Formal representation of the repertoire of IgM antibody specificities
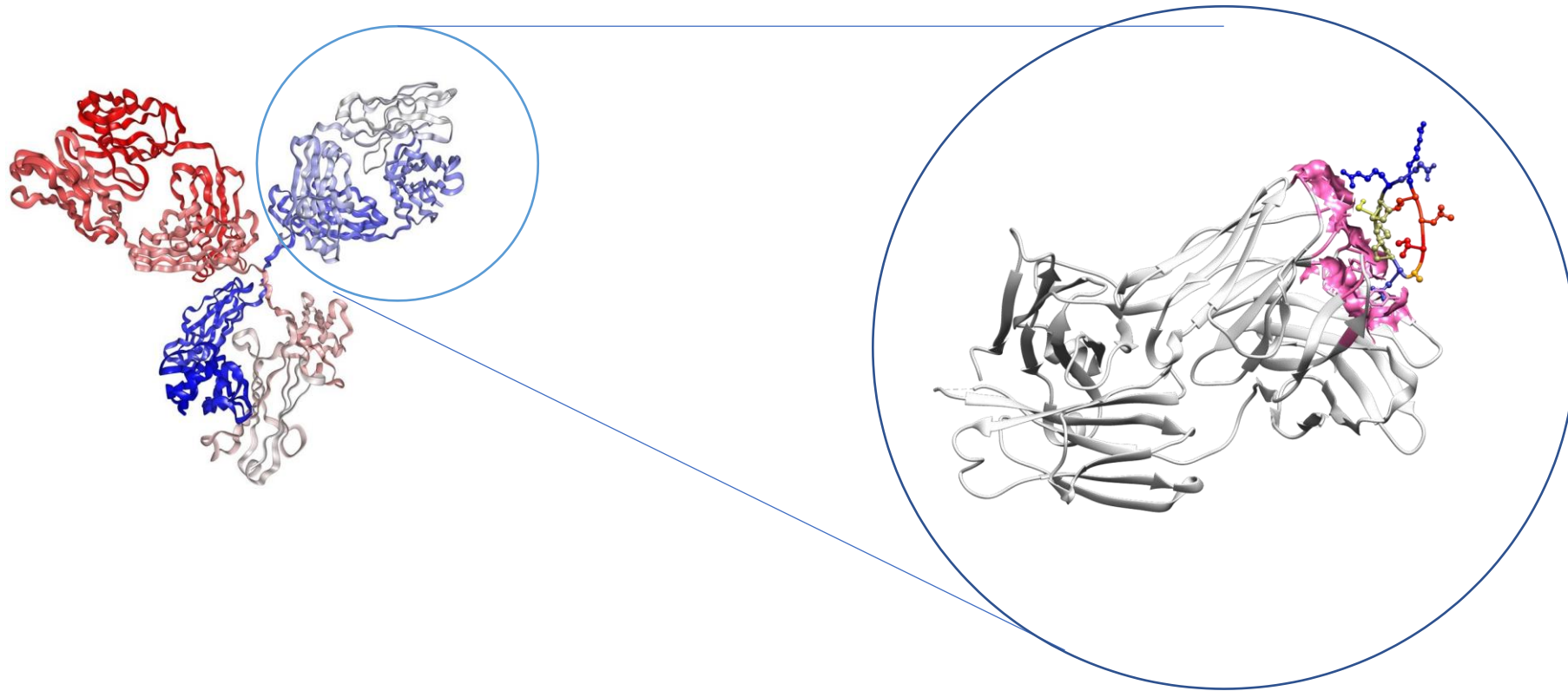
Anastas Pashov[1], Shina Pashova[2], Peter Petrov[3]

[1]Institute of Microbiology, Bulgarian Academy of Sciences

[2]Institute of Biology and Immunology of Reproduction, Bulgarian Academy of Sciences

[3]Institute of Mathematics and Informatics, Bulgarian Academy of Sciences

Mathematics of Life, September 13-16, 2021, Hisarya, Bulgaria

Mathematics of Life 2021

WHY IgM?

Routine practice:
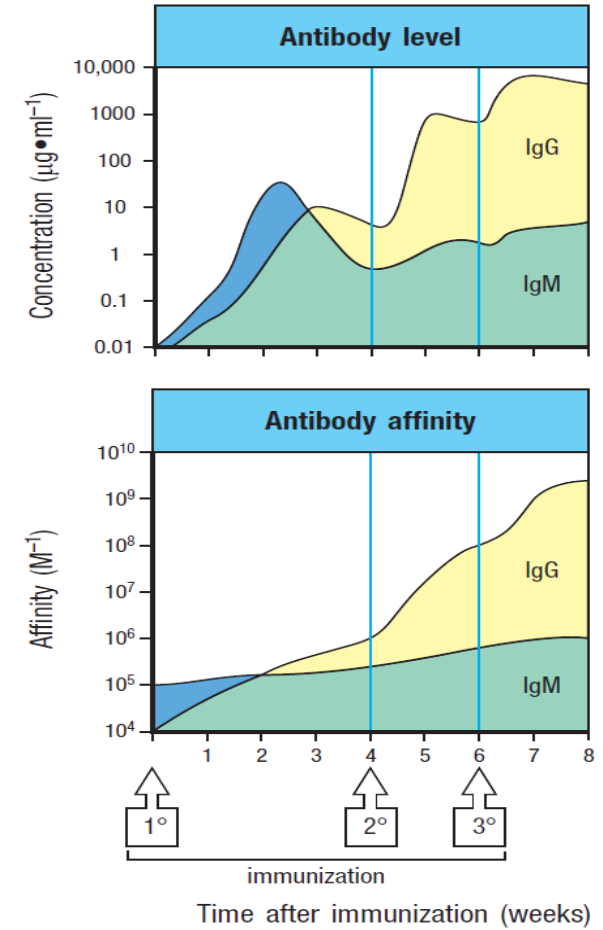
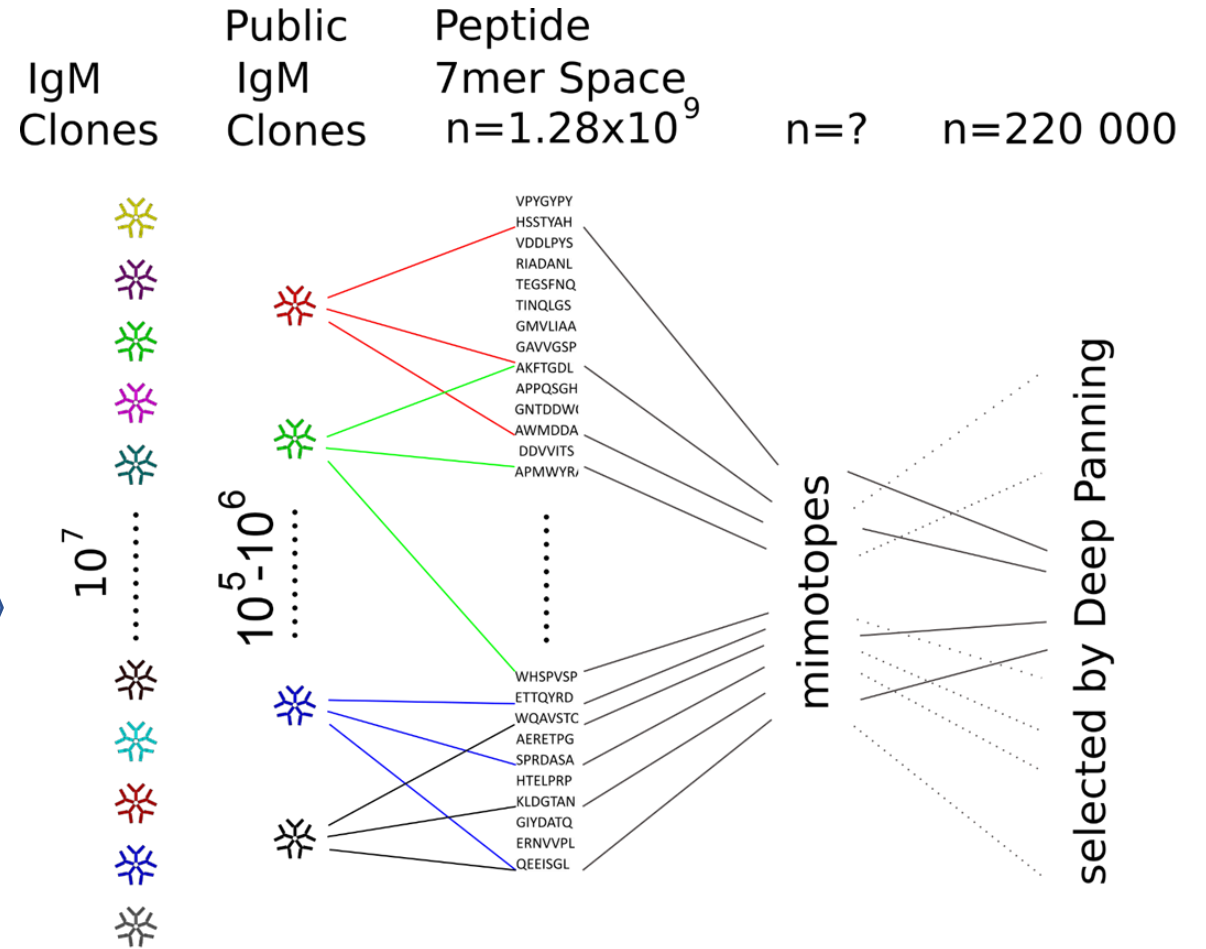- Do I have antibodies to ...

(SARS-Co-V2, HBV, HIV, etc) ?

VS
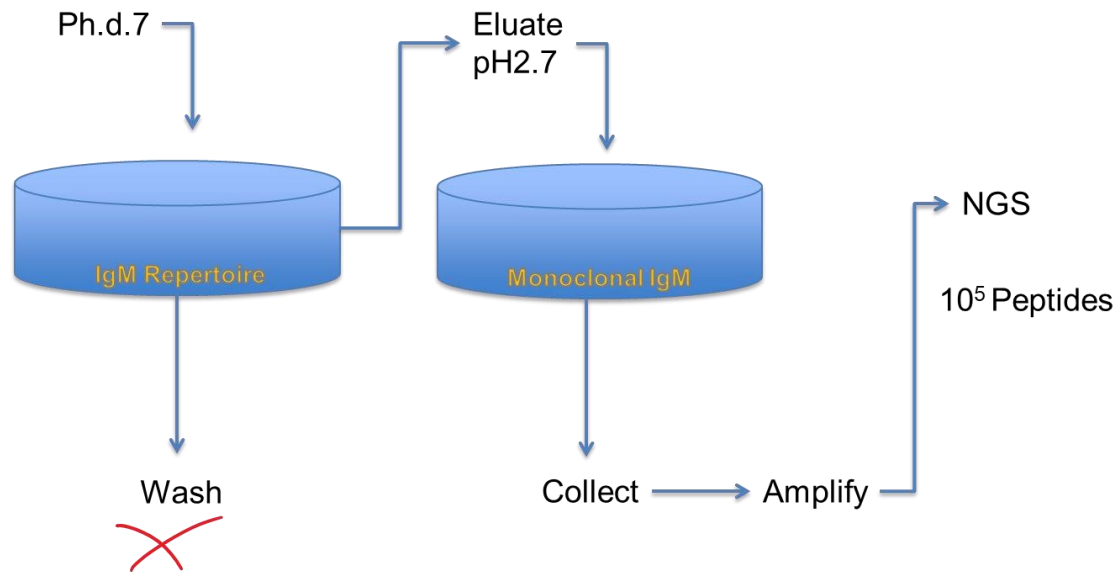
The "omics" view:

-What is the global map of my repertoire?
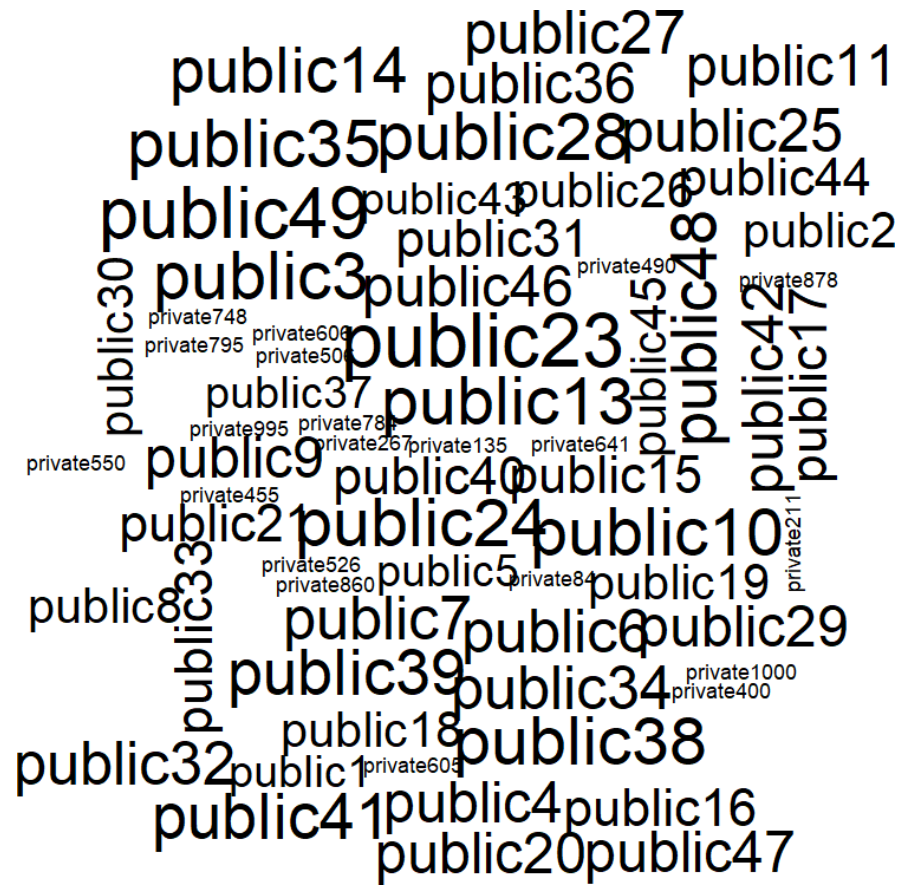
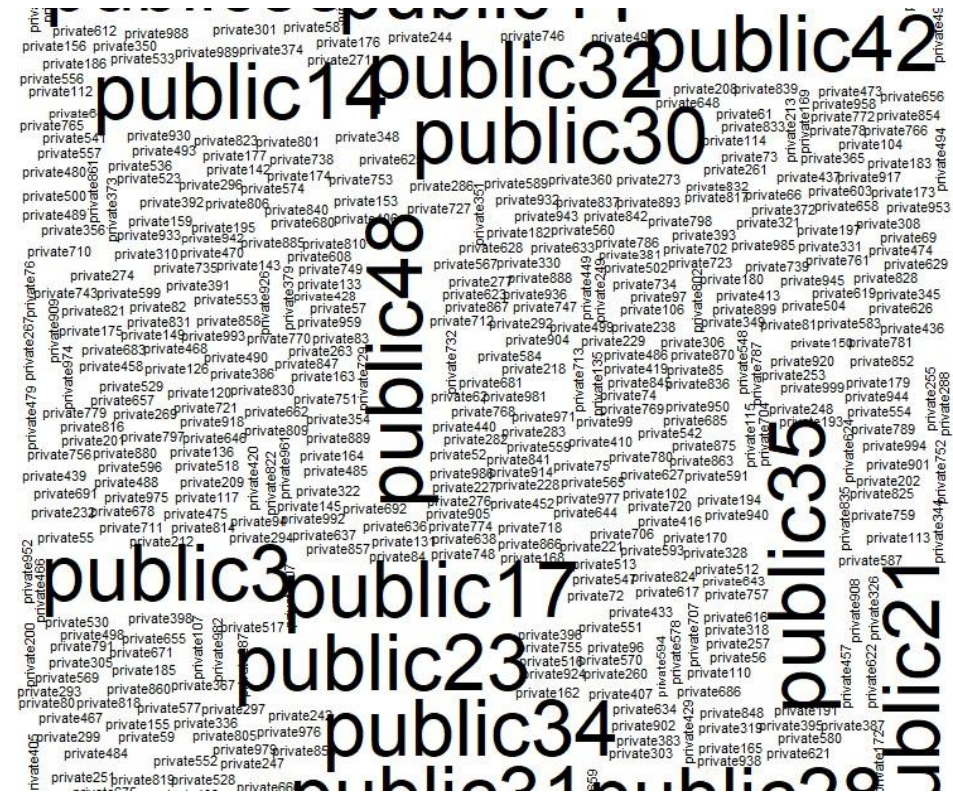# Experimental Setting – "Wet Part"

The Phage Display Library Selection

Each antigen can be represented approximately by a set of k-mers (7mers?) like a word by sylables

Large Pools of Antibodies Emphasize Public Reactivties

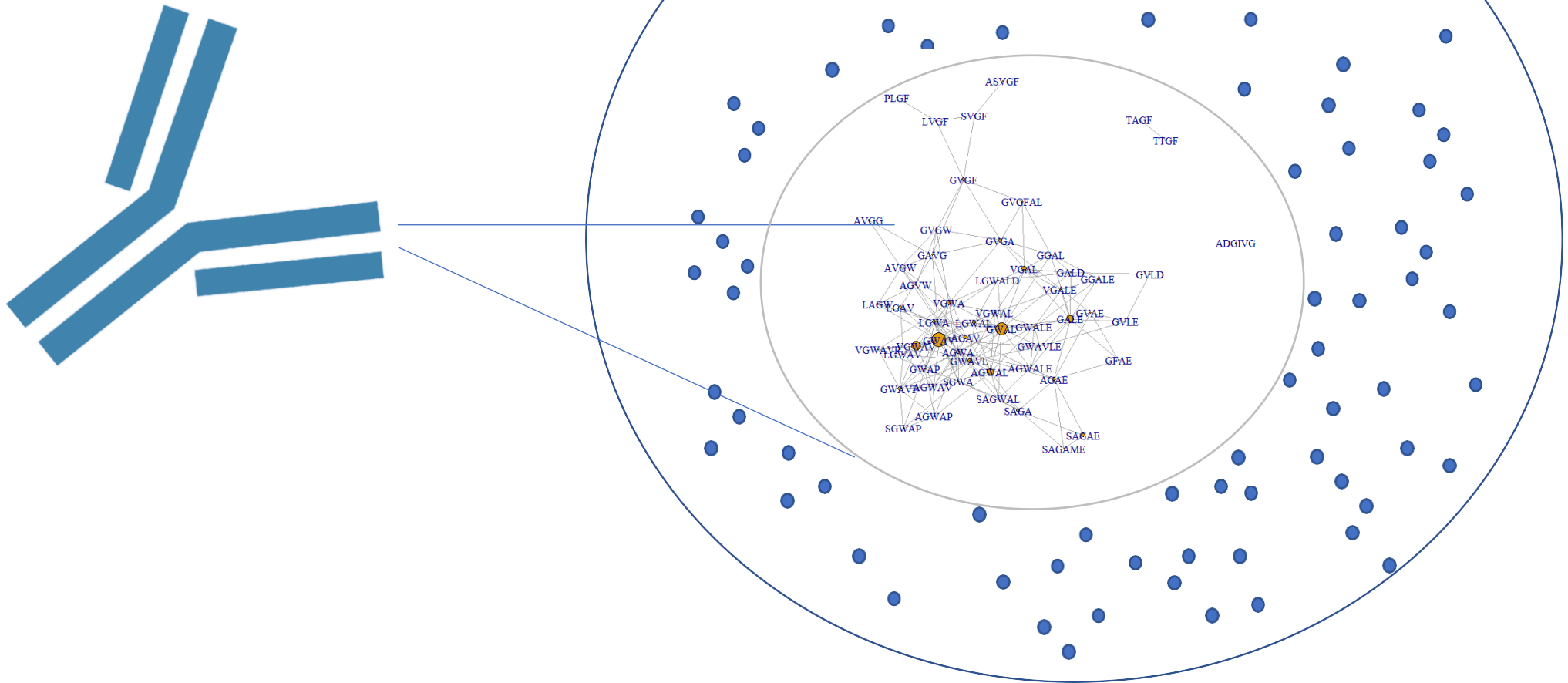Mixture of Serum form 20 Individuals

Mixture of Serum from 10000 Individuals

|	CKIHFFR	AKIHFPP	GKIHFFR	HKIHFLP	NKIHFYP	DKIHFYP	SKIHFPE	AKIHFLP	GKIHFHP	SKIHFLS	SKIHFHP	SKIHFYH	RKLHFLP	EKPHFLV	GYKPTFS	RYYPYFP
CKIHFFR	0	KIHF	KIHFFR	KIHF	KIHF	KIHF	KIHF	KIHF	KIHF	KIHF	KIHF	KIHF	0	0	0	0
AKIHFPP	KIHF	0	KIHF	KIHFP	KIHFP	KIHFP	KIHFP	AKIHFP	KIHFP	KIHF	KIHFP	KIHF	KHFP	0	0	0
GKIHFFR	KIHFFR	KIHF	0	KIHF	KIHF	KIHF	KIHF	KIHF	GKIHF	KIHF	KIHF	KIHF	0	0	0	0
HKIHFLP	KIHF	KIHFP	KIHF	0	KIHFP	KIHFP	KIHFP	KIHFLP	KIHFP	KIHFL	KIHFP	KIHF	KHFLP	KHFL	0	0
NKIHFYP	KIHF	KIHFP	KIHF	KIHFP	0	KIHFYP	KIHFP	KIHFP	KIHFP	KIHF	KIHFP	KIHFY	KHFP	0	0	0
DKIHFYP	KIHF	KIHFP	KIHF	KIHFP	KIHFYP	0	KIHFP	KIHFP	KIHFP	KIHF	KIHFP	KIHFY	KHFP	0	0	0
SKIHFPE	KIHF	KIHFP	KIHF	KIHFP	KIHFP	KIHFP	0	KIHFP	KIHFP	SKIHF	SKIHFP	SKIHF	KHFP	0	0	0
AKIHFLP	KIHF	AKIHFP	KIHF	KIHFLP	KIHFP	KIHFP	KIHFP	0	KIHFP	KIHFL	KIHFP	KIHF	KHFLP	KHFL	0	0
GKIHFHP	KIHF	KIHFP	GKIHF	KIHFP	KIHFP	KIHFP	KIHFP	KIHFP	0	KIHF	KIHFHP	KIHFH	KHFP	0	0	0
SKIHFLS	KIHF	KIHF	KIHF	KIHFL	KIHF	KIHF	SKIHF	KIHFL	KIHF	0	SKIHF	SKIHF	KHFL	KHFL	0	0
SKIHFHP	KIHF	KIHFP	KIHF	KIHFP	KIHFP	KIHFP	SKIHFP	KIHFP	KIHFHP	SKIHF	0	SKIHFH	KHFP	0	0	0
SKIHFYH	KIHF	KIHF	KIHF	KIHF	KIHFY	KIHFY	SKIHF	KIHF	KIHFH	SKIHF	SKIHFH	0	0	0	0	0
RKLHFLP	0	KHFP	0	KHFLP	KHFP	KHFP	KHFP	KHFLP	KHFP	KHFL	KHFP	0	0	KHFL	0	0
EKPHFLV	0	0	0	KHFL	0	0	0	KHFL	0	KHFL	0	0	KHFL	0	0	0
GYKPTFS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RYYPYFP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Hypothetical image of an antibody's reactivity space in the mimotope space

Mimotope Space

Selecting the right string metric

Classification of known "isospecific" mimotope pairs compared to random peptide pairs using different metrics:

- ham   Hamming distance,
- lcs   Longest common substring distance.
- lv   Levenshtein distance,
- dl   Damerau-Levenshtein distance,
- osa   Optimal string aligment,
- qgr   q-gram distance,
- cos   Cosine distance between q-gram profiles,
- jac   Jaccard distance between q-gram profiles,
- pwa   Pair-wise alignment (AA substitutin matrix based),
- ..._h   Versions of editing distances
          with recoded AA based on the Phys\Chem properties

ROC curves
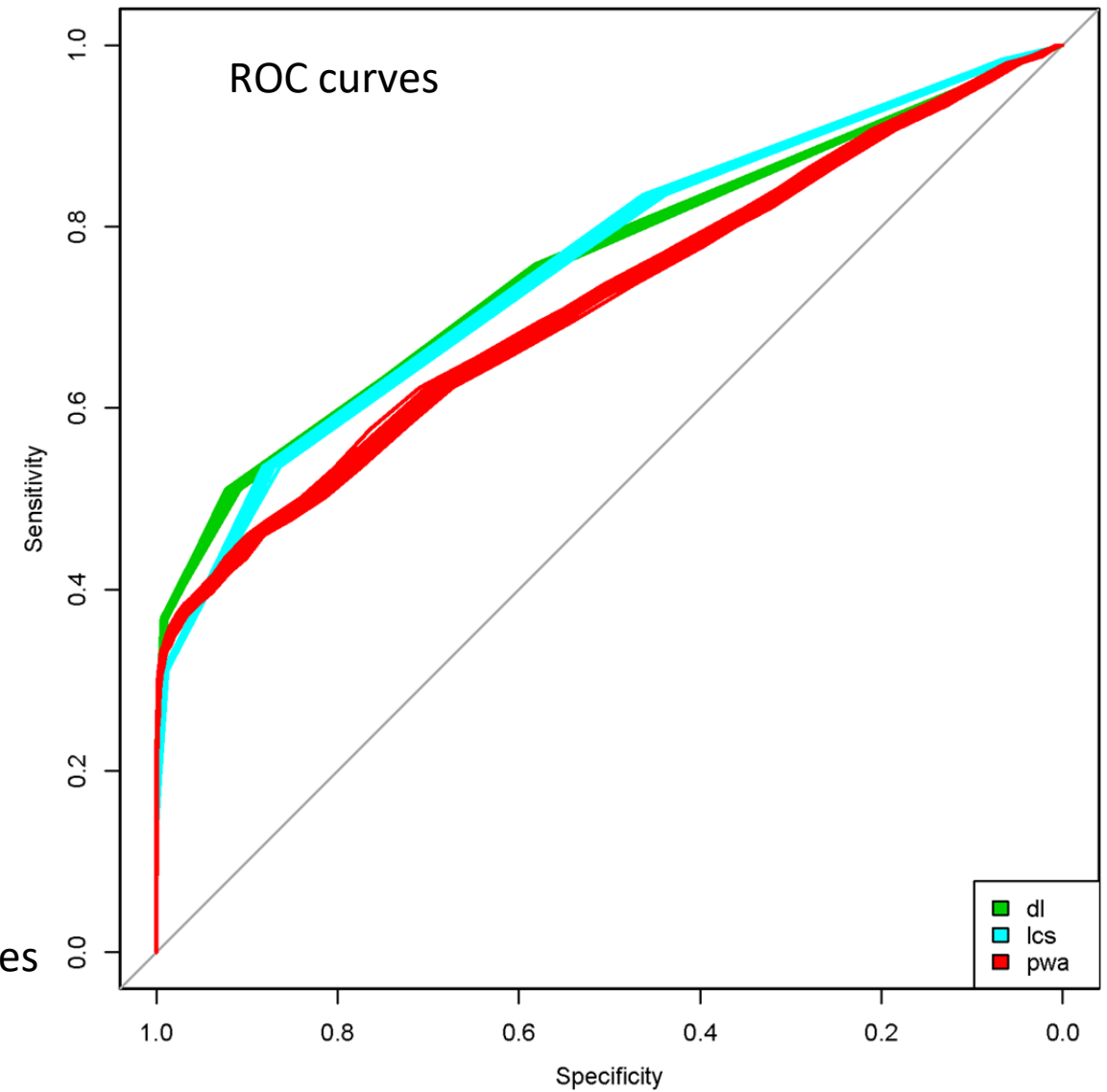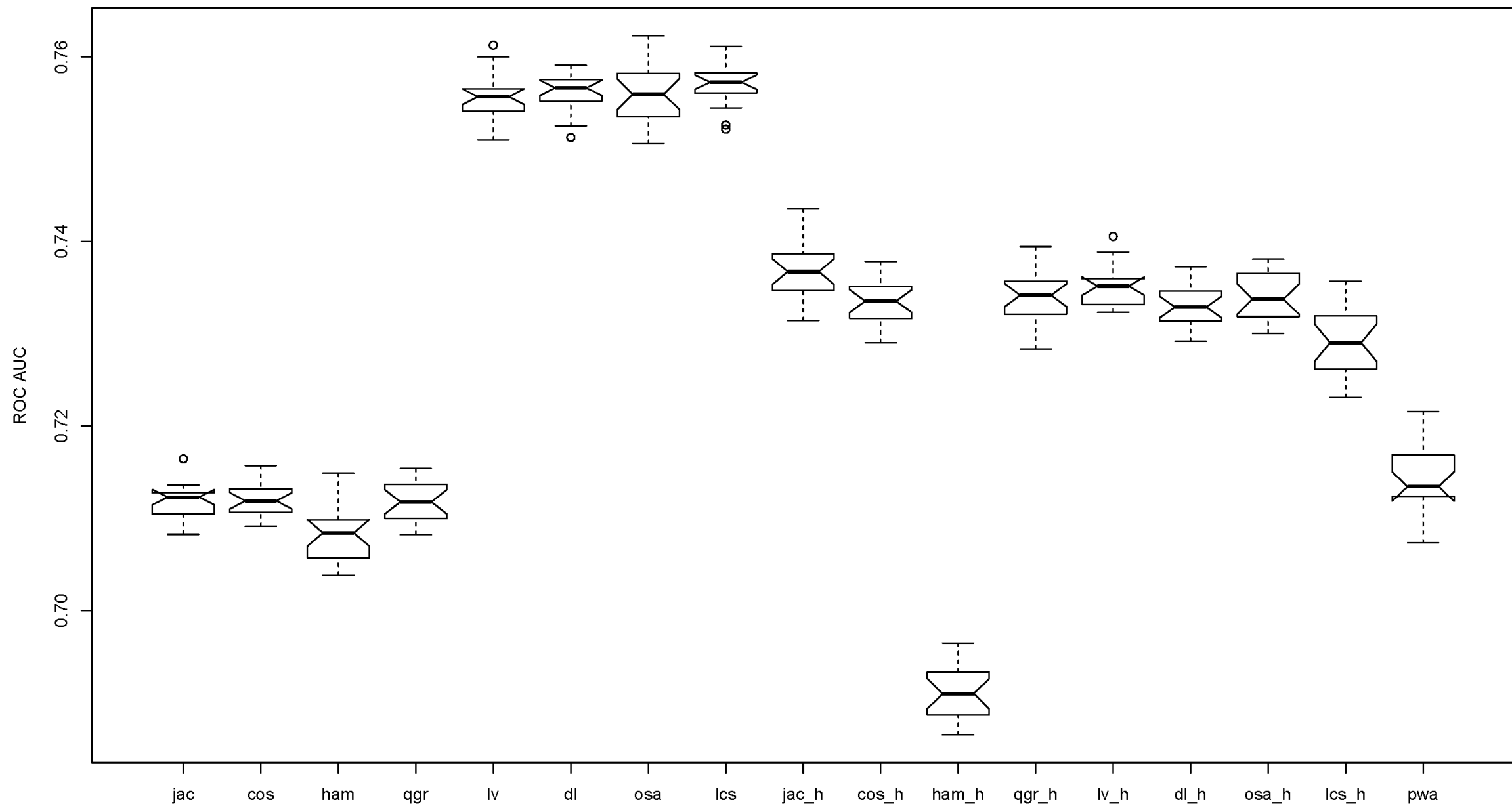
Building the Graph

Longest common subsequence

A	P	V	N	T	L	G
A	P	V	N	L	T	L
A	V	F	V	N	T	L

Conservative threshold - d < 5
(2 differences**) -** high specificity but
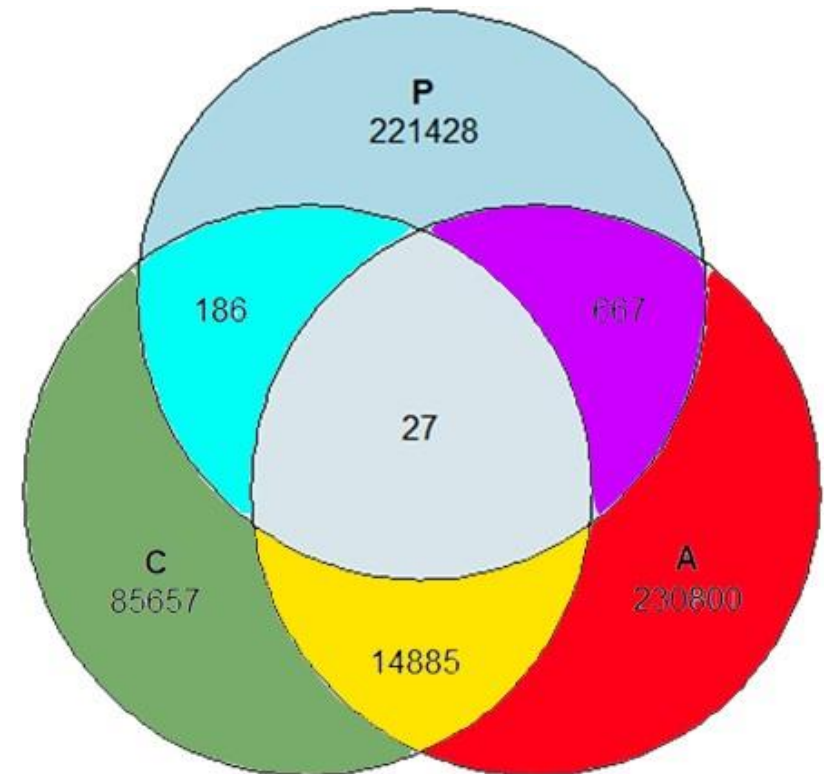lower sensitivity

Example
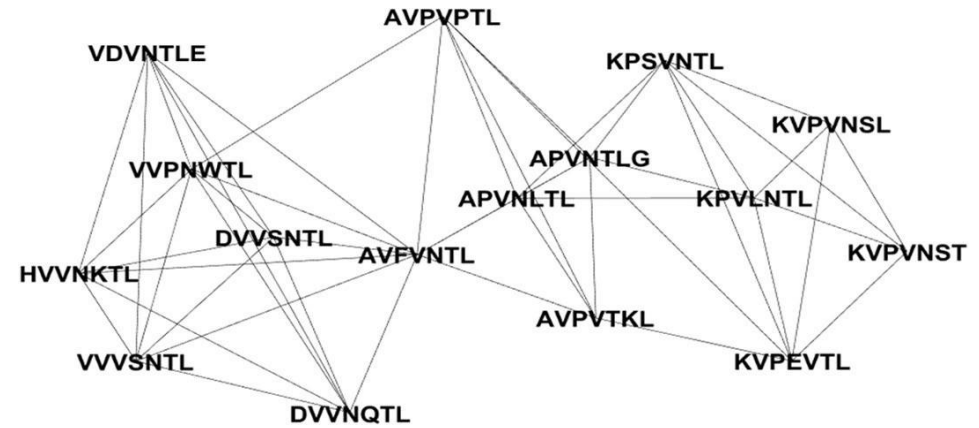
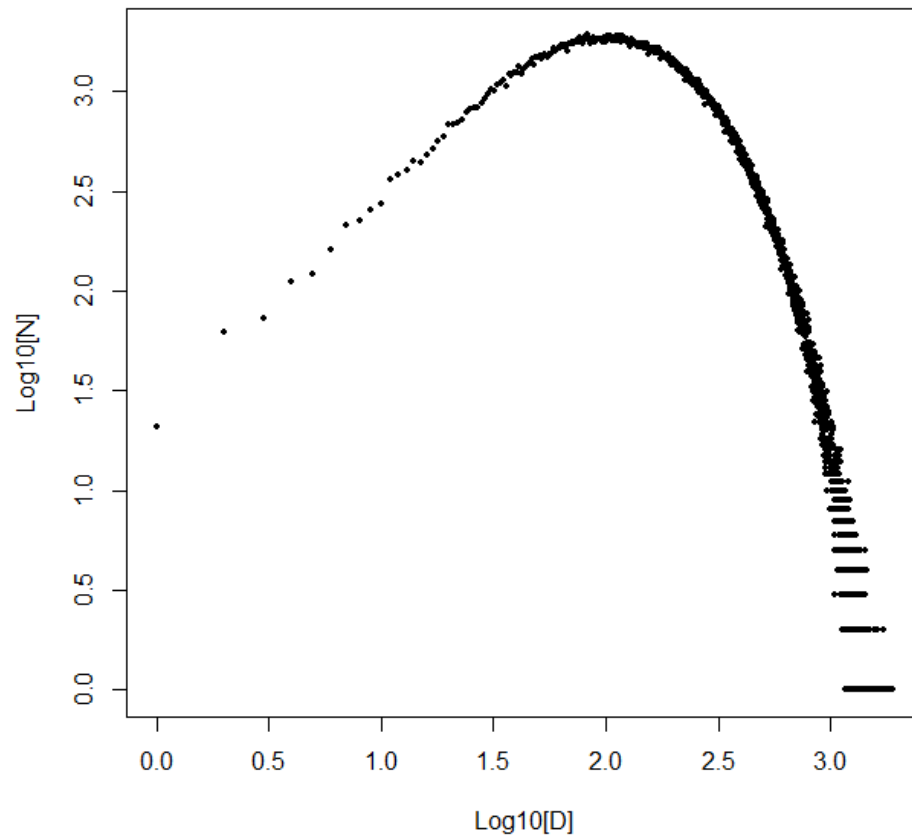## The Data - APLS Study

Sera from:

C(ontrol) – Healthy women, age matched (n=20),
A(PLS) – Women with antiphospholipid autoantibodies
        and habitual abortions (n=24),
P(ublic) – IgM from IVIgM (Biotest, Dreireich) –
        global repertoire of 10000 healthy donors.

The composition of the
mimotope libraries

The Graph

553 634 vertices and 65 873 585 edges

Distribution of mimotope classes
in the neighborhood of each vertex

| Sequence	C	A	C+A	P	P+C	P+A	C+A+P
AAHPAPK	29	155	9	76	0	0	0
AAHPRQT	38	163	11	114	0	0	0
AAHQLRL	47	184	16	121	0	1	0
AAHSLRL	74	303	24	189	0	4	0
AAHSRIL	29	190	17	139	0	2	0
AAHSTSD	40	180	13	120	0	3	0
AAHSYPA	34	188	15	115	1	0	0
...	...	...	...	...	...	...	...

+ mapping of
4×10^6 IgJ
region 7-mers

The Volcano Plot

Comparing mimotopes selected by
the **IgM repertoires of APLS and healthy women**,
the neighborhoods in the graph
populated predominantly
by mimotopes of the healthy
far exceeded those dominated by
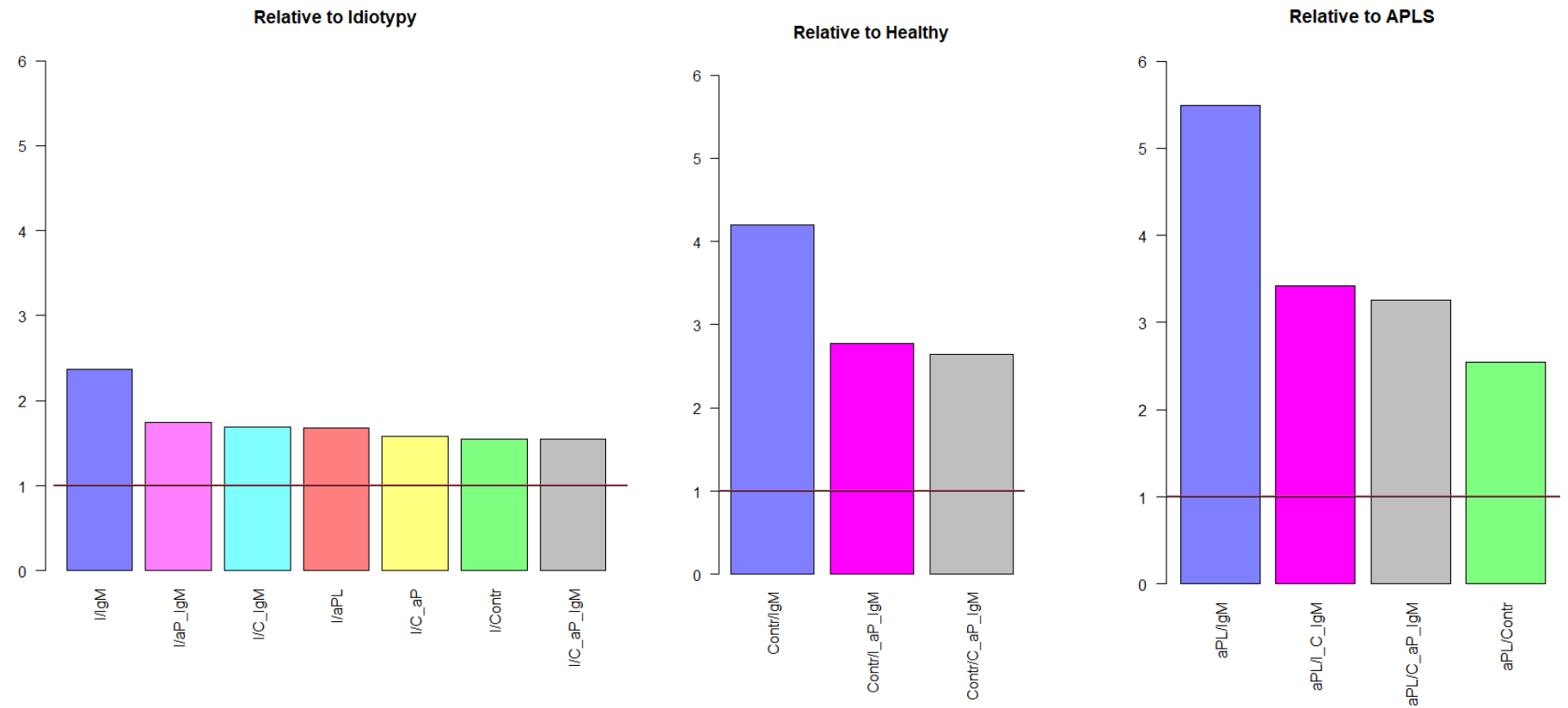mimotopes of the APLS repertoire.
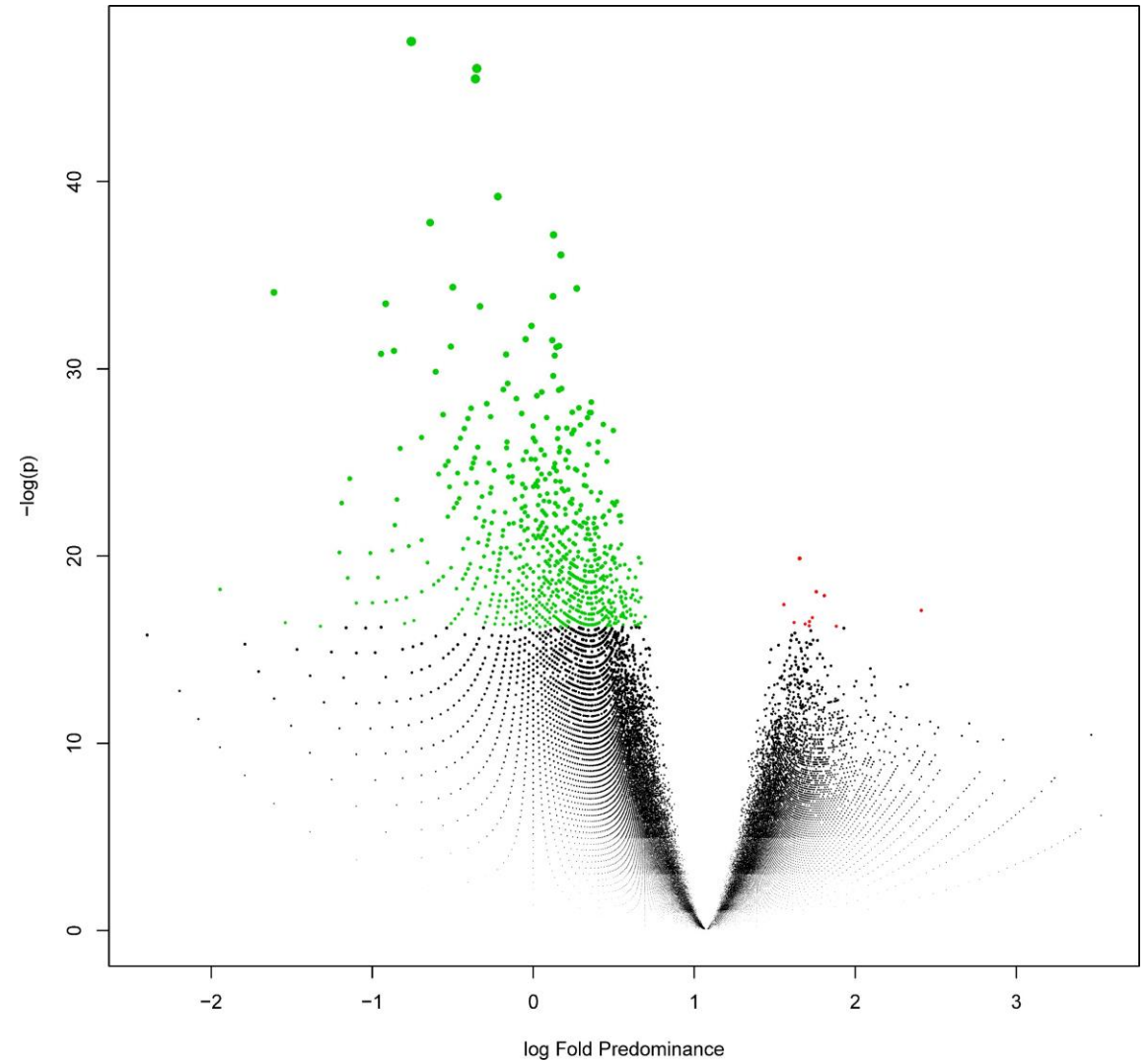
Profiles of the calls
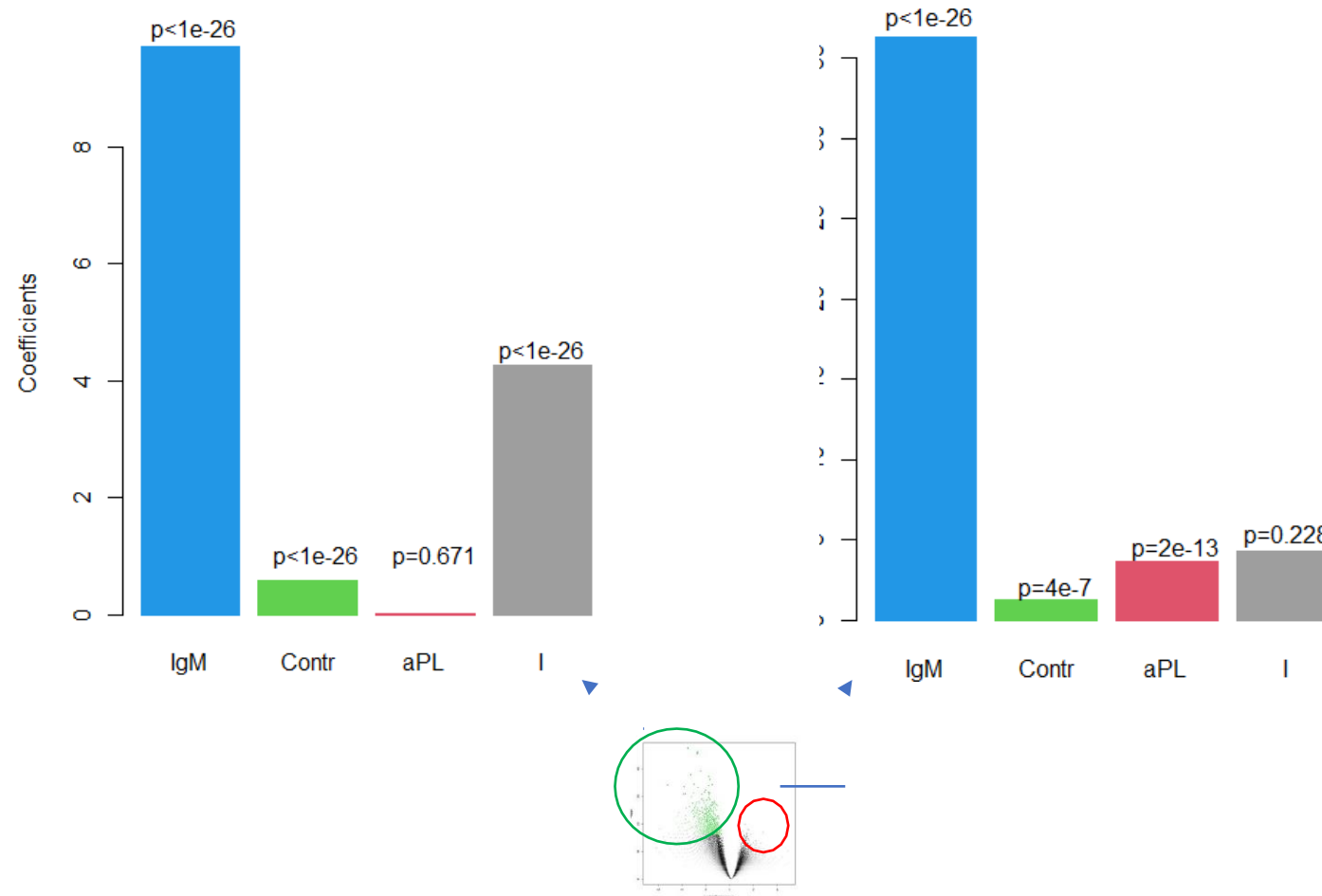
Colocalization between reactivities in different repertoires

The neighborhood distributions were dichotomized at the medians of each class and 2x2 tables were constructed for each two-way comparison.
The 2x2 tables were used to calculate odds ratios.

Odds Ratios

CONCLUSION

Autoimmune antiphospholipid syndrome is characterized by pathological IgG auto-antibodies but the main change in the IgM repertoire is

loss of specificities which are public and idiotypically connected.

Thanks to:

Lubomir Balabansky, Doctor Malinov Clinic, Ltd, Sofia
Gabriel Elmadjiyan, IBIR, BAS
Dr. Velizar Shivarov, Sofiamed Hospital, Sofia