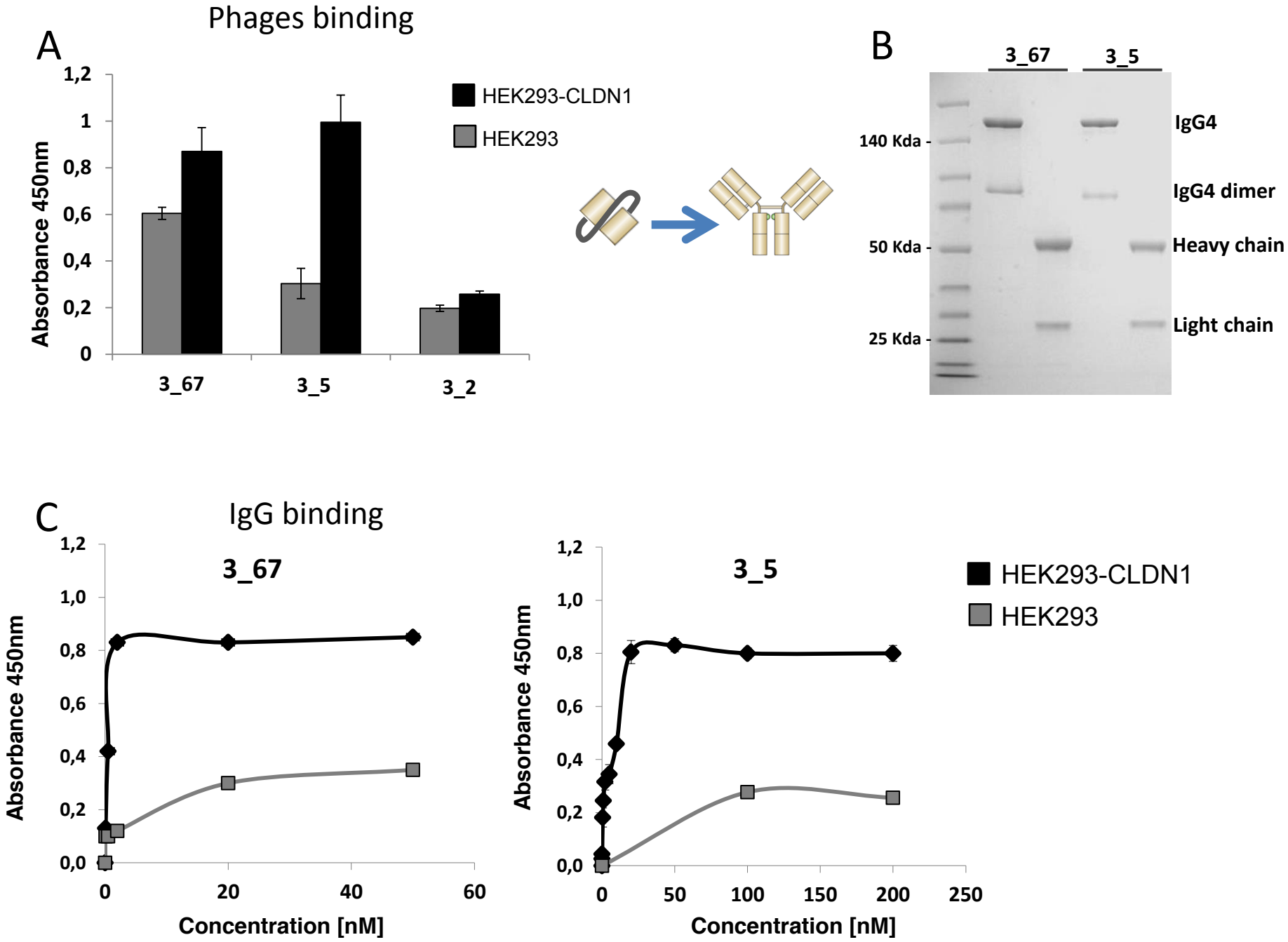
Growing difficulty in rescue



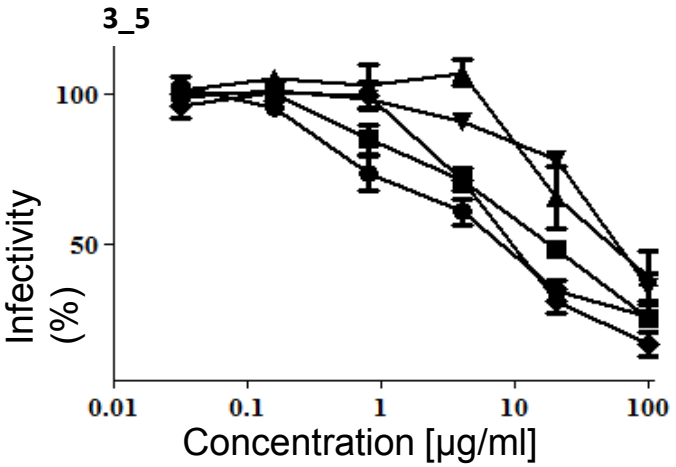






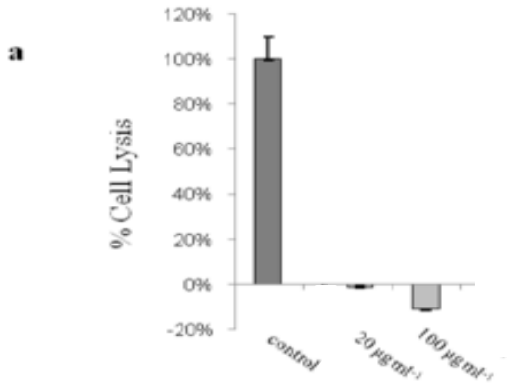

A
Phages binding
Absorbance 450nm
1,2
1
0,8
0,6
0,4
0,2
0
3_67
3_5
3_2
HEK293-CLDN1
HEK293
B
3_67
3_5
140 Kda
50 Kda
25 Kda
IgG4
IgG4 dimer
Heavy chain
Light chain
C
IgG binding
3_67
3_5
Absorbance 450nm
1,2
1,0
0,8
0,6
0,4
0,2
0,0
0
20
40
60
50
100
150
200
250
Concentration [nM]
HEK293-CLDN1
HEK293

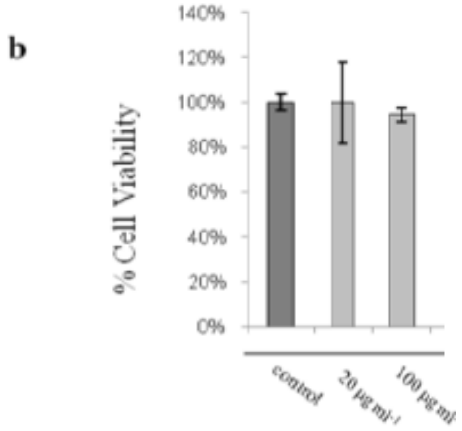
3_5 antibody counteracts HCV infection



		EC50 (µg/ml)				
mAbs	Kd (nM)	H77/JFH-1	JFH1	2B1.1	2B2.8	J6
D10X	2.9	11	8	57	61	17

Neutralization curves of 3_5 anti-claudin 1 antibody. The mAb was tested on Huh7.5 cells treated with 2B2.8 (▼), JFH-1 (●), J6 (■), 2B1.1 (▲) and H77/JFH-1 (♦). 3_5 has capabilities against the isolates tested.





a) percentage of cell lysis after incubation with 3_5 anti-CLDN-1 mAb; as positive control, cells were treated with 1% triton X-100.

b) percentage of cell viability after 72 h incubation with 3_5 anti-CLDN-1 mAb measured by MTT.

Comparing the classical ELISA with NGS to identify positive scFvs during phage library selections

High throughput	Classical (ELISA)
Enrichment evaluation and sequence analysis are contextual	Information in phage clones enrichment needs sequencing after each step
	Repetitive isolation of the same clones within and between selection steps
Minimal handling of clones during selection steps	Continuous handling of clones during selection steps (isolation, growth, ELISA)
Sequence analysis limited to VH	Possibility to sequence the whole scFv
Post-analysis recovery of clones	Positive clones recovered after each cycle

Conclusions

- This novel approach can guarantee rapid and cheap isolation of antibodies for virtually any antigen involved in human diseases, for therapeutic and/or diagnostic applications.
- NGS analysis of phage display libraries allows to identify good binder clones
- New potential binders are identified by NGS compared to the classical method
- Novel human anti-CLDN-1 mAbs represent useful reagents for anti-HCV therapy.

Future perspectives

- Application of high throughput screening for isolation of news monoclonal antibodies.
- Overcome problems related to short reads of NGS technologies to obtain the whole scfv sequence.

Nicola Zambrano

Alfredo Nicosia

Claudia De Lorenzo

Riccardo Cortese

Francesco D'Auria

Guendalina Froechlich

Rolando Paciello

Gennaro Riccio

Richard A. Urbanowicz

Patrick C. McClure

Jonathan K. Ball

Epitope or «in vivo» display?