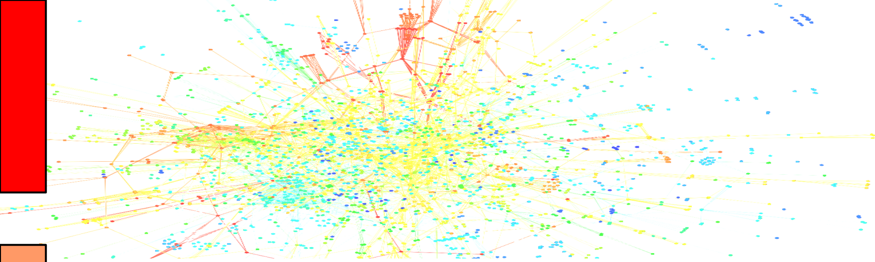


Close-to-the-experiment Data analysis

Dr. Werner Van Belle

werner.van.belle@gmail.com, werner@onlinux.be

Bridging The Gap

- Core of most biological research is experimental

Experiments

Data Reduction

Interpretation

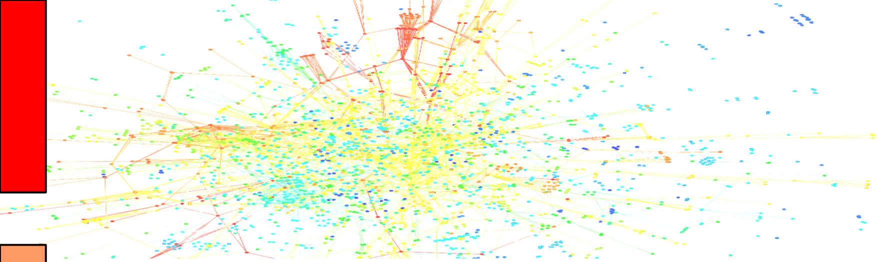
Bridging The Gap

- Core of most biological research is experimental

Experiments

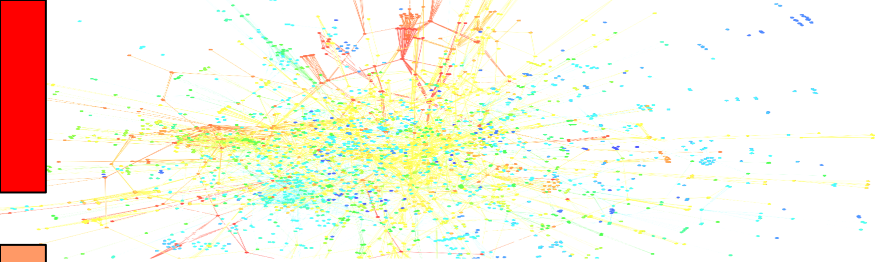
Data
Analysis

Interpretation



Contents

- Part 1. 2DE Gel Correlation Analysis
- Part 2. Maldi Tof Artefacts / Denoising
- Part 3. Accuracy Analysis of a Micro-Array Experiment
- Part 4. Protein Interaction Map Integration



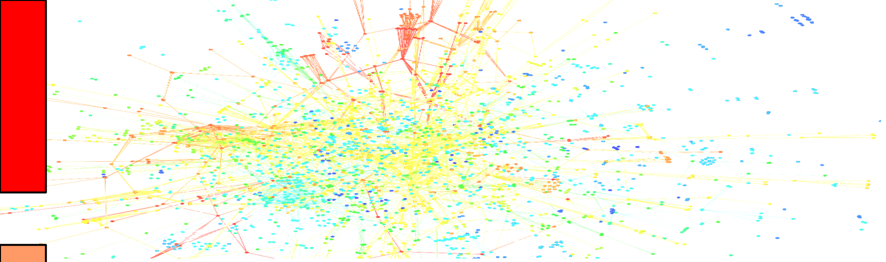
Part 1. 2DE Gel Analysis

Werner Van Belle

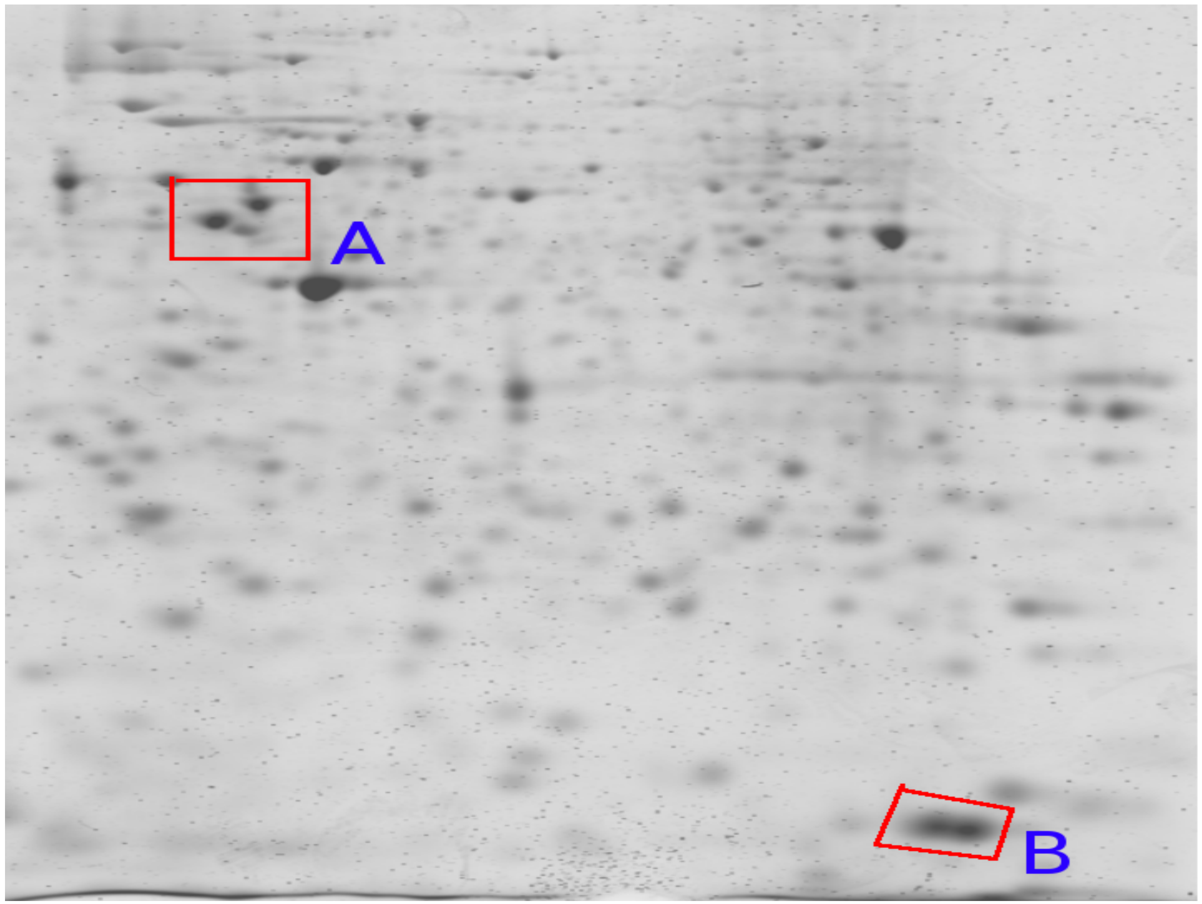
werner.van.belle@gmail.com, werner@onlinux.be

In cooperation with: Bjørn Tore Gjertsen, Nina Ånensen

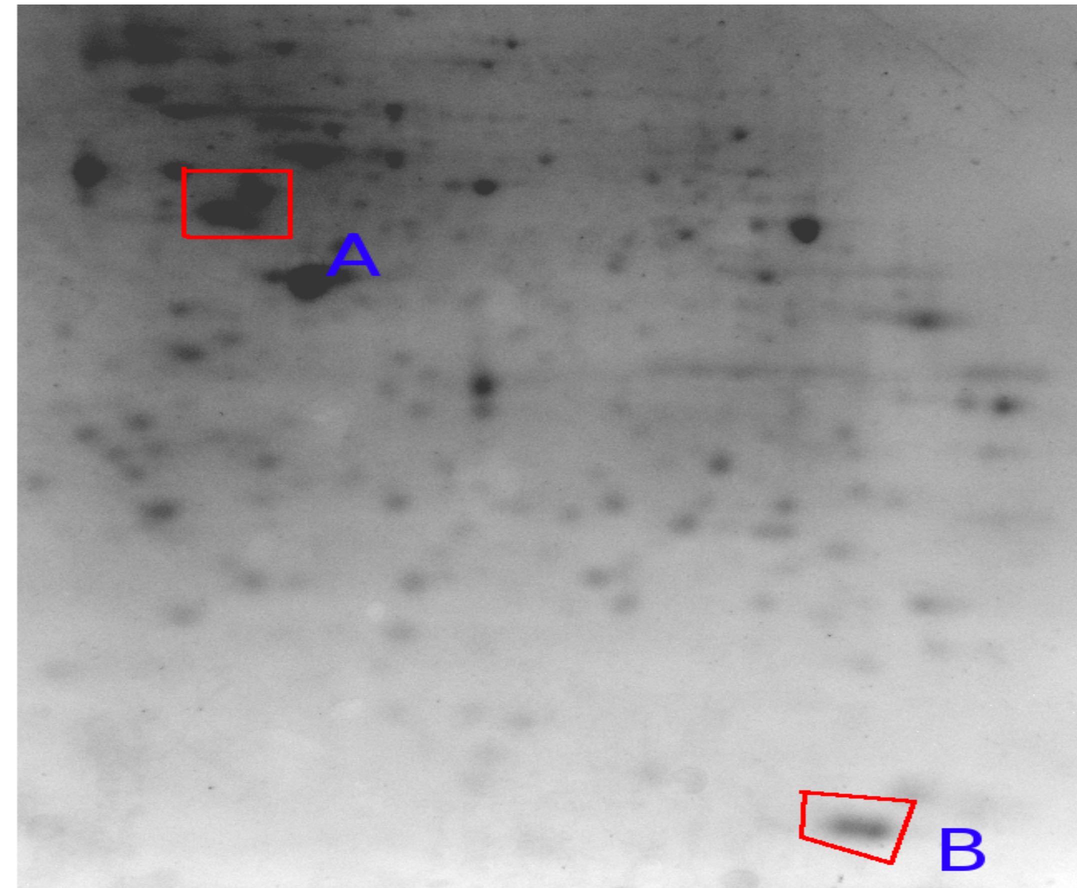
Ingvild Haaland, Gry Sjøholt, Kjell-Arild Høgda



2D Gels



Patient #1
Age: 57



Patient #2
Age: 46

Courtesy Gry Sjøholt, Nina Ånensen & Bjørn Tore Gjertsen

Initial Problem

- The question we were asked
  - Is there a relation between various parameters of AML/ALL cancer patients and their P53 biosignatures / isoforms ?
- Gels: +/- 97 gel images of different patients
- Biological Parameters:
  - FAB Classification (AML/ALL), AML Class, Flt3 (WT/ITD)
  - Resistance AML, Resistance ALL, Survival AML, Survival ALL
  - BCL2, Stat5 GMCSF, Stat3 IL3, Stat1 Ifng, CD4, C34

Standard Solution

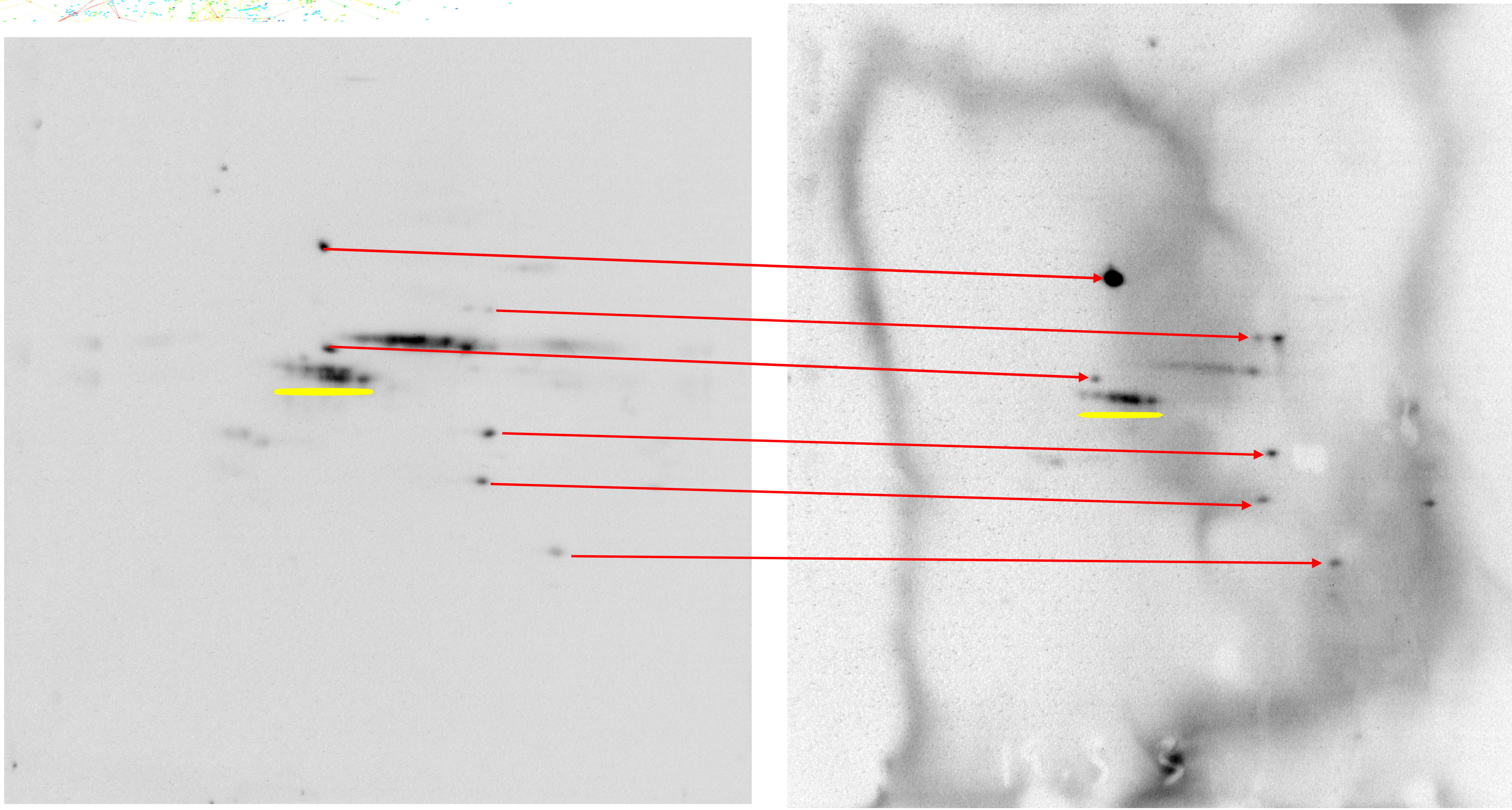
- Detect Spots, Measure Spot Volumes, Compare
- Non Trivial Solution
  - Spot identity unknown, often no calibration spots
  - Manual interpretation dangerous; shifts of spots are difficult to interpret
  - Some PTM influence spot positioning, complicating the matter

Complicated method

Tedious work

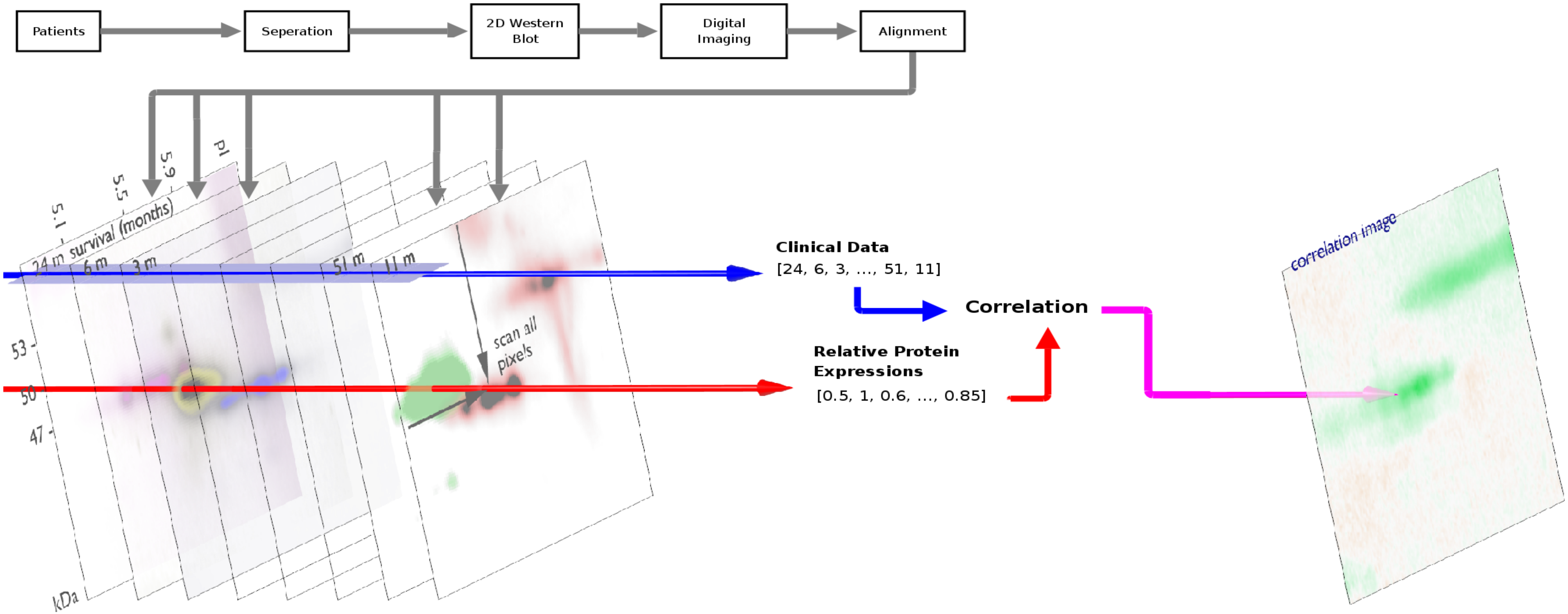
Lousy results

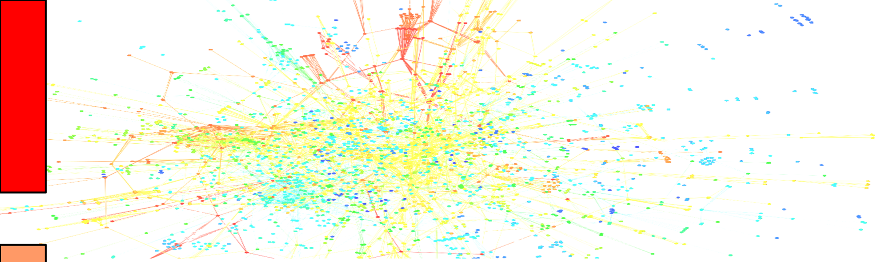
Manual Comparison



2D Gel Analysis

- Step 1: Image registration: rotate, scale & translate





Different Operations

- Scaling (Zoom)
- Rotation
- Translation

Calibration spots

Antibody spots

Manual annotation

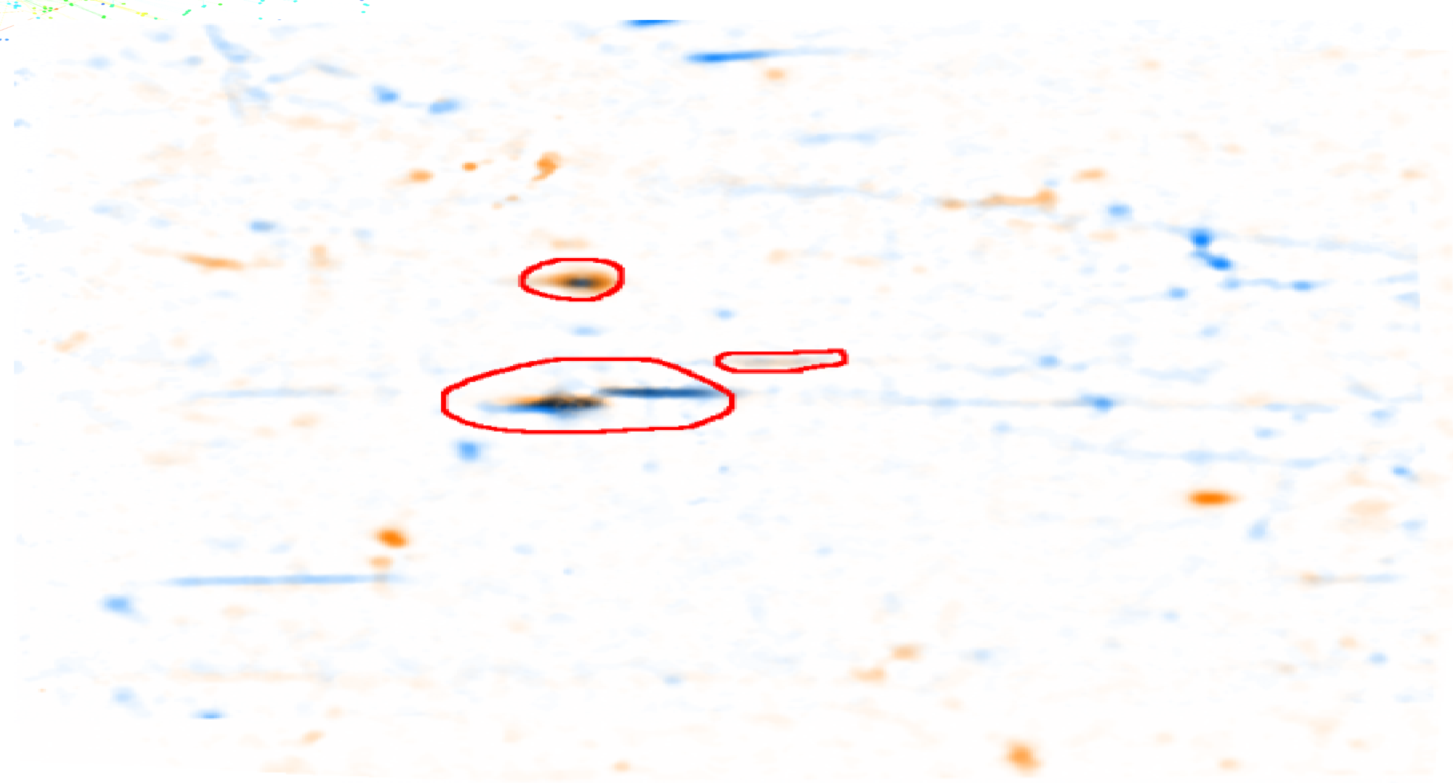
Image registration techniques

Geocoding

Landmark tracking

Standard spot detection

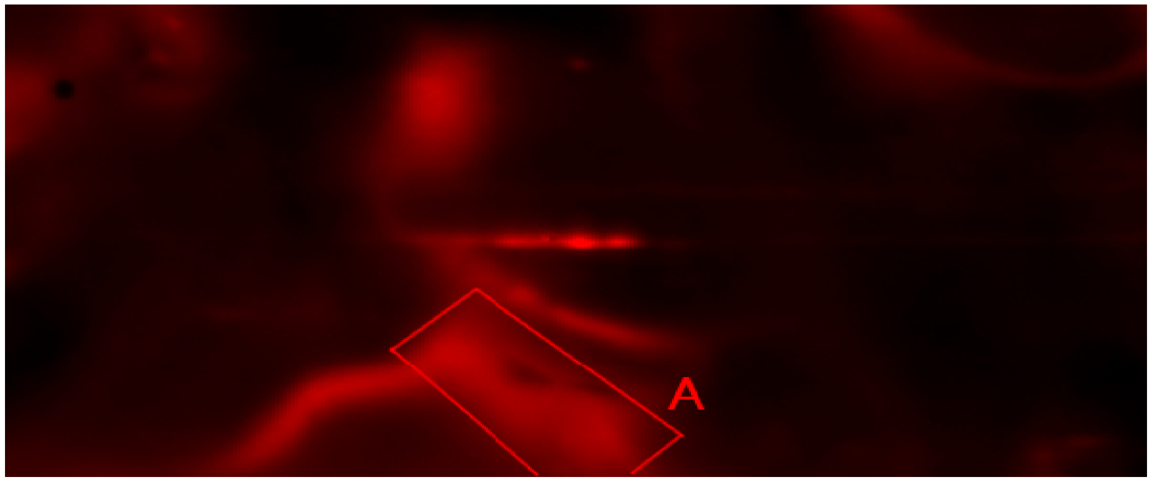
Pairwise Image Alignment

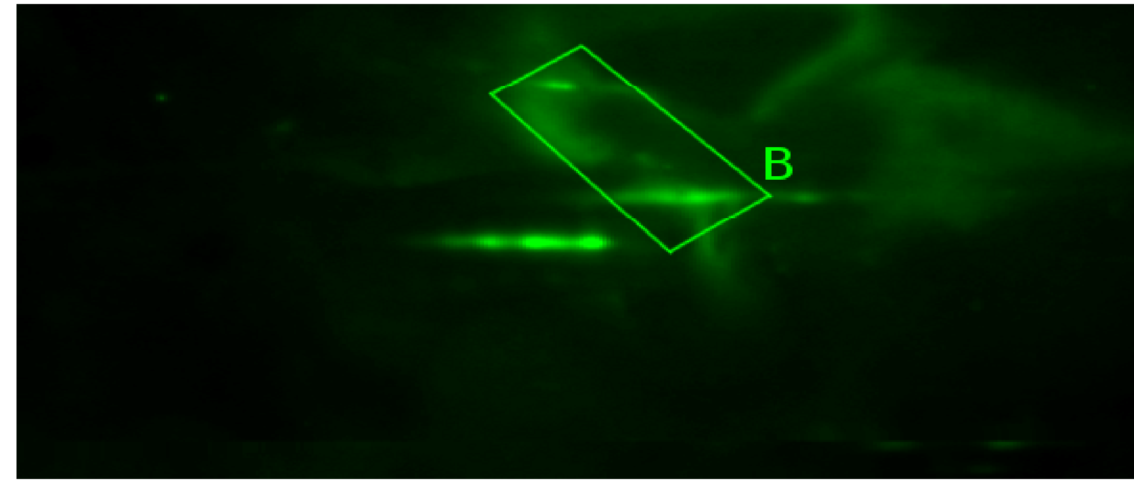


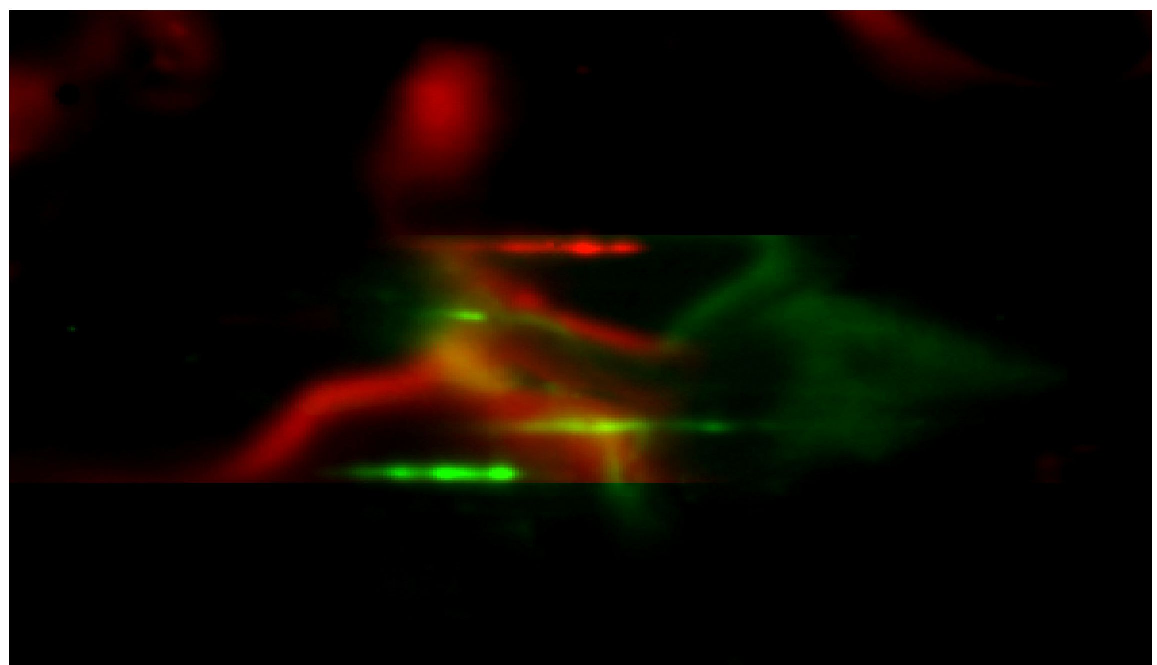
The alignment between two images M and N, is obtained by finding the translation of image N which maximizes $|M - translate(N)|^2$. This is the position in the cross correlation image with the highest correlation. The cross correlation image can be calculated fast

$$FFT^{-1}(FFT(M) \times FFT(N)^*)$$

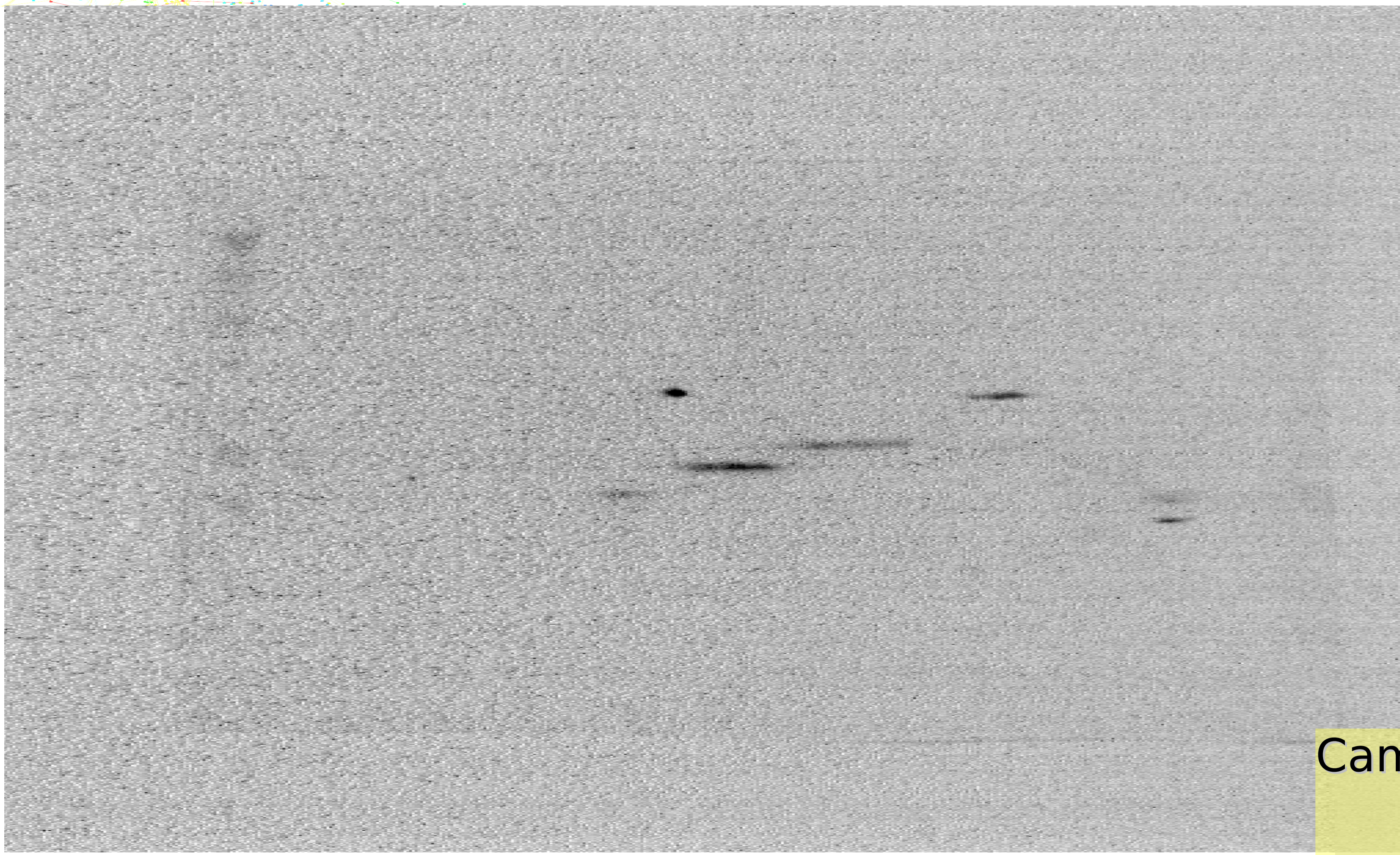
Noise hinders pairwise alignment







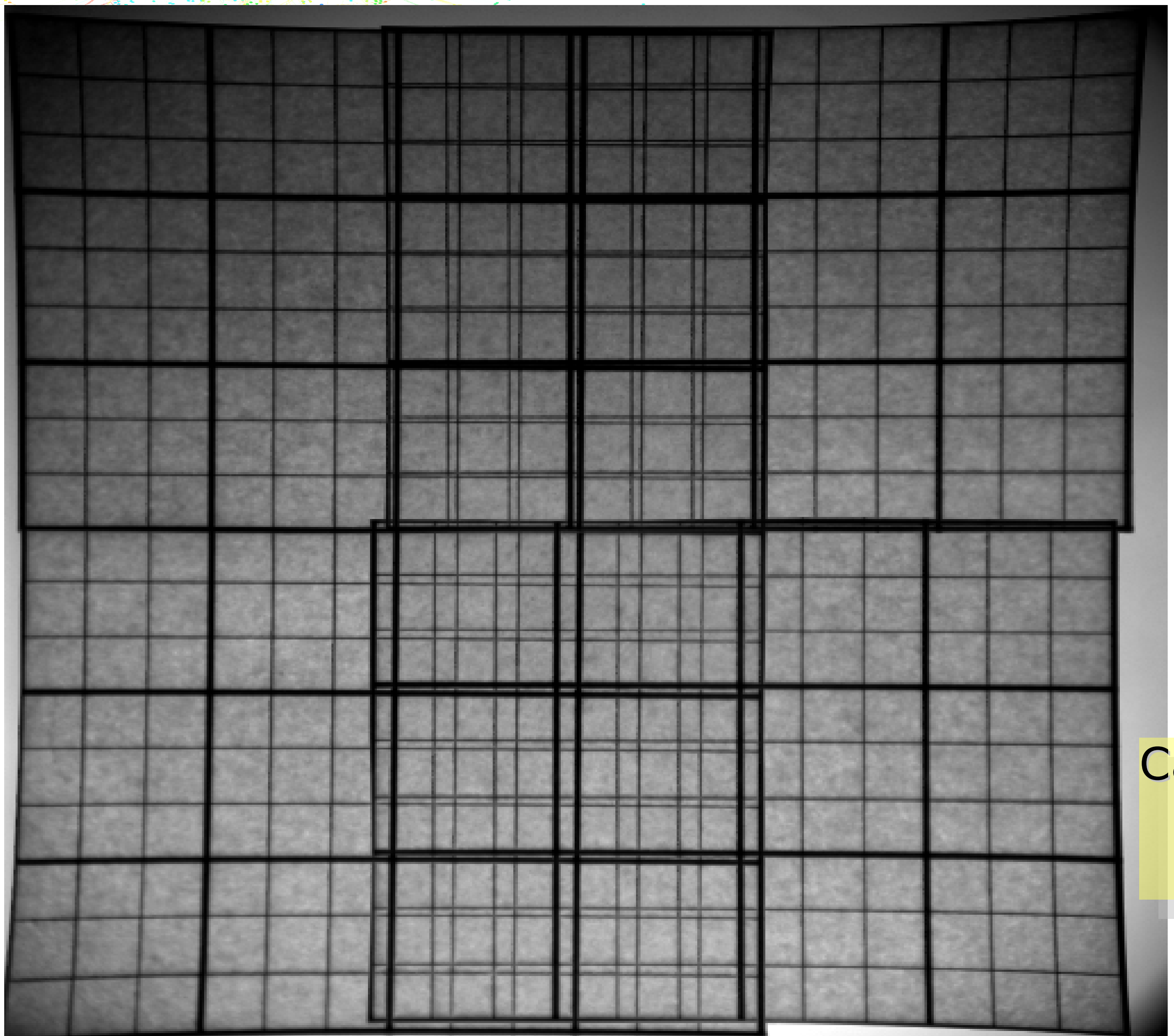
1. Artefacts in 2D gels



Camera Noise

Gels have been altered to respect NDA

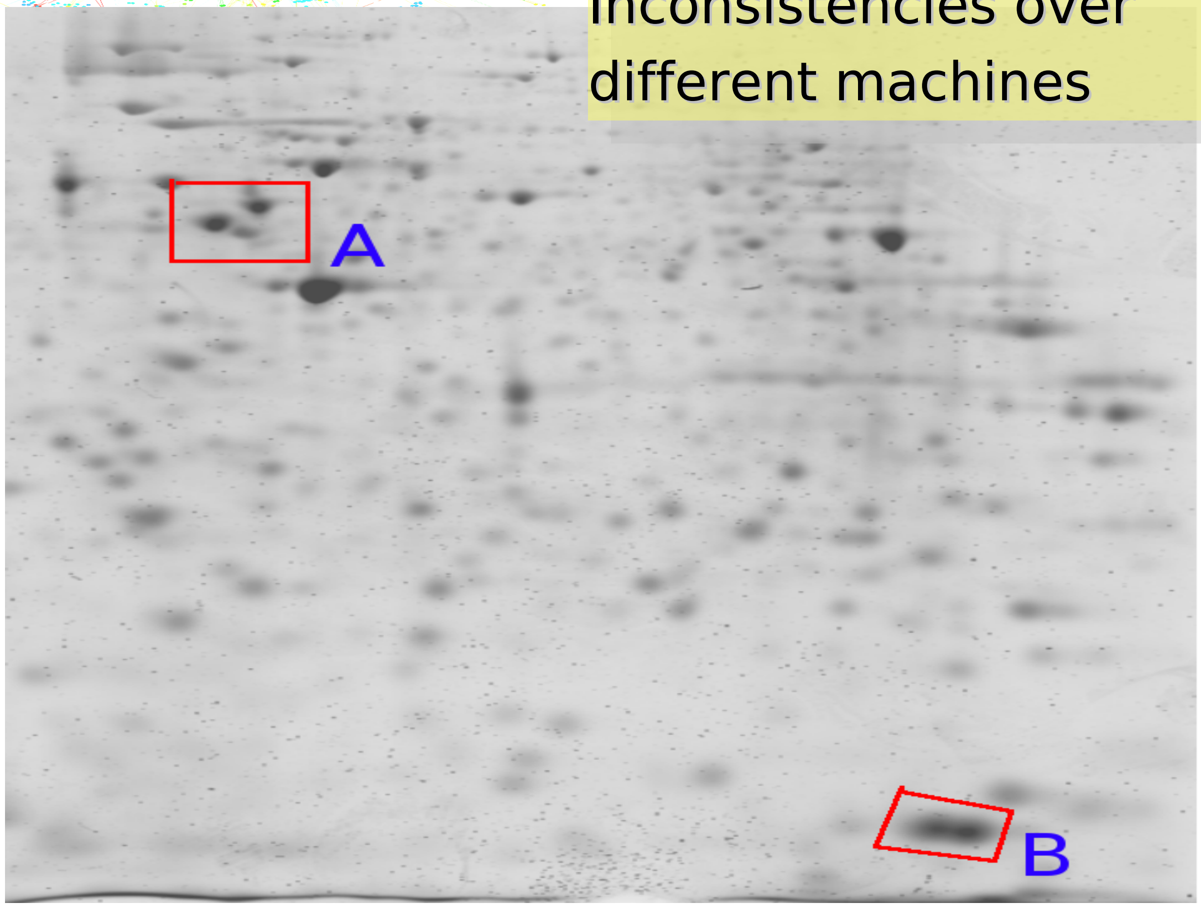
2. Artefacts in 2D Gels



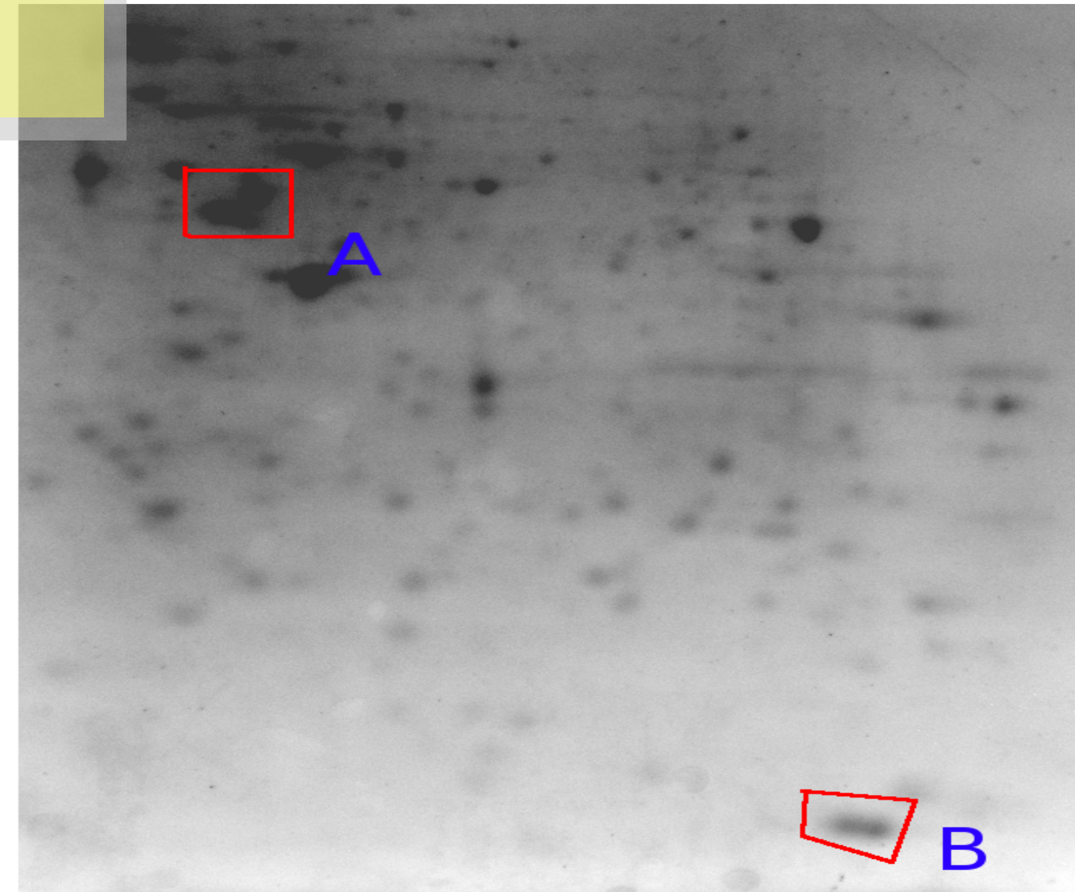
Camera Warping

3. Artefacts in 2D gels

Inconsistencies over different machines



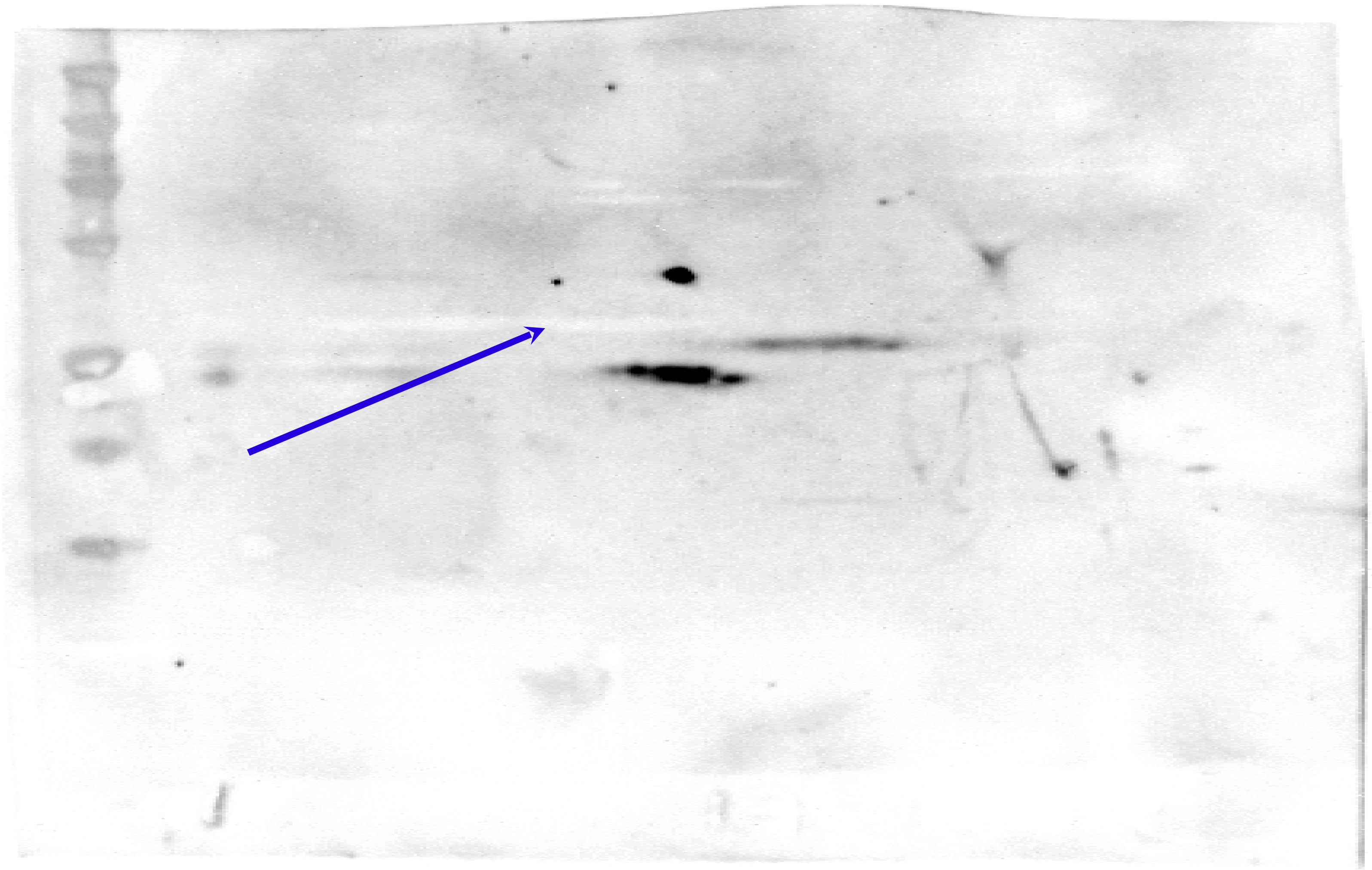
Typhoon Image Station



KODAK Image Station

Courtesy Gry Sjøholt, Nina Ånensen & Bjørn Tore

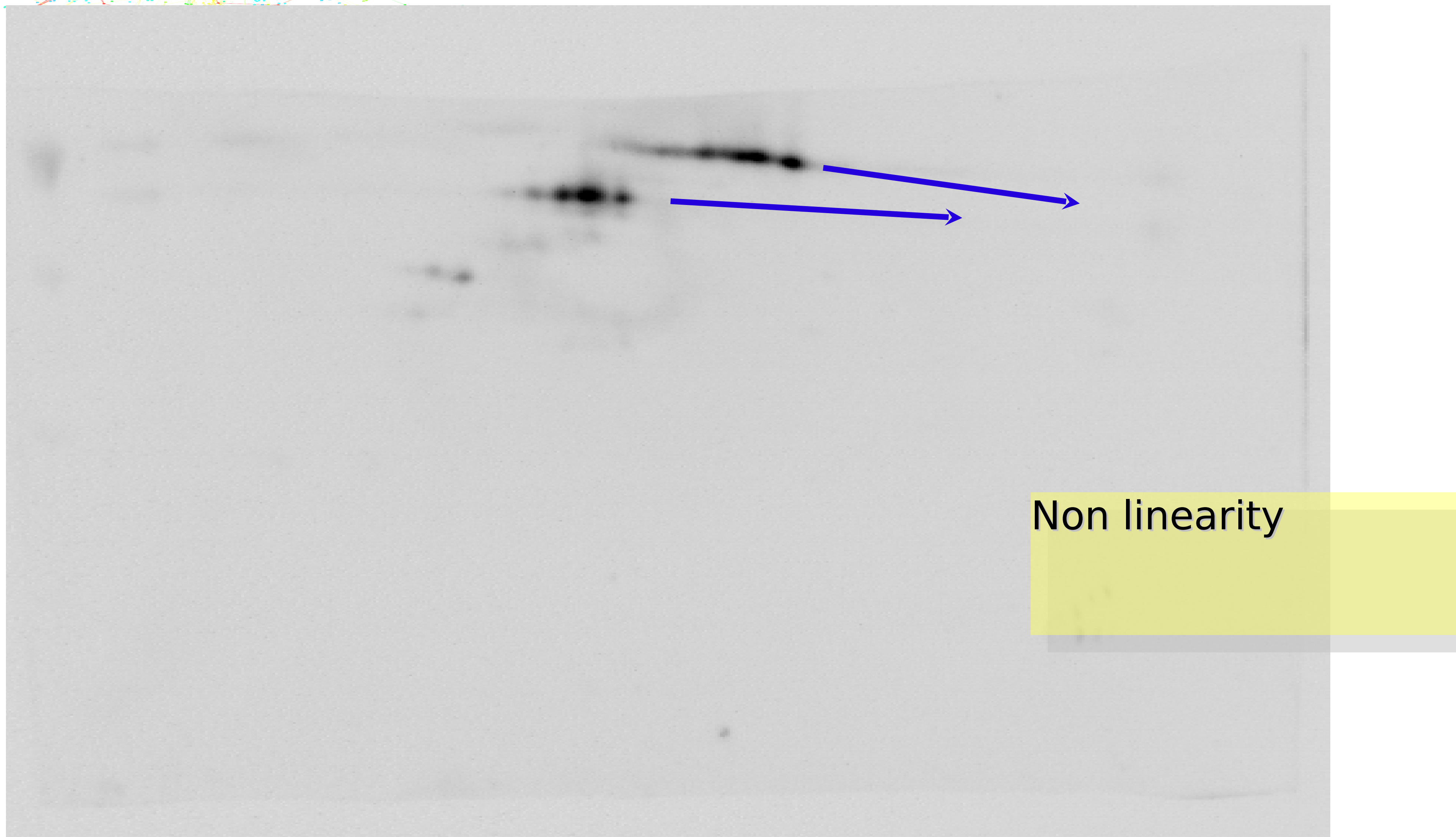
4. Artefacts in 2D gels



Underexpressed tails

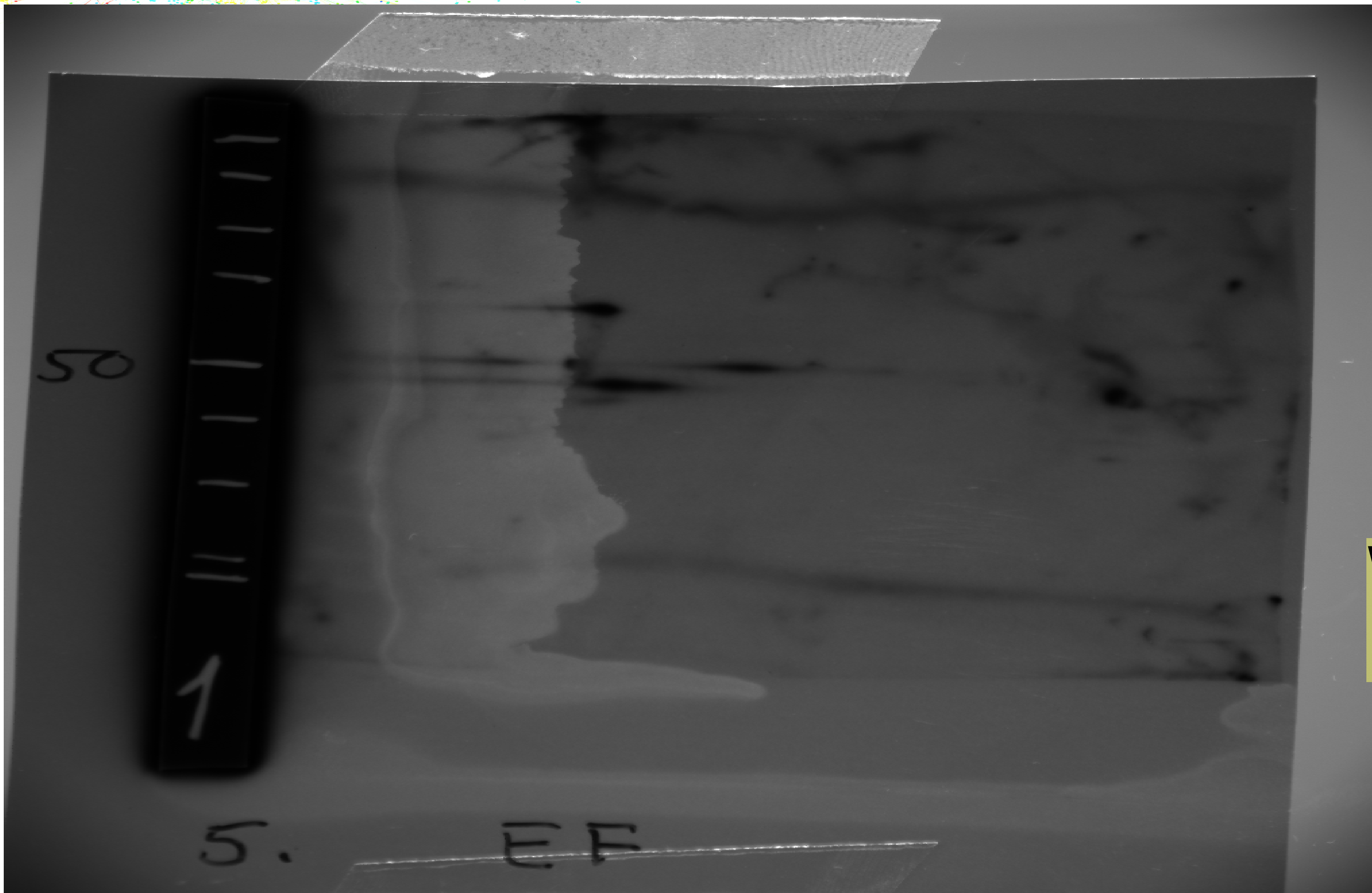
Gels have been altered to respect NDA

5. Artefacts in 2D gels



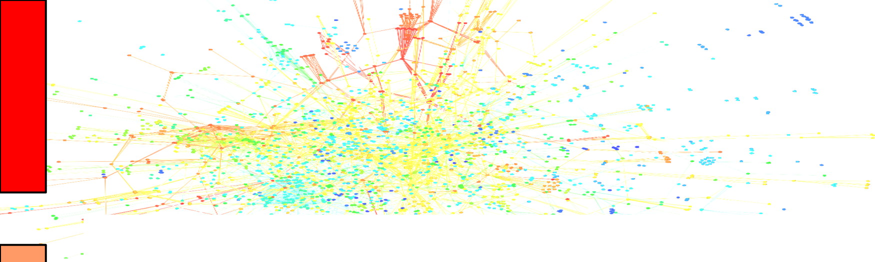
Gels have been altered to respect NDA

6. Artefacts in 2D gels

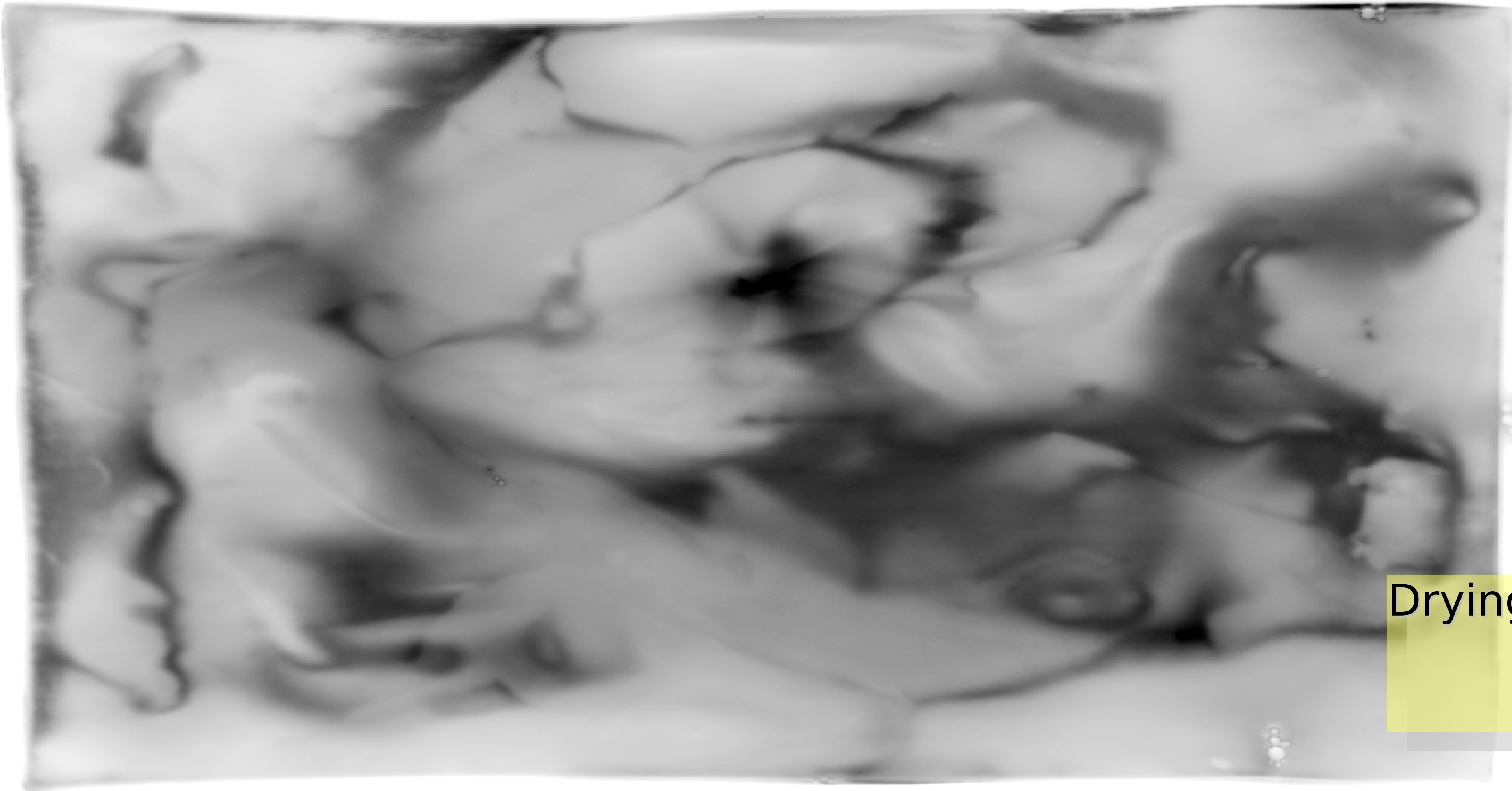


Washing

Gels have been altered to respect NDA



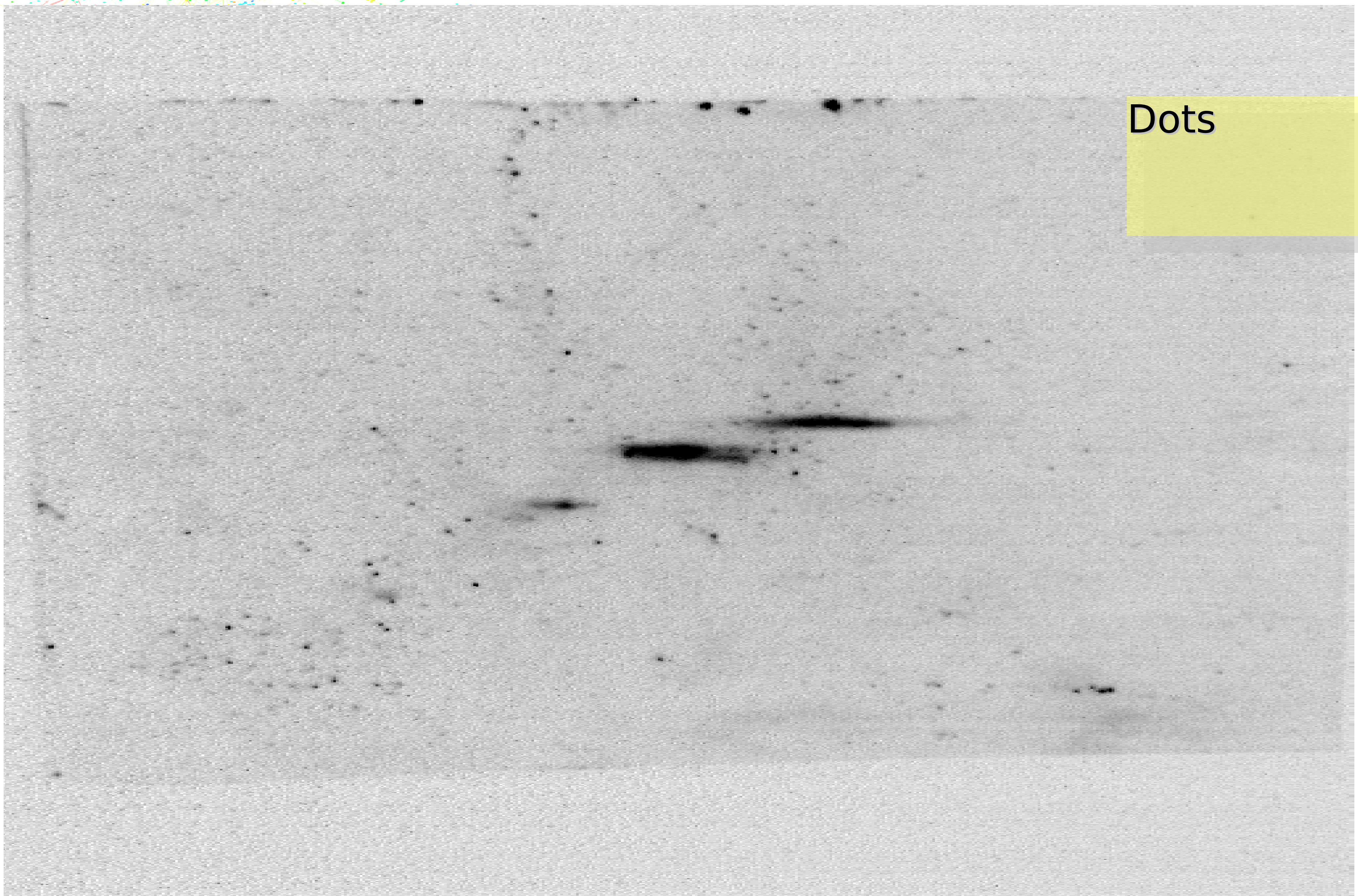
7. Artefacts in 2D gels



Drying

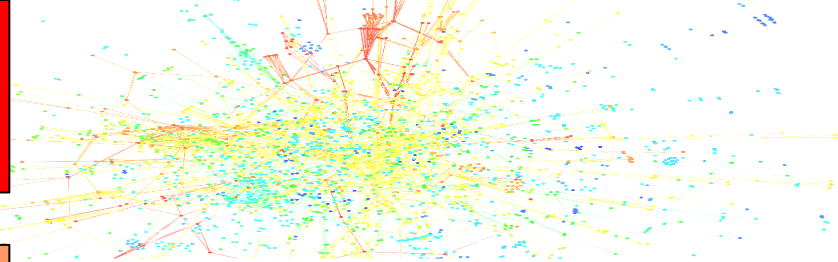
Gels have been altered to respect NDA

8. Artefacts in 2D gels

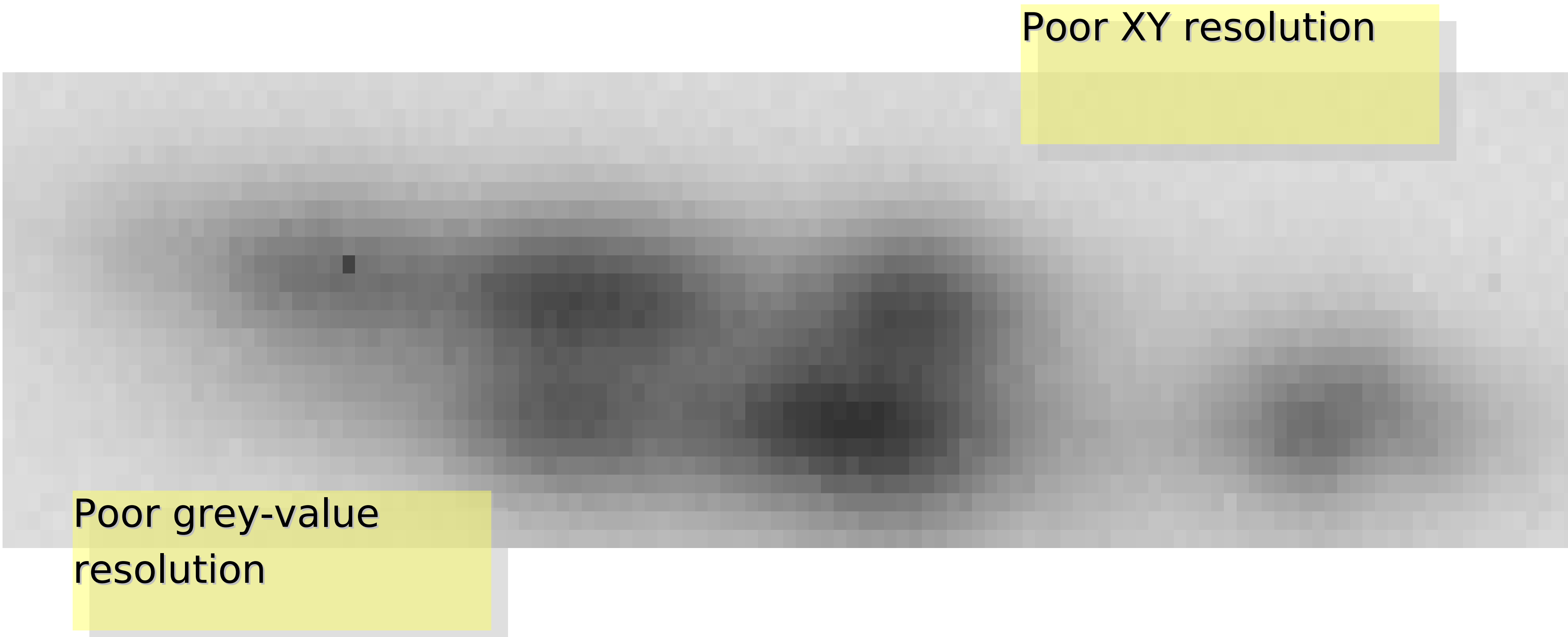


Dots

Gels have been altered to respect NDA

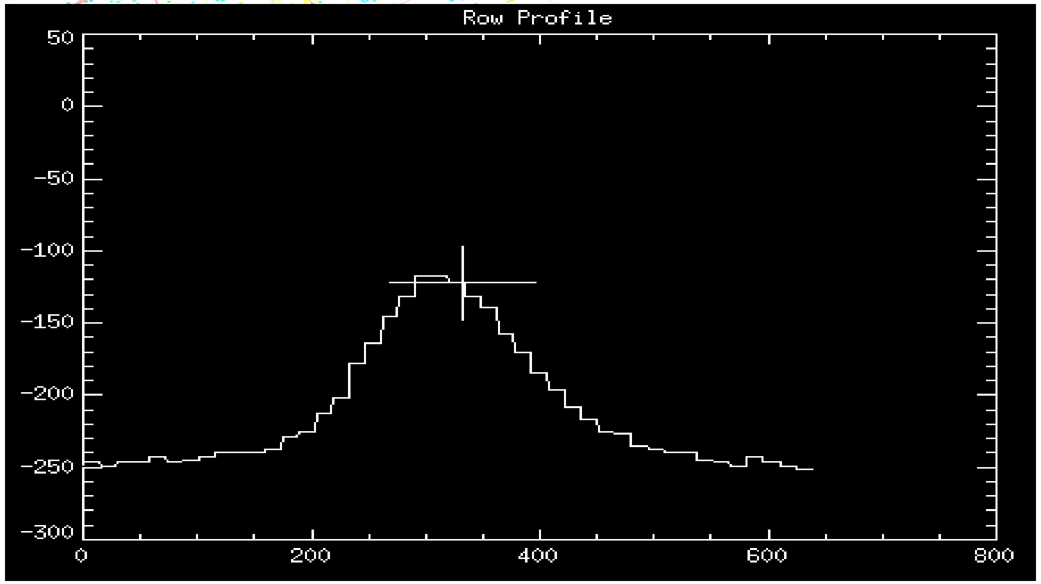


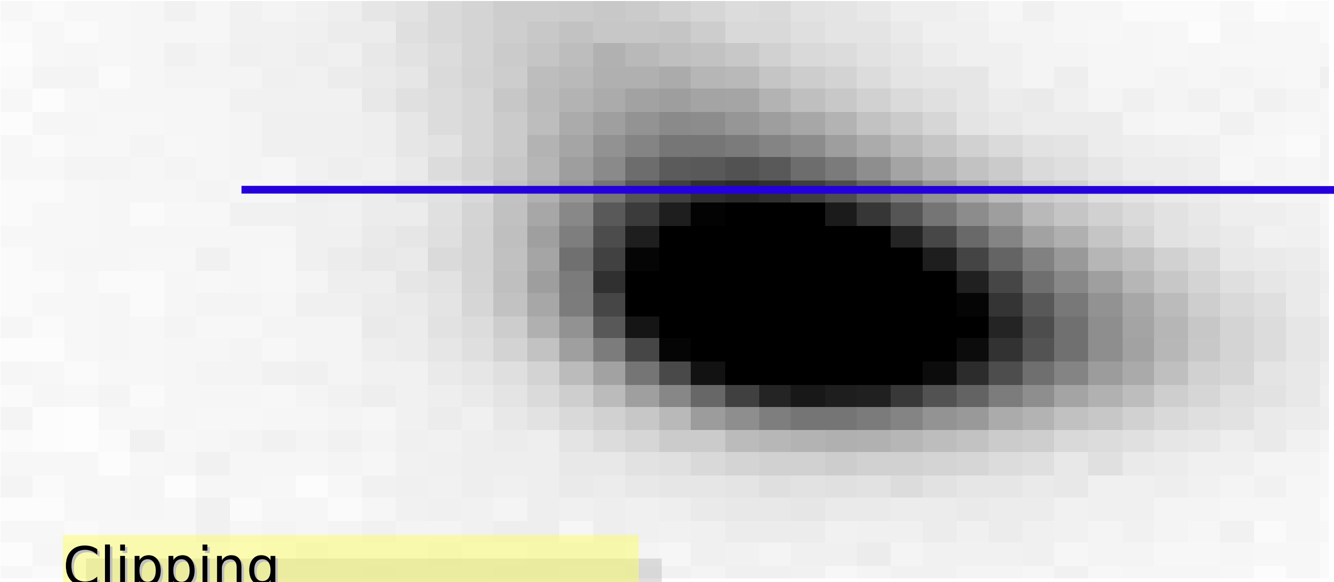
9. Artefacts in 2D gels



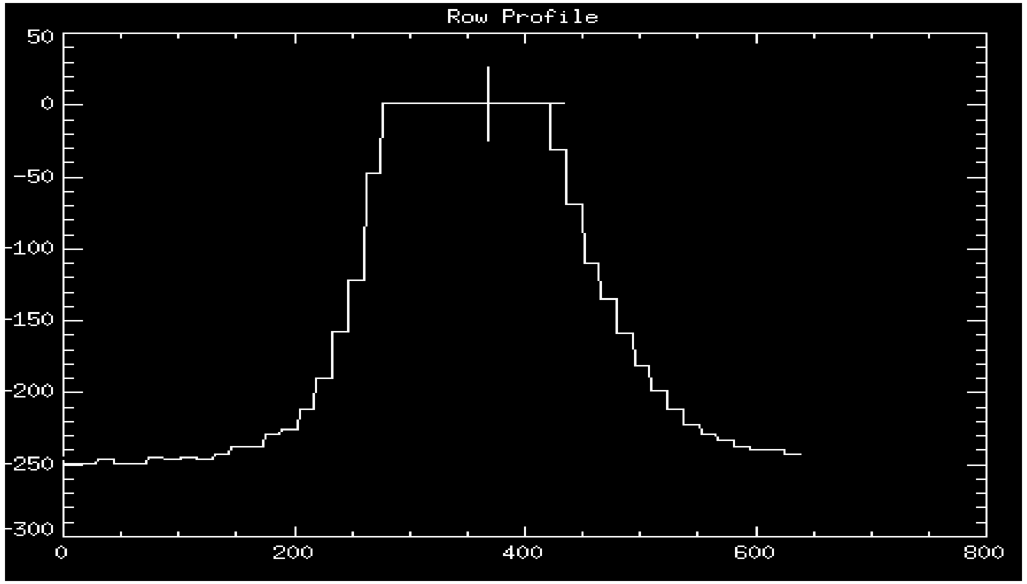
Gels have been altered to respect NDA

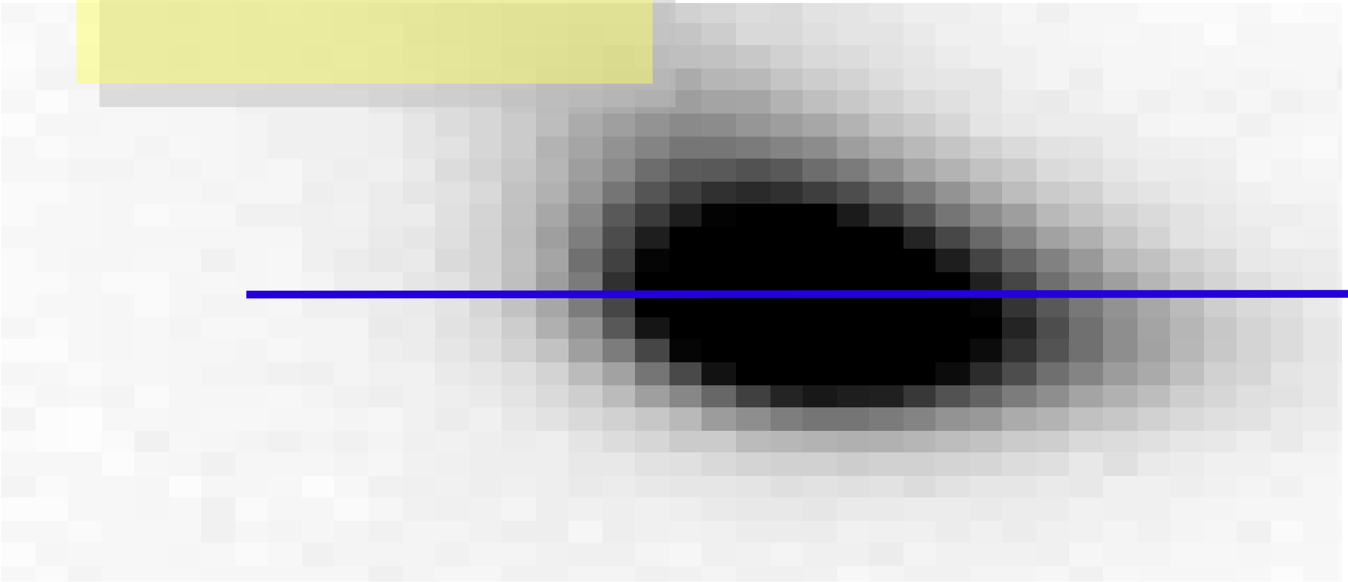
10. Artefacts in 2D gels





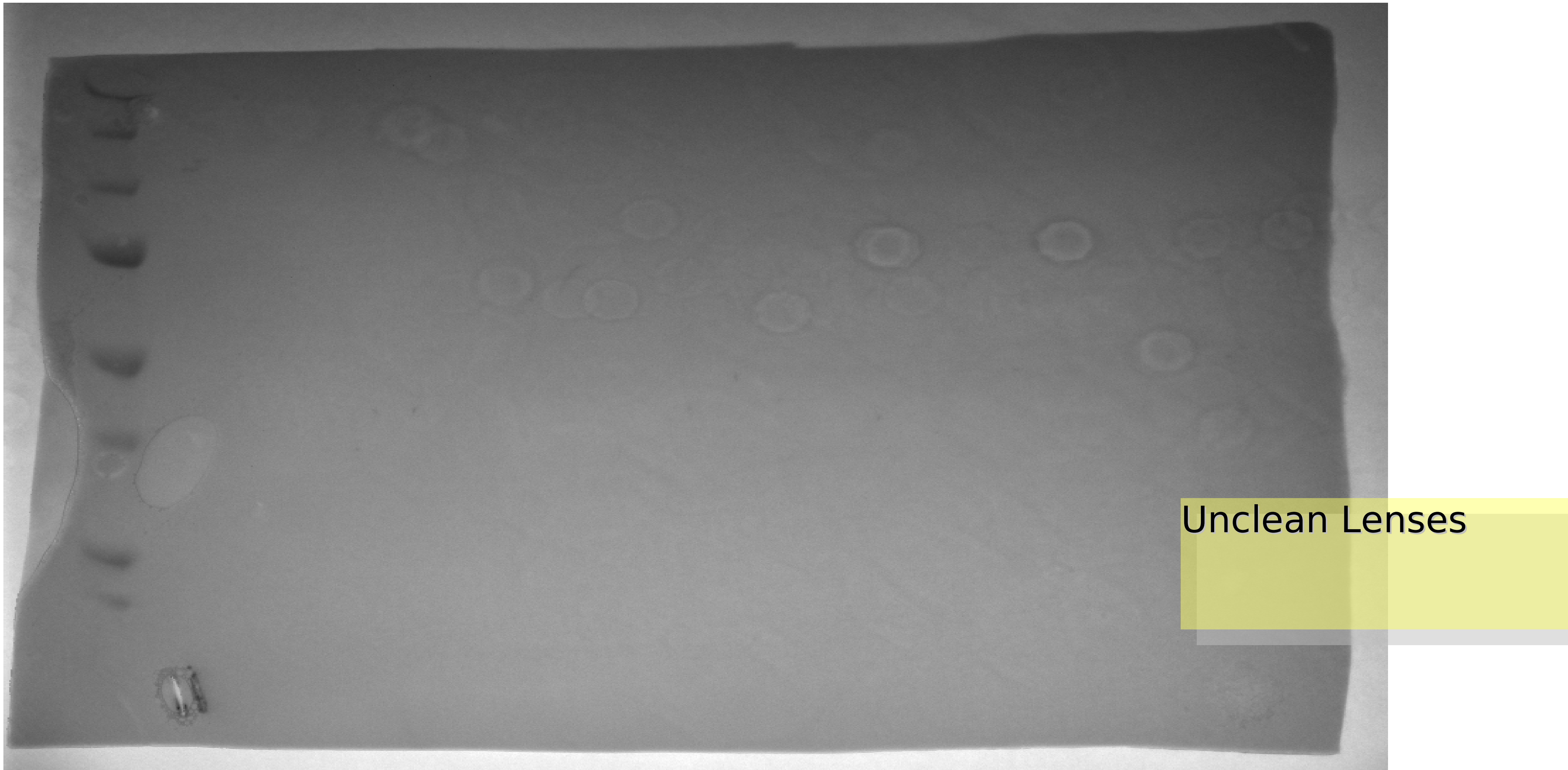
Clipping



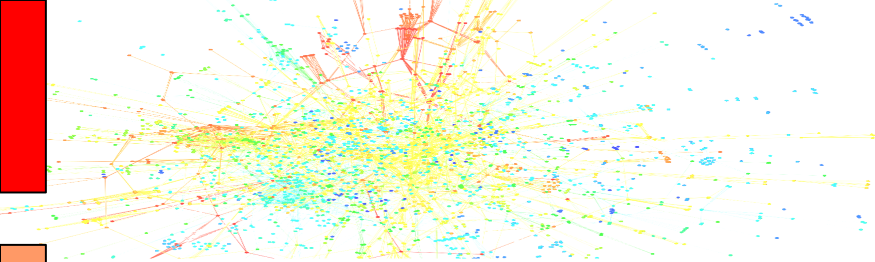


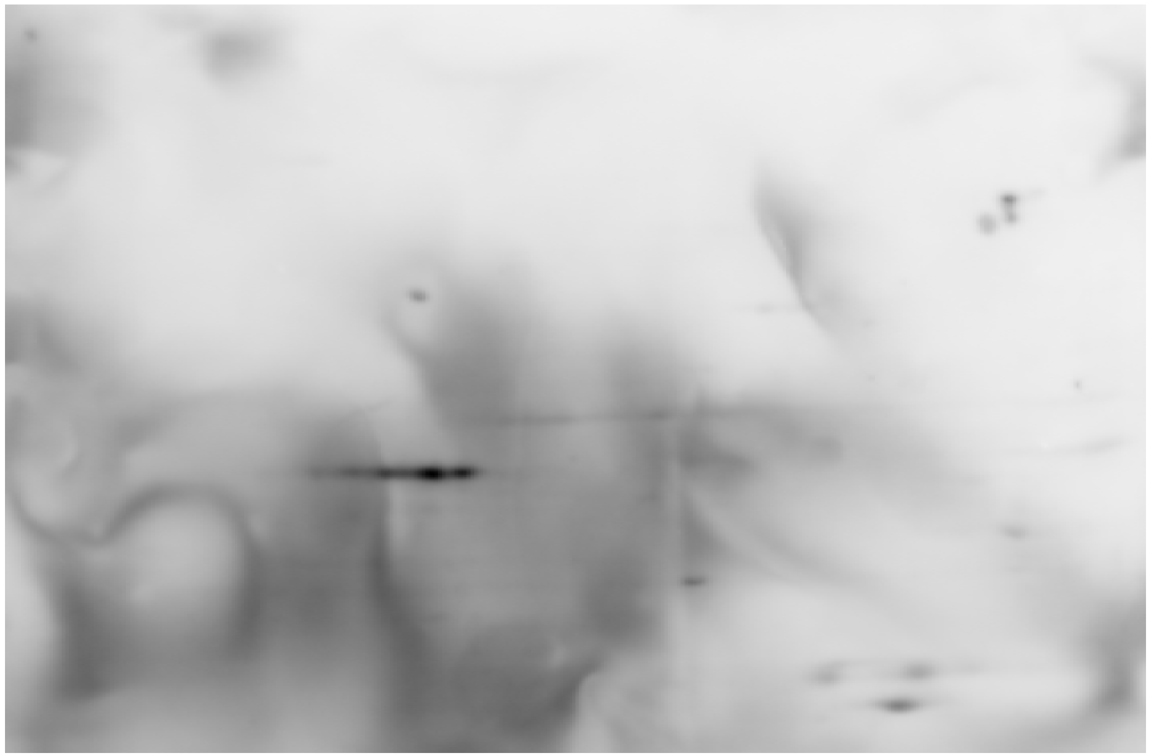
Gels have been altered to respect NDA

11. Artefacts in 2D gels

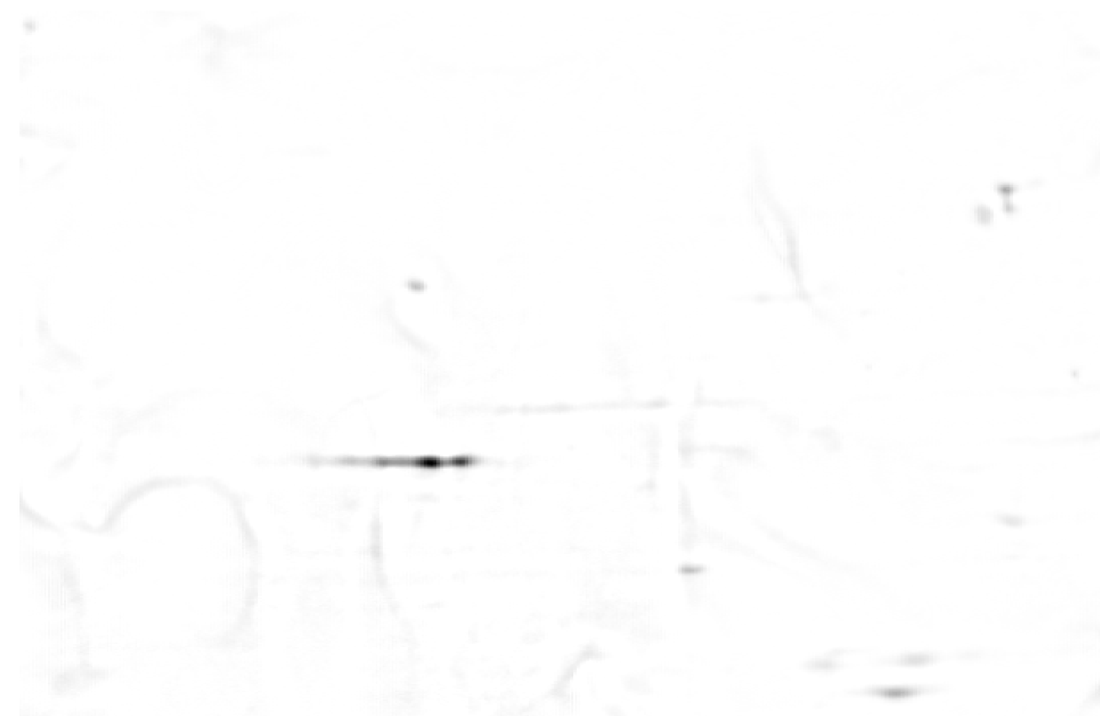


Denoising 2DE gels

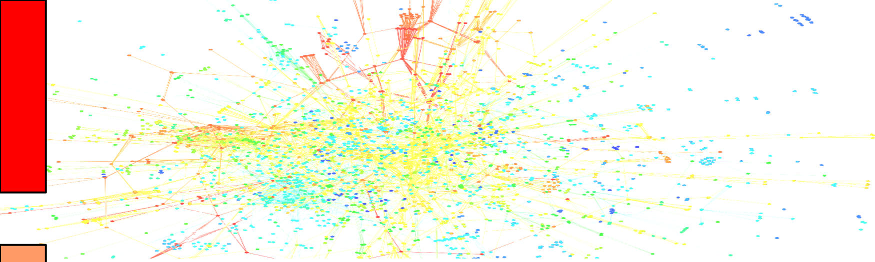




Before



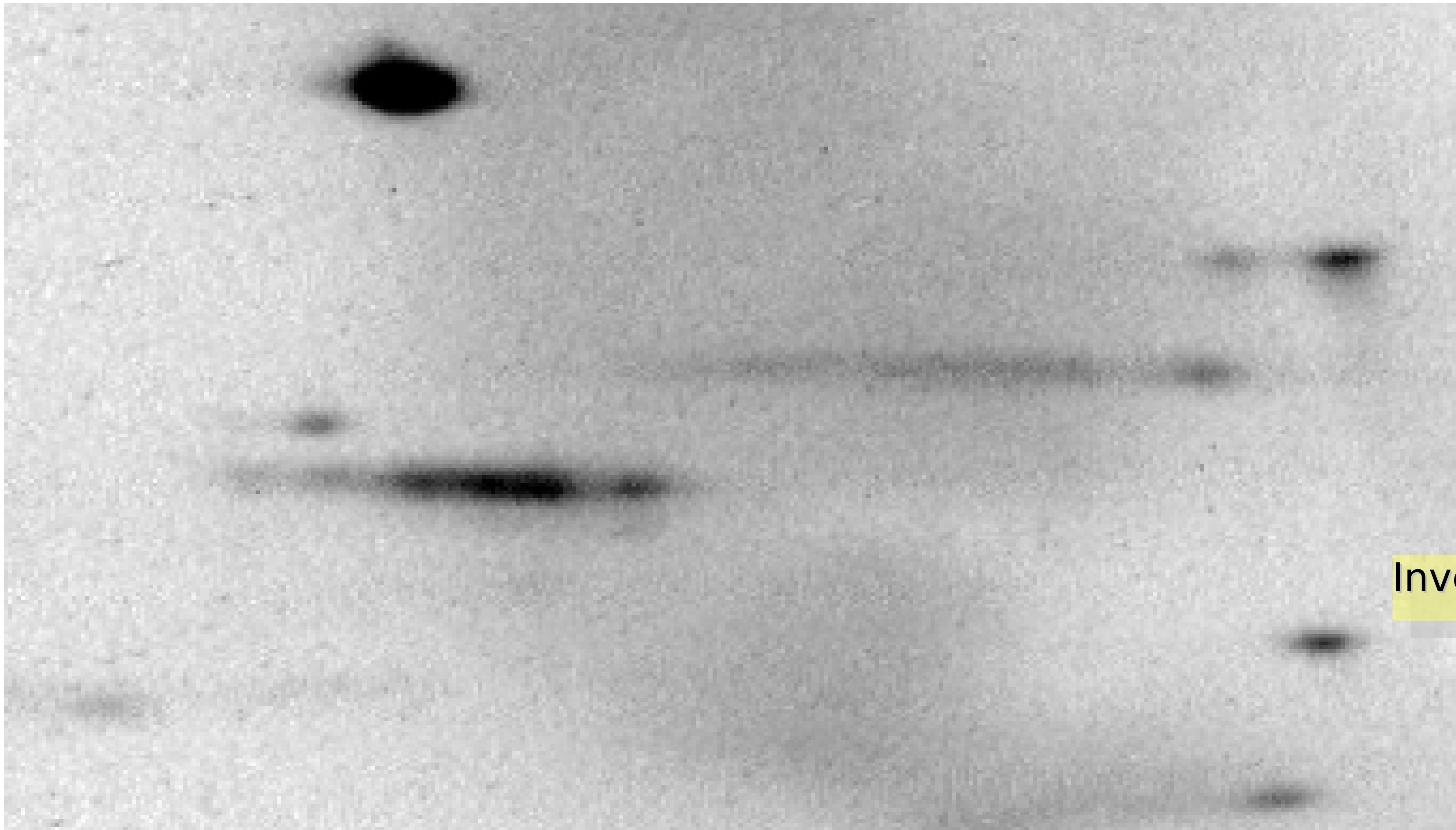
After



Denoising I

Input Image

Denoising II



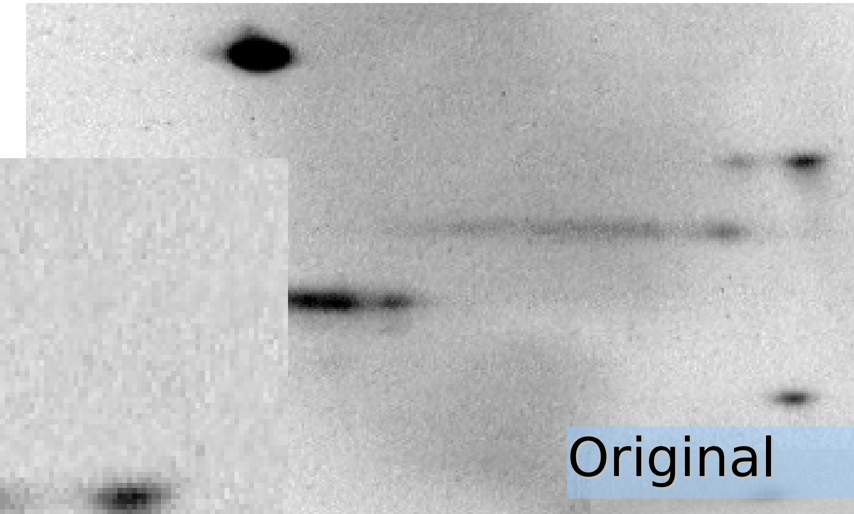
Inverted Image

Denoising III

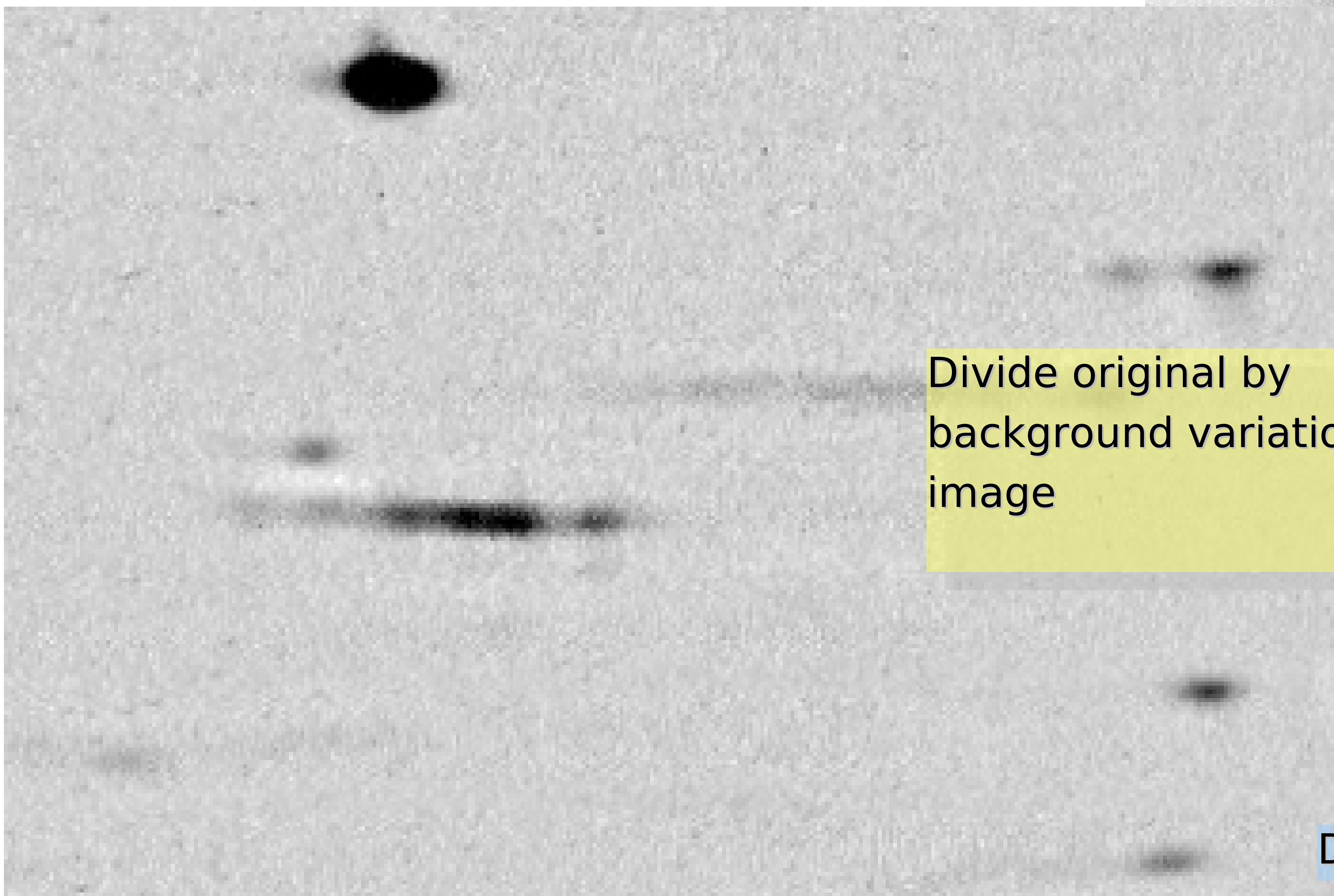
- Calculate Background Variations

Denoising IV

- Remove backround variation



Original

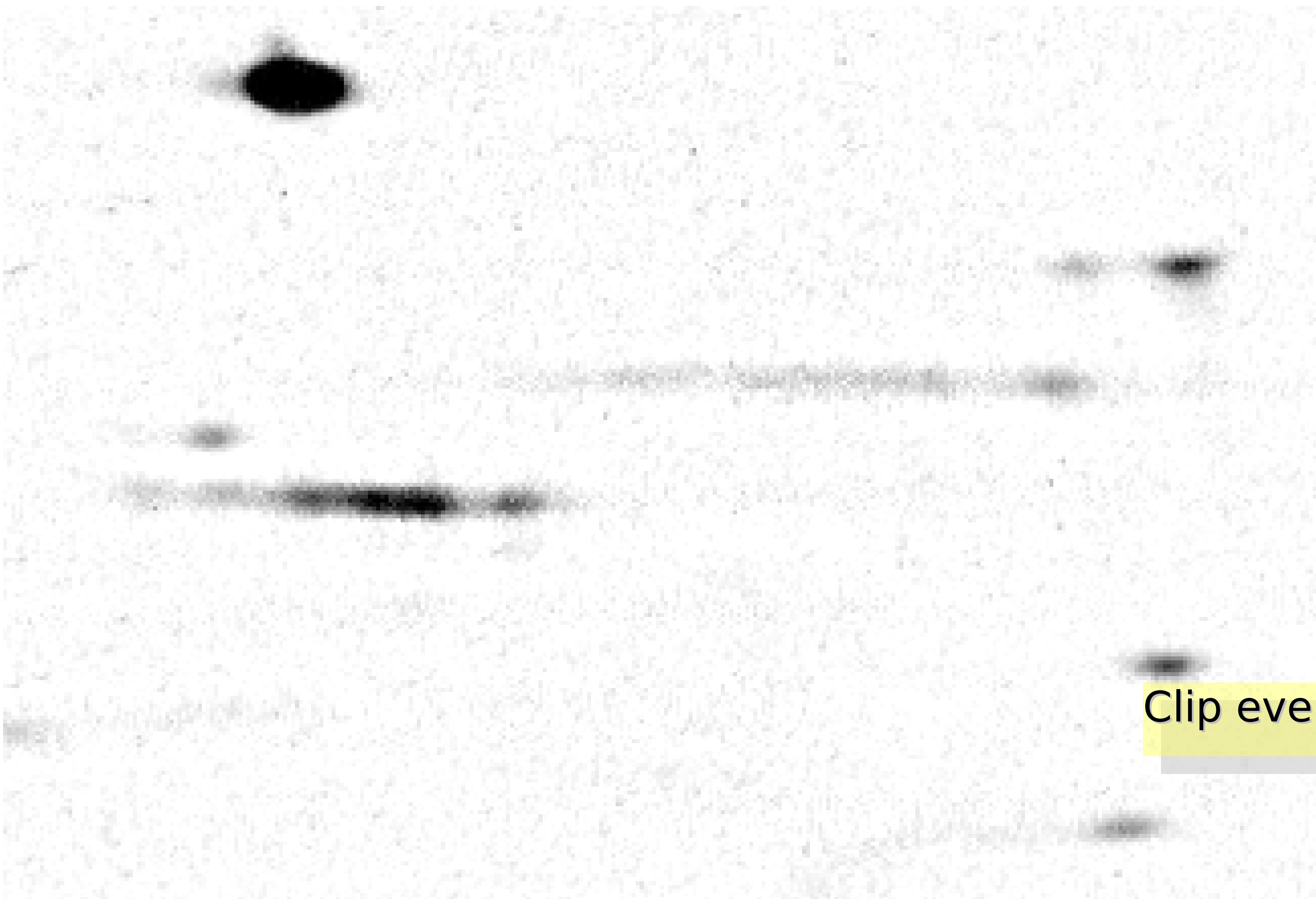


Divide original by background variation image

Denoised

Denoising V

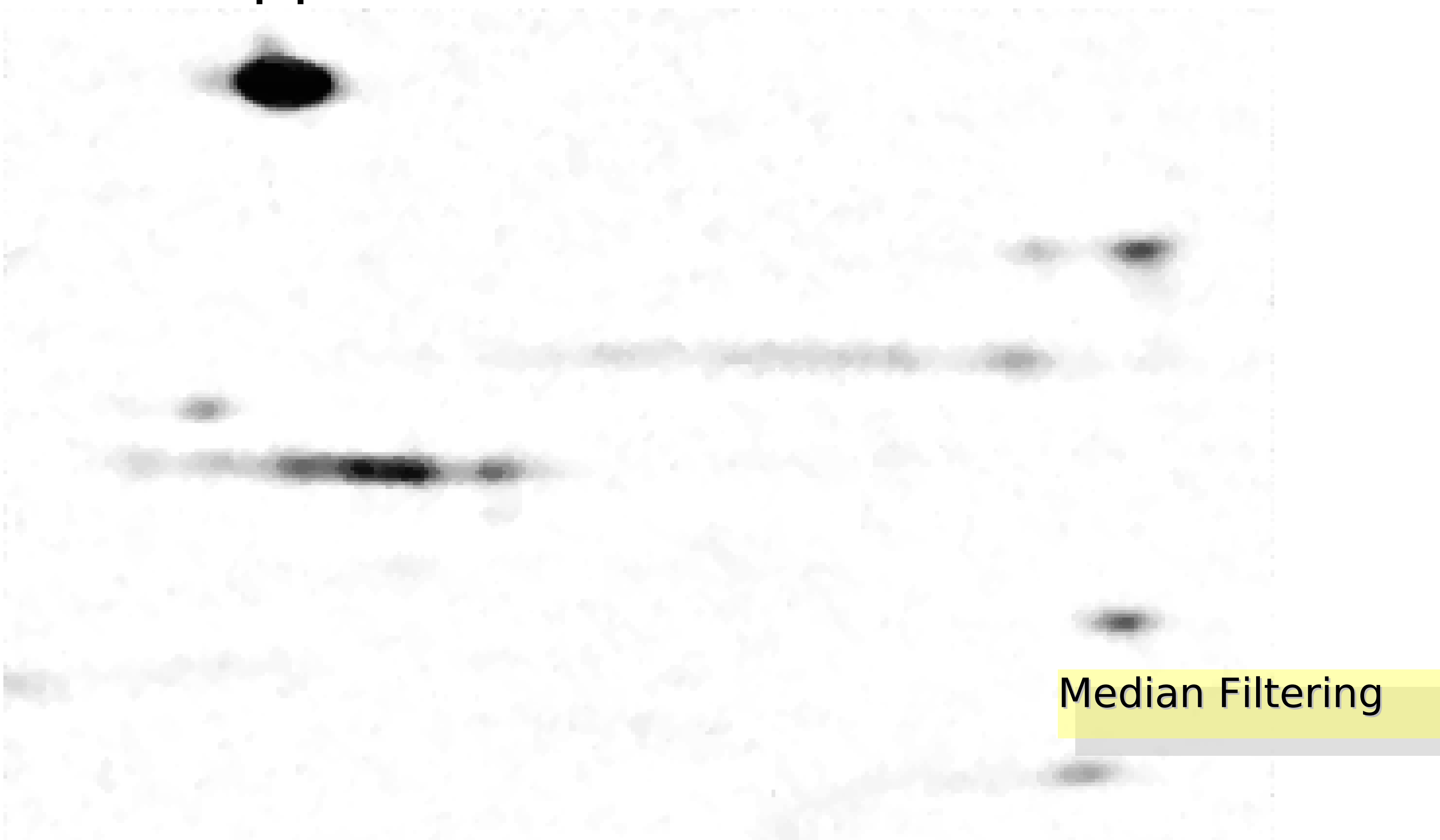
- Thresholded

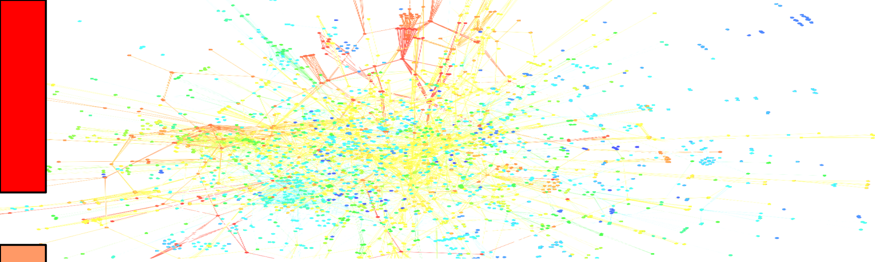


Clip everything > 1

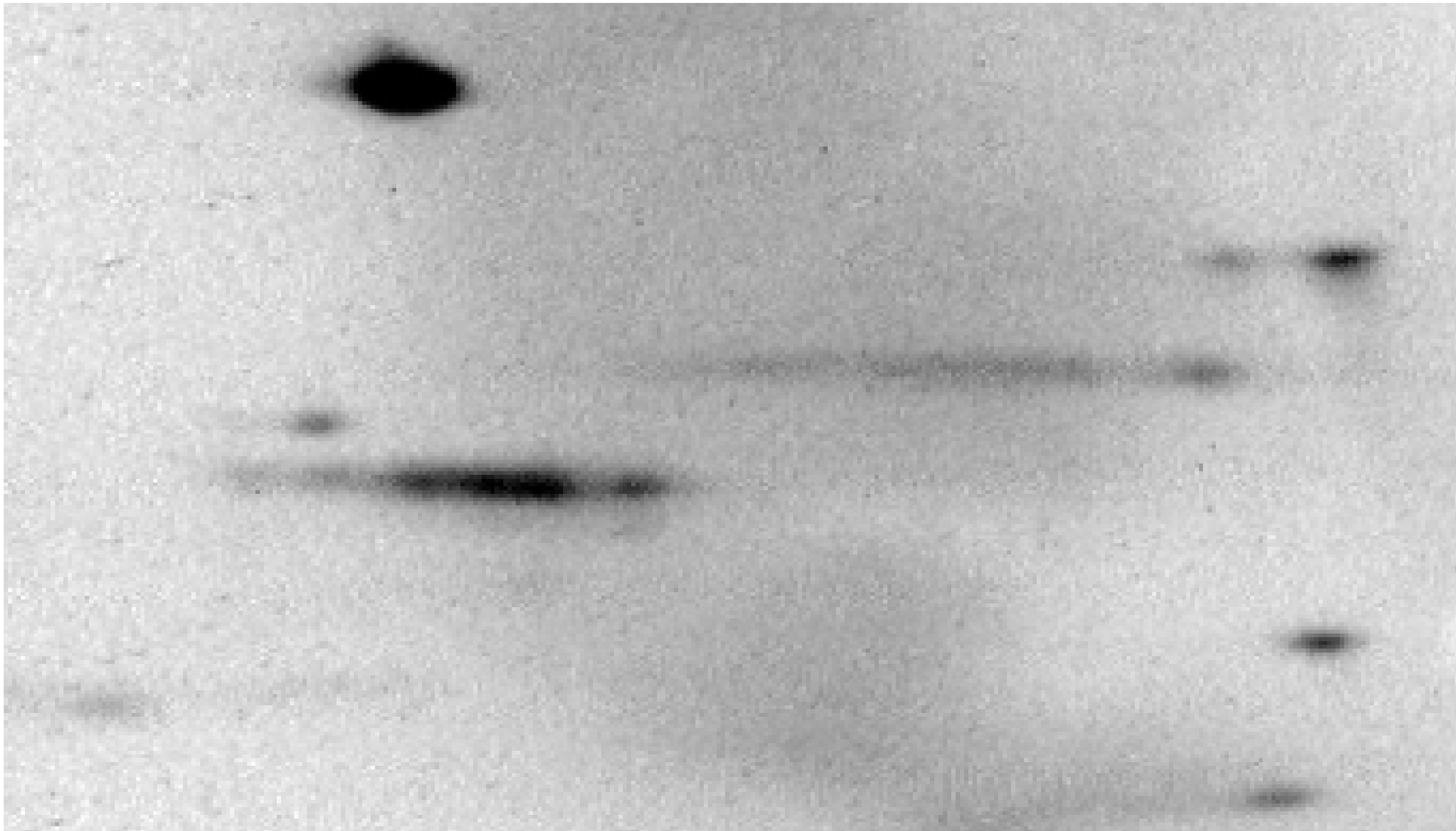
Denoised result

- Salt & Pepper removal

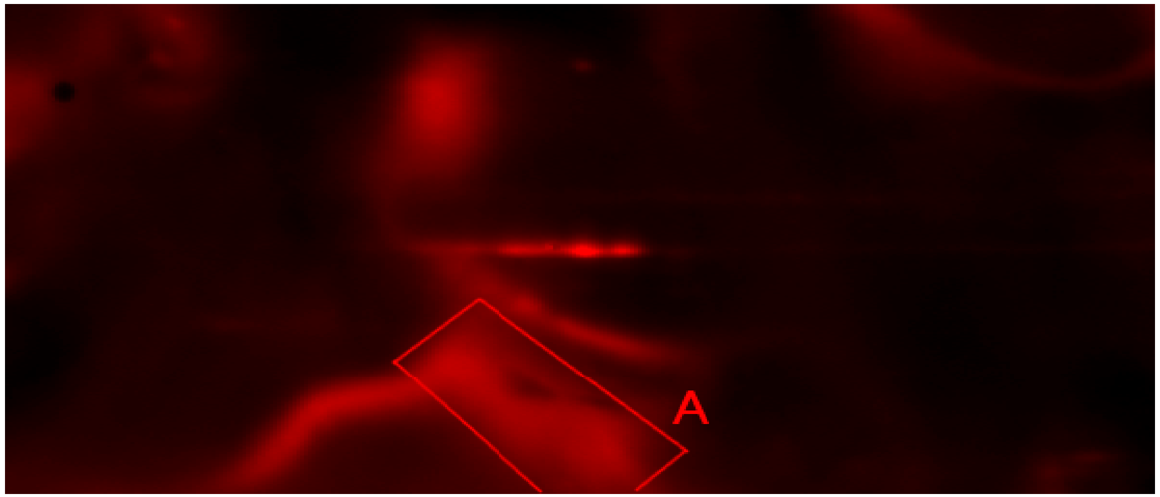


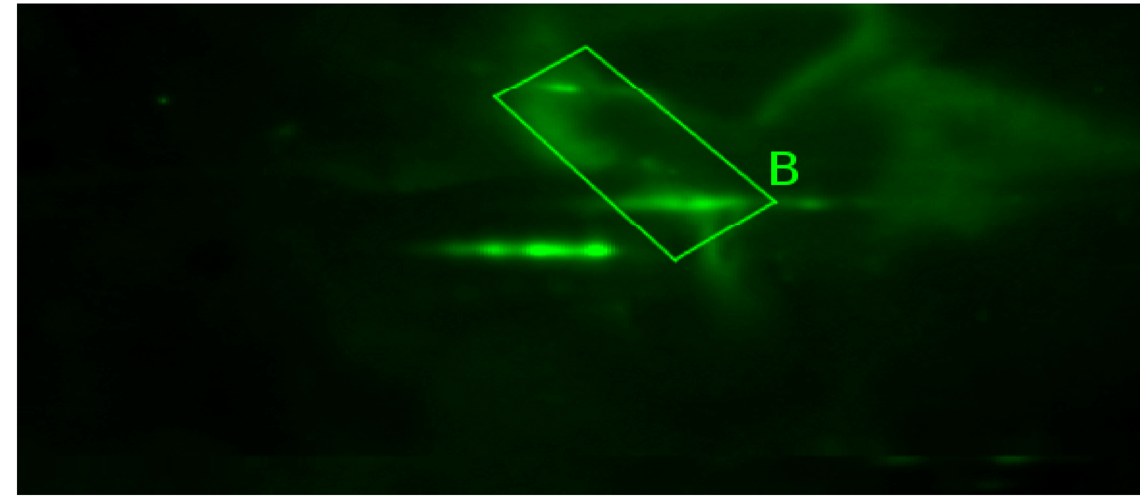


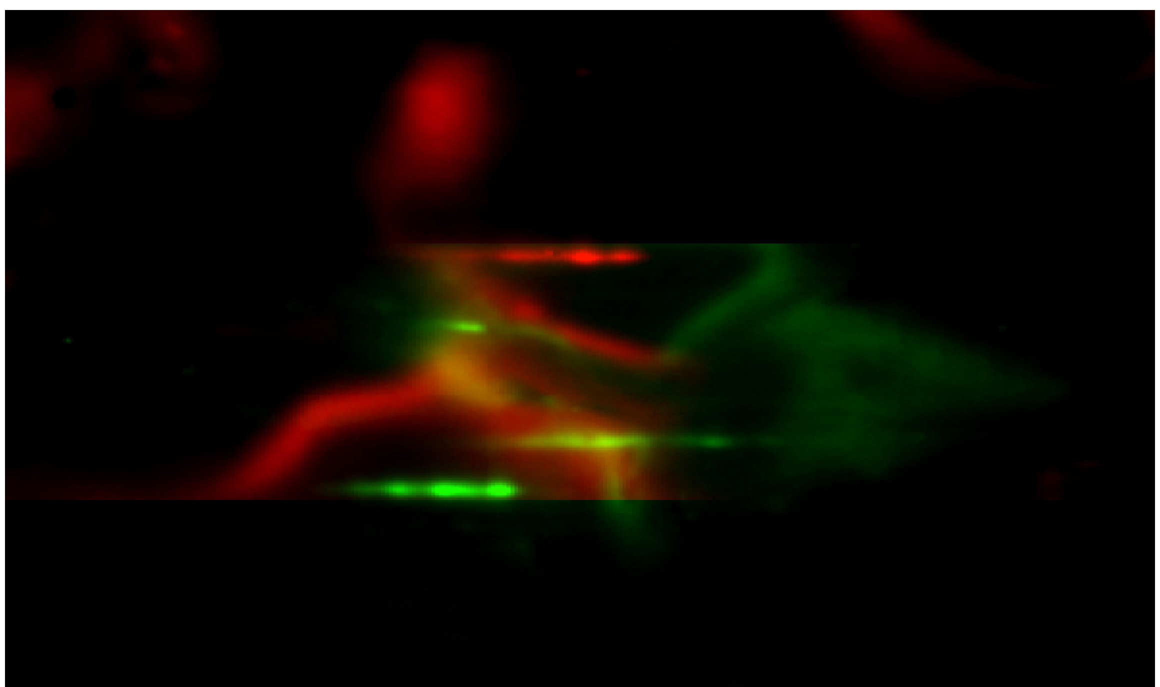
Original

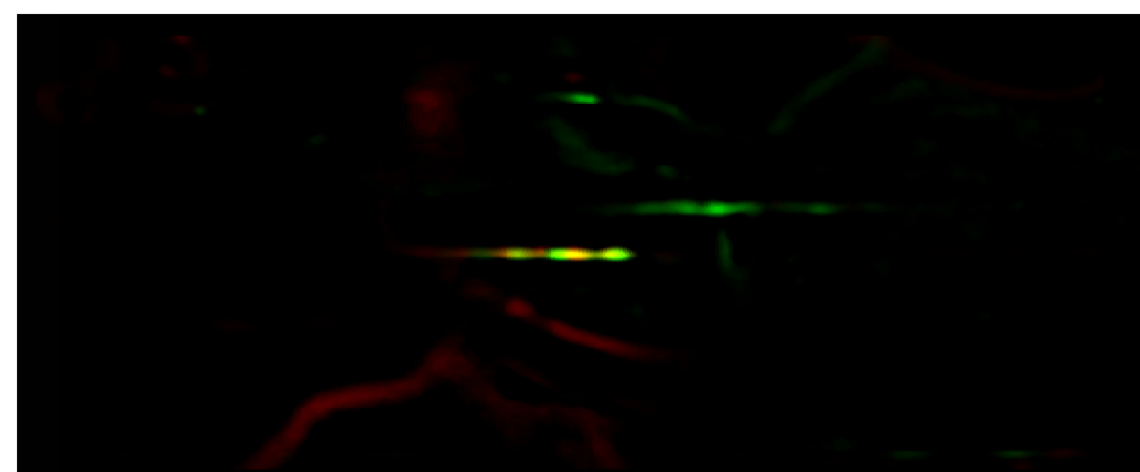


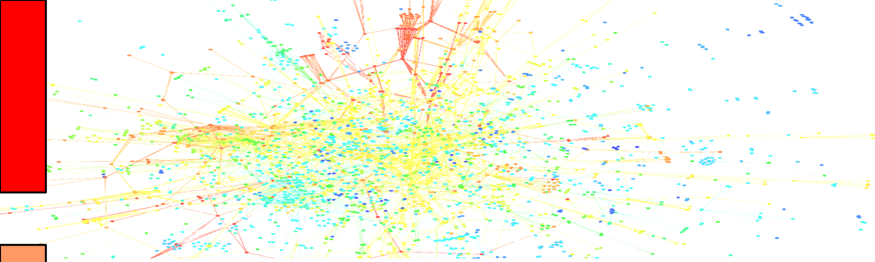
Denoising enables pairwise alignment











Outline

- Correlation analysis
  - Requires multiple aligned gels
    - Multiple gel alignment based on pairwise alignment
      - Pairwise alignment difficult due to many artefacts
        - Developed denoising algorithm
      - Pairwise alignment possible
    - **How to align multiple gels ?**

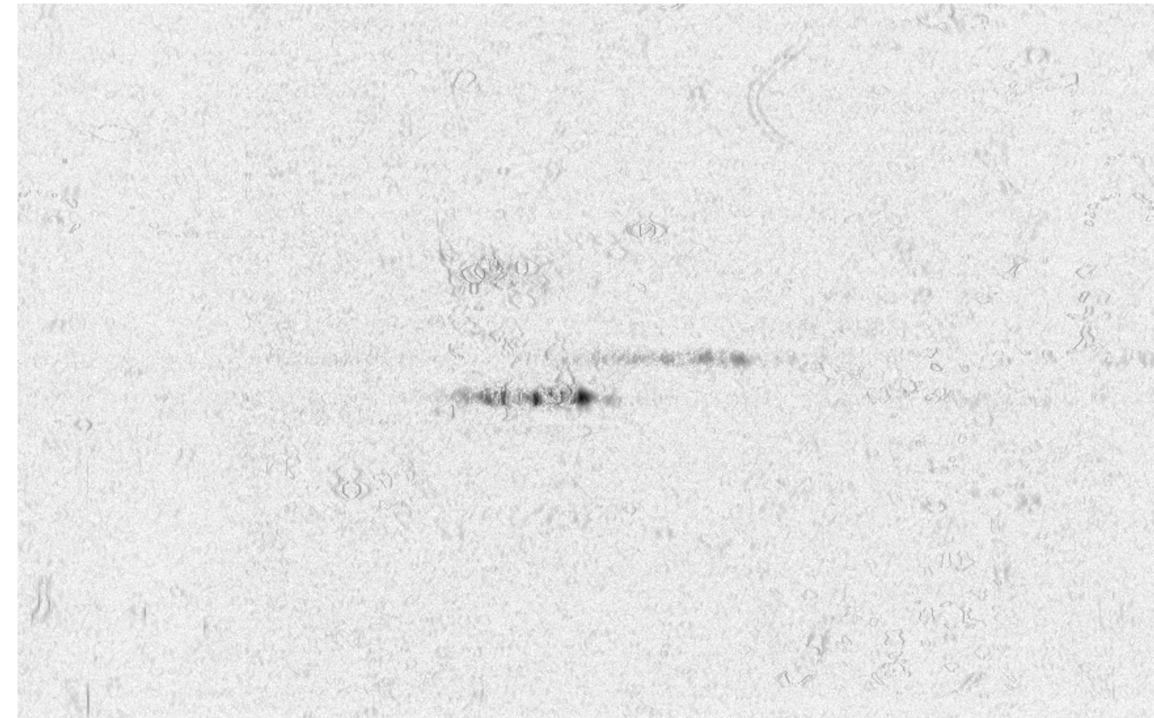
Cummulative Superposition

- Idea
  - take first gel, superimpose second gel
  - take third gel, superimpose on projection of previous gels
  - repeat process for all gels

This does not work,
we merely find a suitable superposition to reflect the first images.

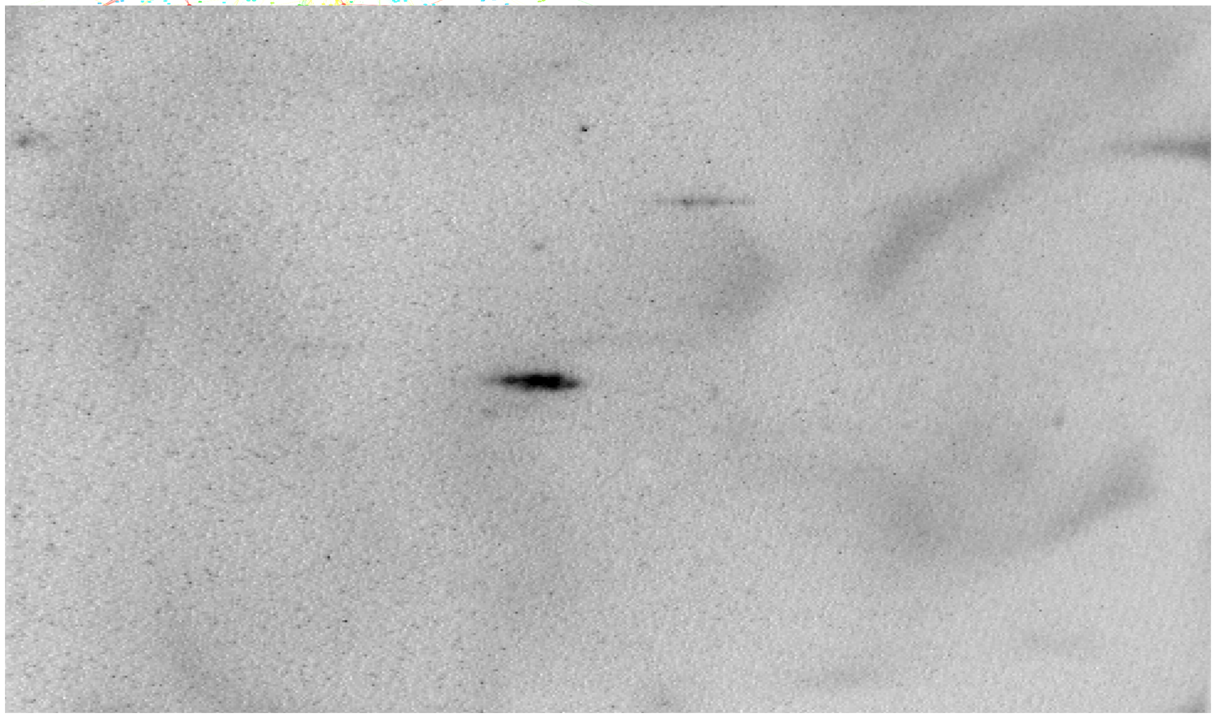
Cummulative Superposition

Initial 2DE Gel Image

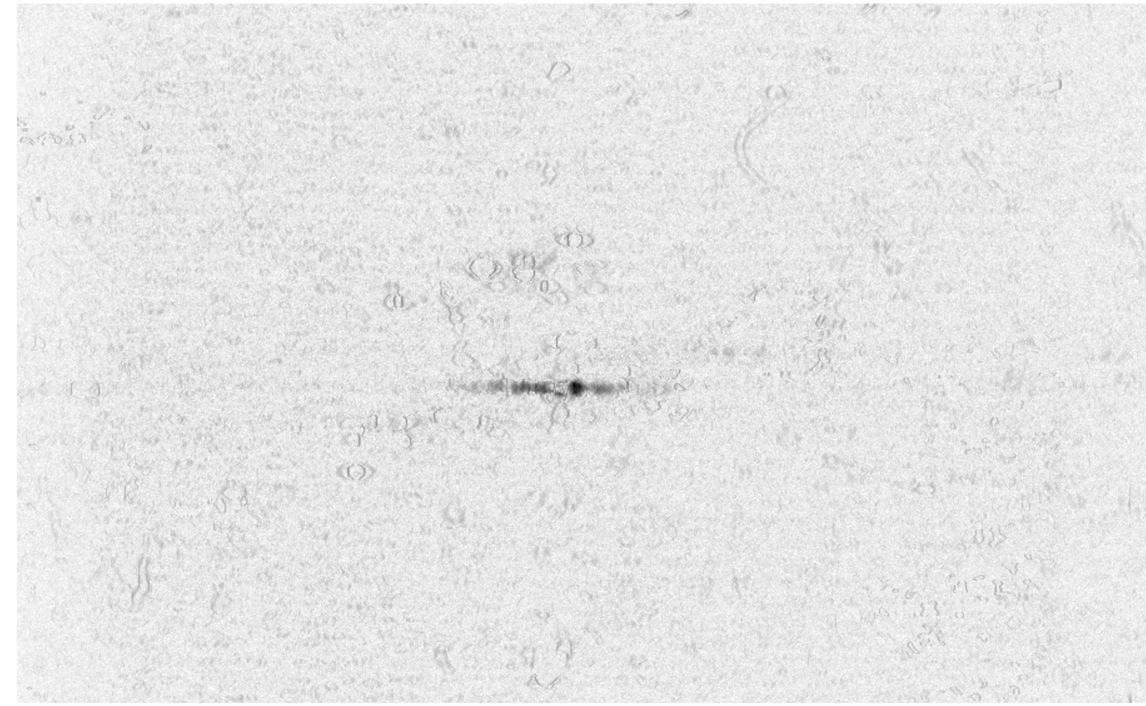


Final Overlay Image

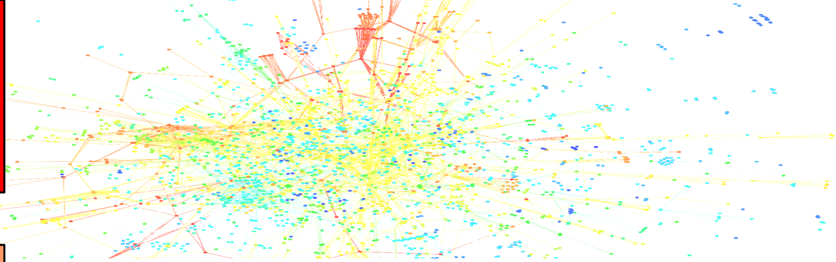
Cummulative Superposition



Initial 2DE Gel Image



Final Overlay Image



Multi Gel Alignment

1- align all image pairs -> X.X alignments

2- find an optimal (x,y) position that minimizes the overall alignment error

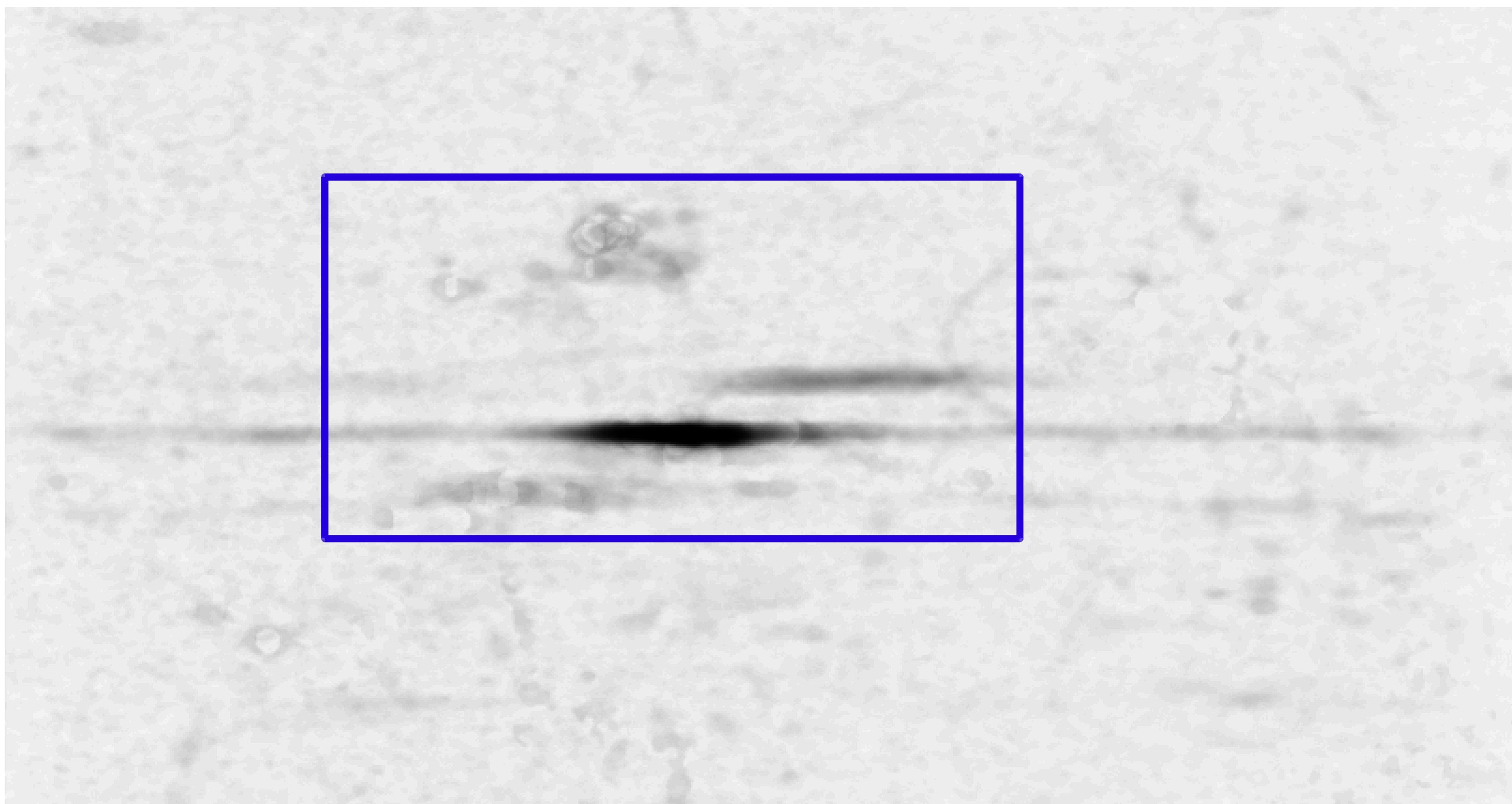
	A	B	C	D
A		(50,80)	(0,-20)	(30,5)
B	(2,45)		(-12,0)	(-12,70)
C	(23,-156)	(15,-73)		
D				
E				

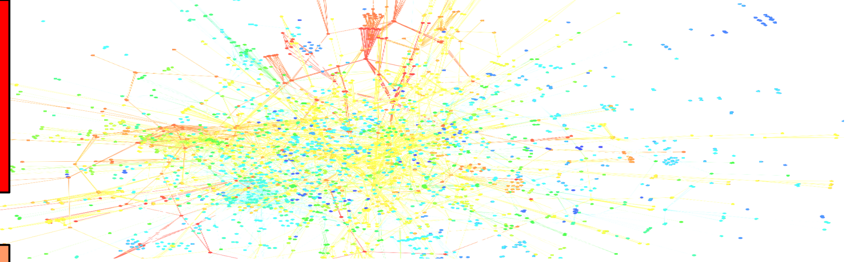
100 images at 1024 x 1024
65011712 operations per cross correlation
5000 cross correlations
325058560000 operations in total
325.10^9 FLOP
theoretical = 2.7 hours
practical = 3 days

2D Gel Overlays

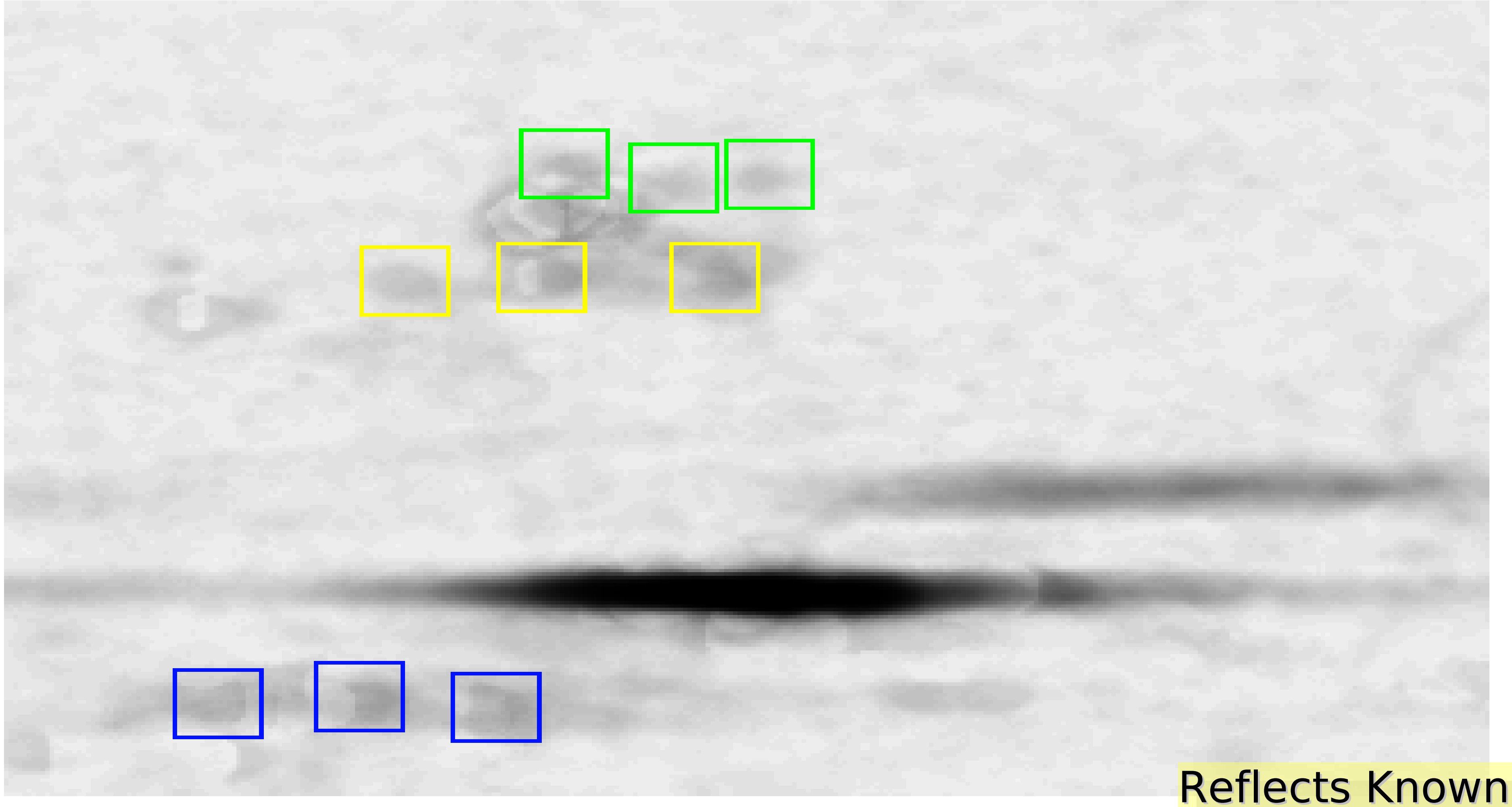
Superposition of all images

Mother image

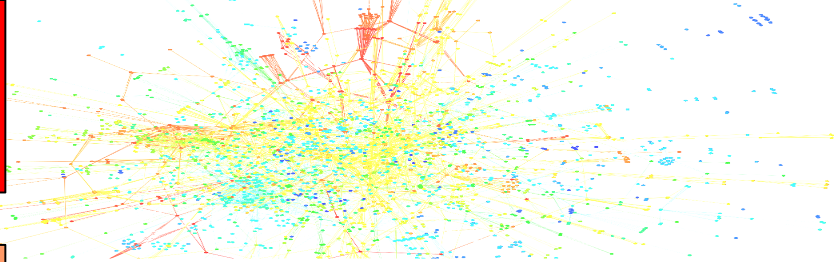




2D Gel Overlays



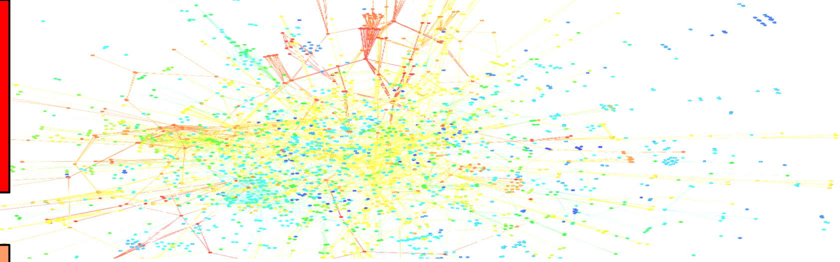
Reflects Known Protein Isoforms



Step 1: Alignment

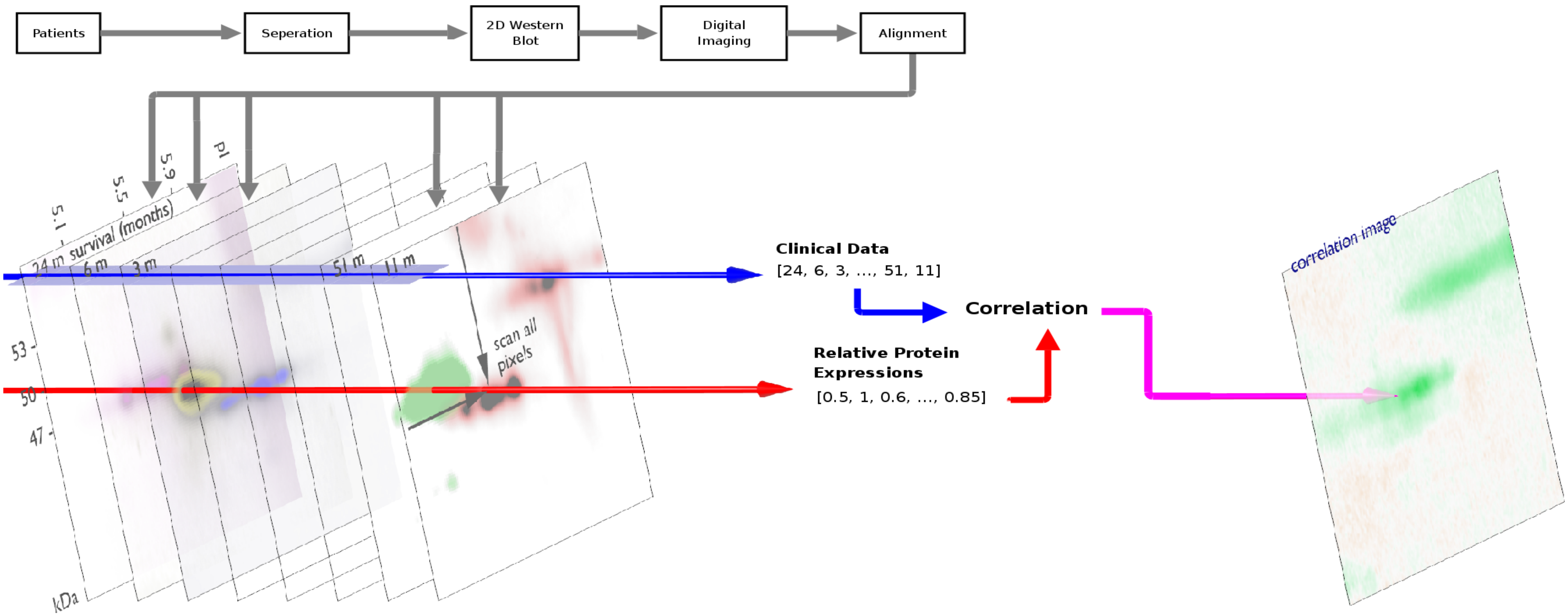
Step 1: Alignment and registration

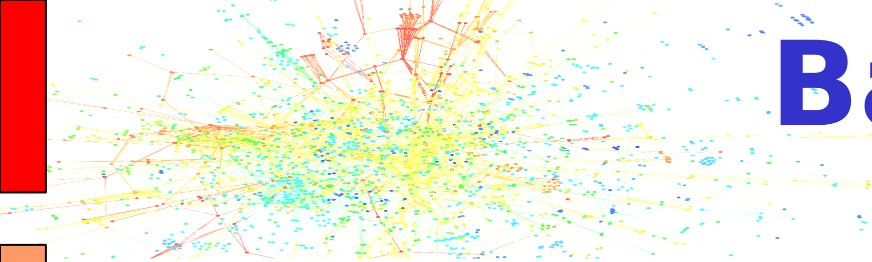
The method requires proper direction and alignment of all gels. Presence of calibration spots facilitates this process, otherwise techniques such as Hough transformation [26, 52] for gel direction measurement and cross correlation [53] for multiple gel alignment can be used. Once the gels are aligned, further basic warping and registration [45] techniques are useful to account for small shifts between the different gels. The aligned images are denoted A'_z.



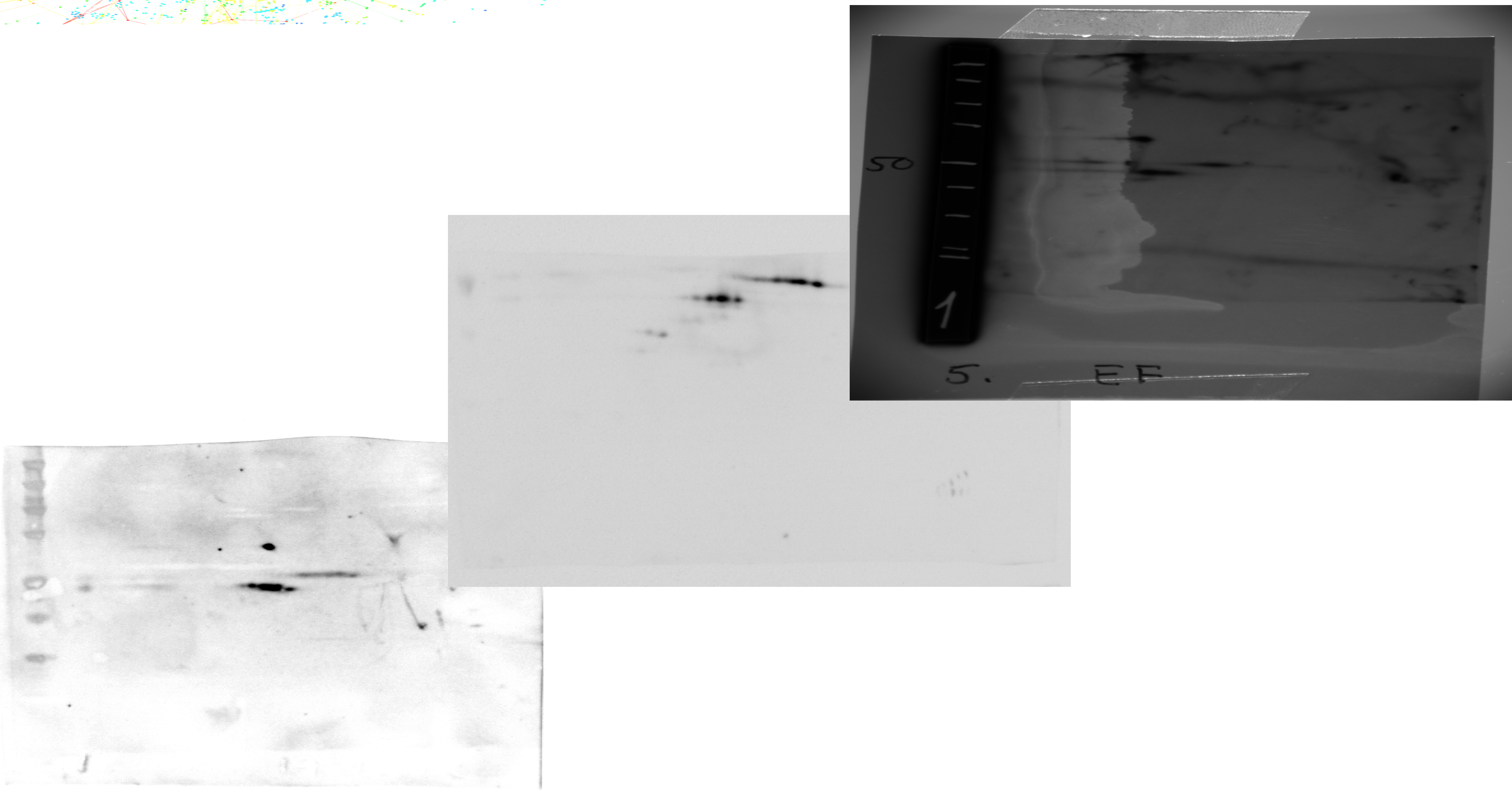
2D Gel Analysis

- Step 2: Intensity Normalization

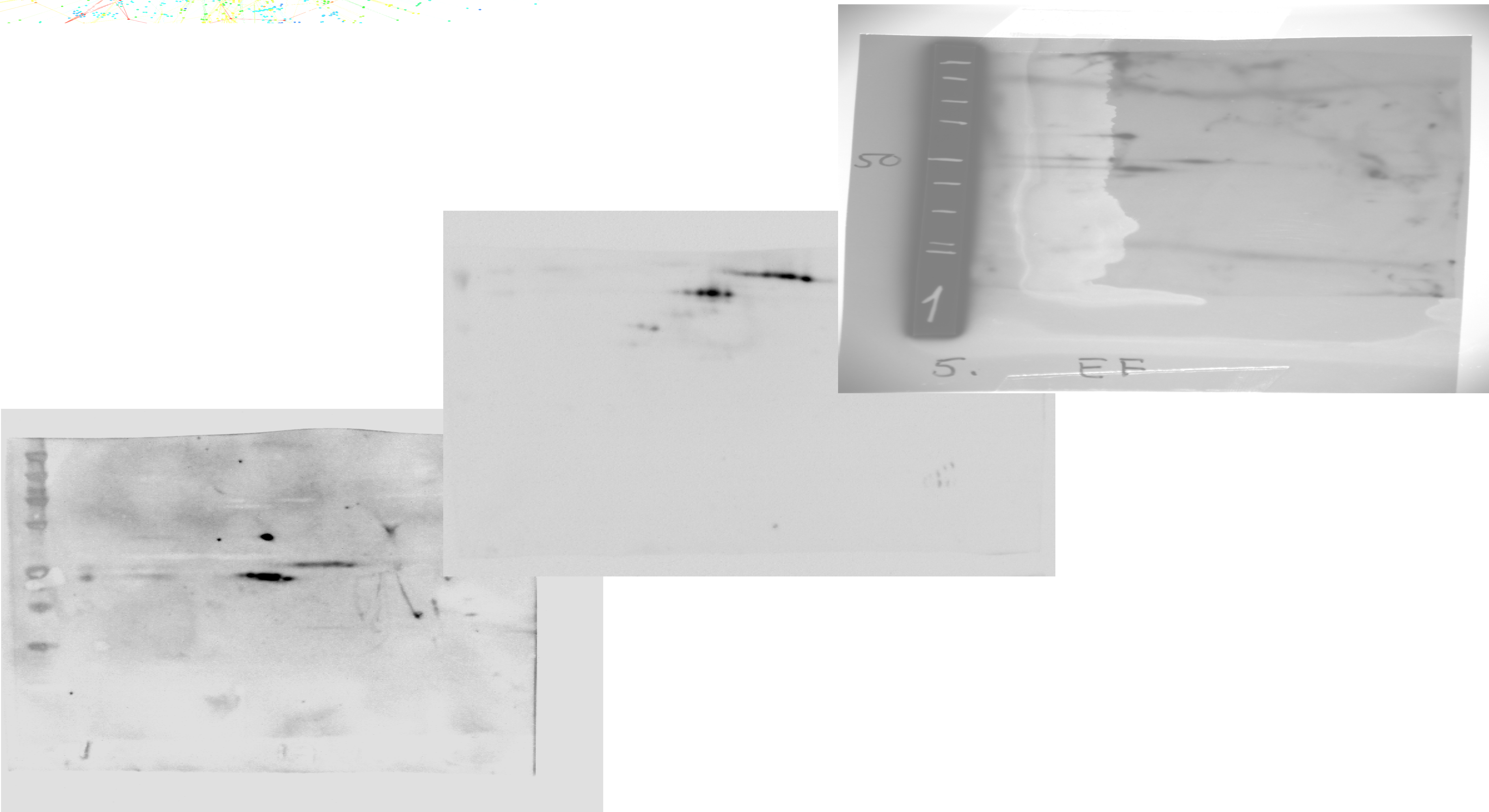




Background Differences



Background Differences

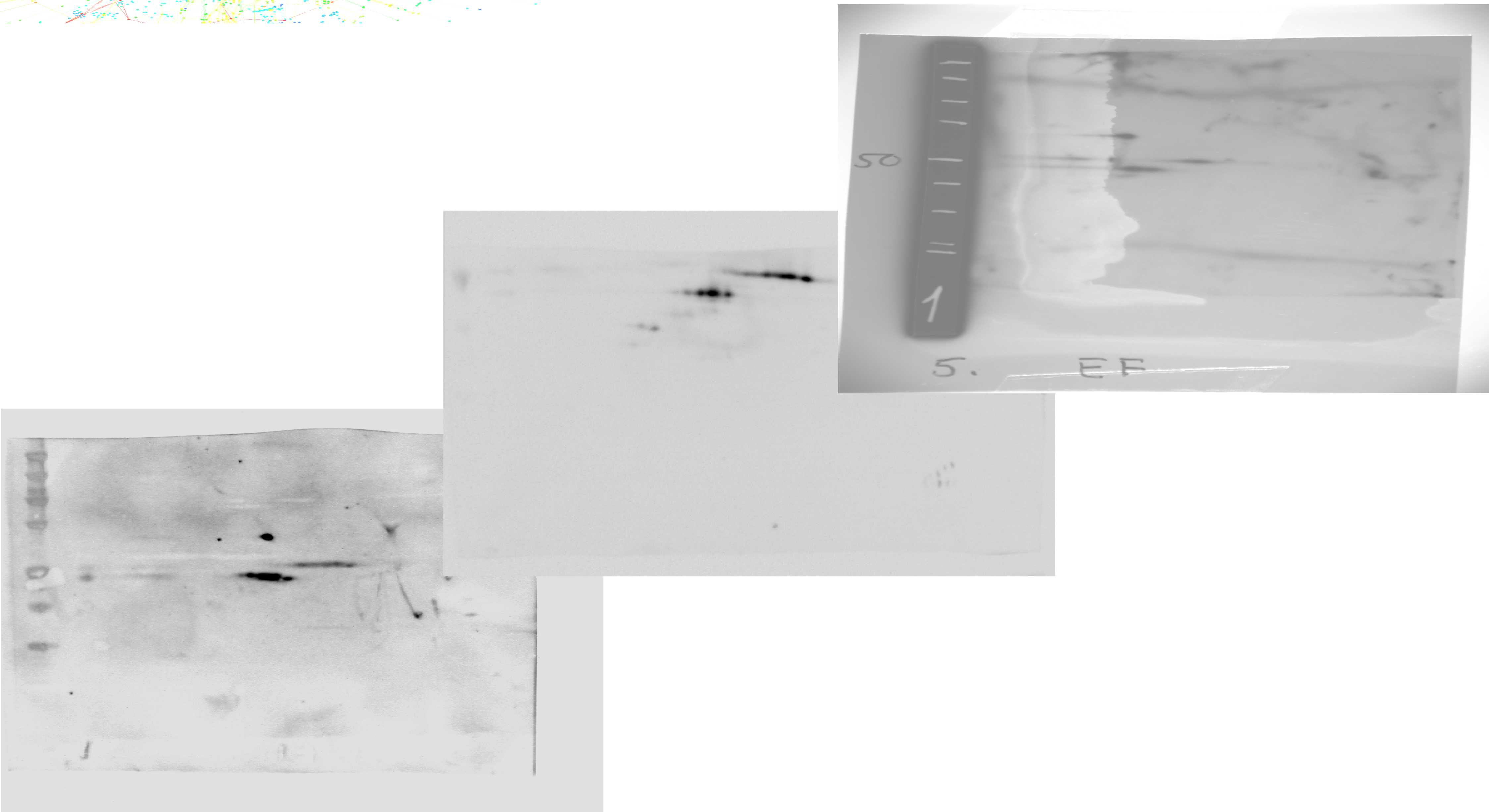


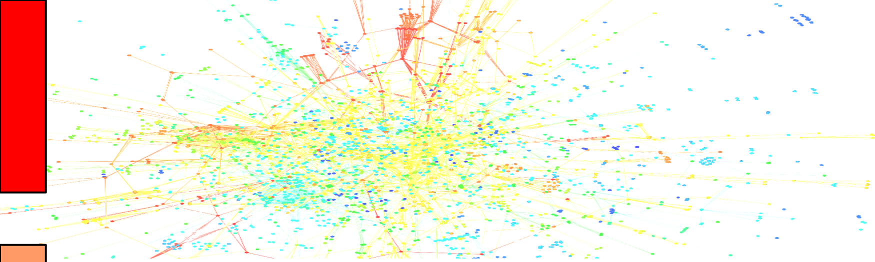
Step 2a: Background Intensity

Step 2a: Background intensity

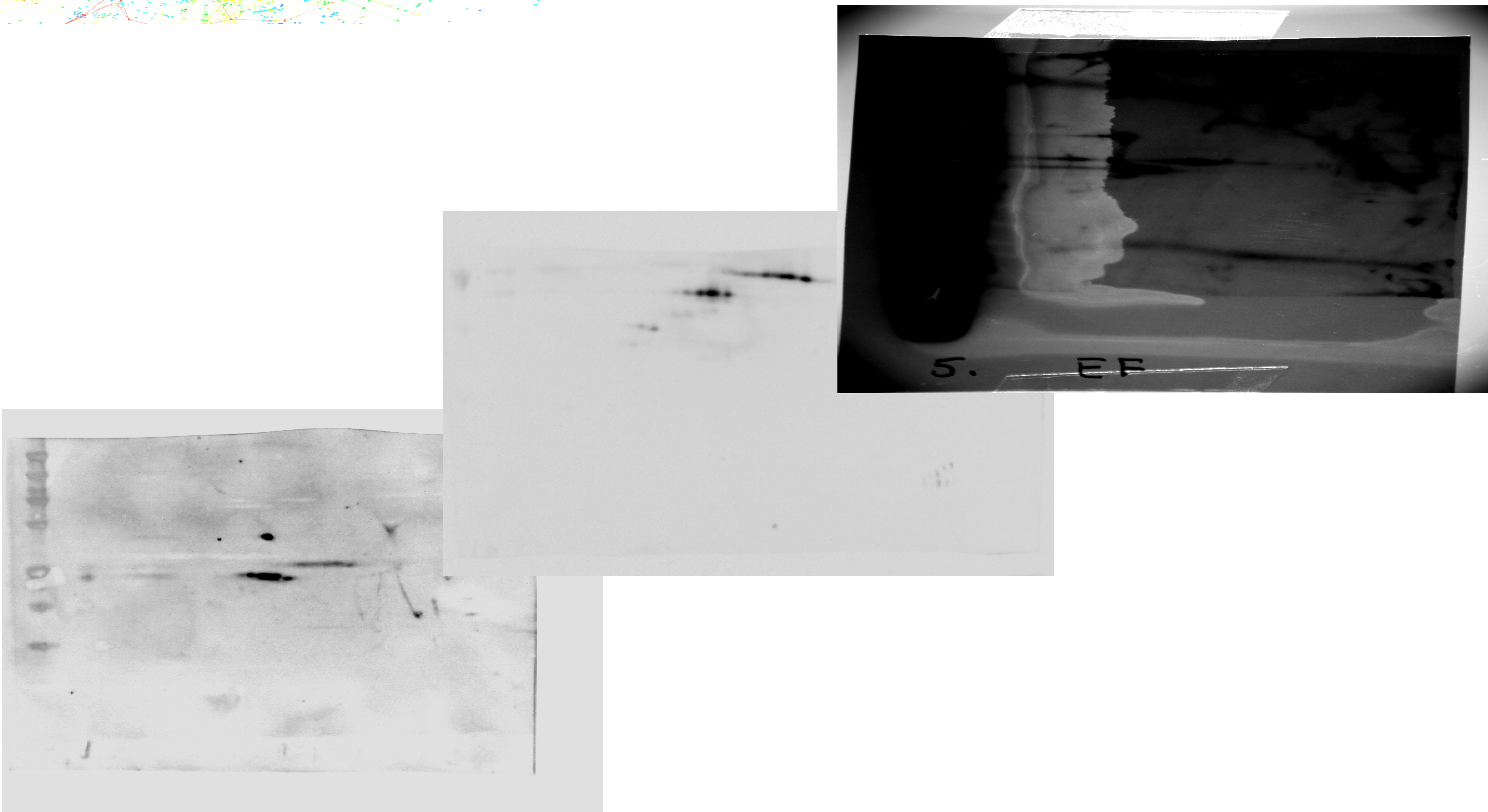
The background floor of a 2DE image refers to the brightness of empty gel areas. Different capture techniques produce different background floors. Background signal can be either added to all pixel values (additive background), or it can accumulate with a decaying signal (multiplicative background). As previously observed [44], most cameras introduce a mixture of additive and multiplicative backgrounds. Removal of additive noise can be done through subtracting the mean ($A''_z := A'_z - \overline{A'_z}$) or median value ($A''_z := A'_z - median(A'_z)$). Removal of multiplicative noise can be done through $A''_z := \frac{A'_z}{\overline{A'_z}} - 1$. We would emphasize that whatever normalization scheme is used in this step, it should be performed on an individual gel basis.

Contrast





Contrast



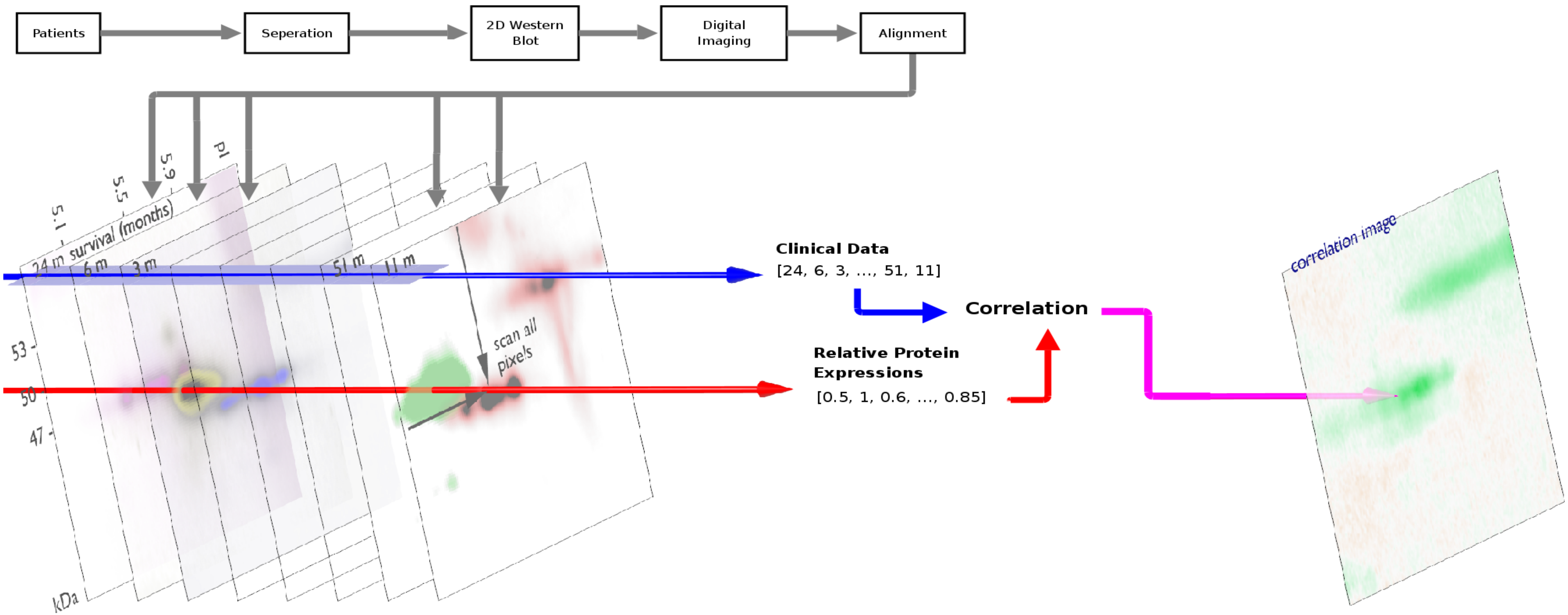
Step 2b: Intensity Normalization

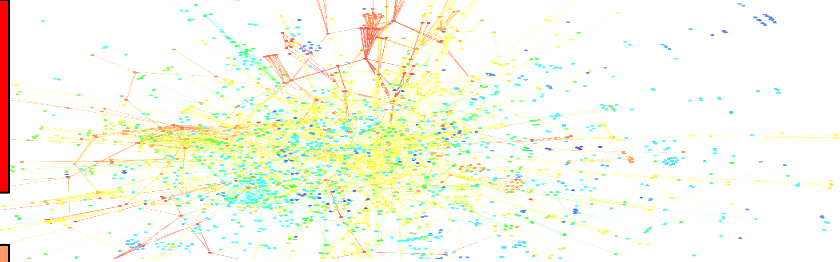
Step 2b: Scaling of gel intensity

After removal of the background floor, the dynamic range of the image is normalized through scaling of gel intensities. The presence of a calibration spot eases this process. If A' is the non-relative image and (x, y) is the calibration spot position, then the image $A'' := \frac{A'}{A'_{x,y}}$ defines the normalized image. Without calibration spot the total energy content (sum of all intensities or RMS value) forms a very reasonable scaling means: $A''_z = \frac{A'_z}{RMS(A'_z)}$

2DE Gel Analysis

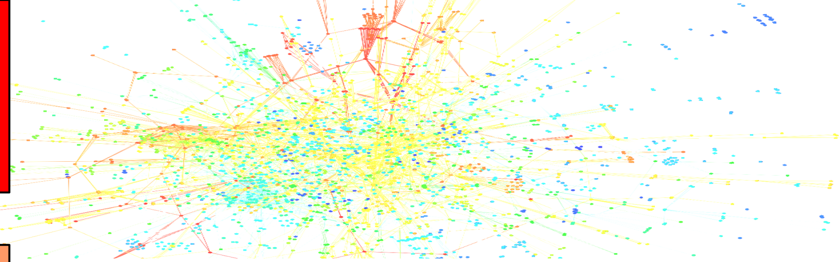
- Step 3: Correlate





Step 3: Correlation

Step 3: Correlation image The correlation image is composed of pixels, each testing one position on the gel. The result of each test is a number between -1.0 (anti-correlation) and 1.0 (correlation), which, after appropriate scaling, defines the pixel color in the correlation image. The two vectors participating in the test are $A^{"}_{x,y}$ and B. The first vector contains the gel expression levels at position (x, y). Given 89 gel images, $A^{"}_{x,y}$ will contain 89 different expression values; one for each gel. The second vector B contains 89 external values associated with every gel. Repeating this correlation test for every pixel results in the correlation image C (Eq. 1)



Step 3: Correlation

$$C_{x,y} = \rho(A^{"}_{x,y}, T) \tag{1}$$

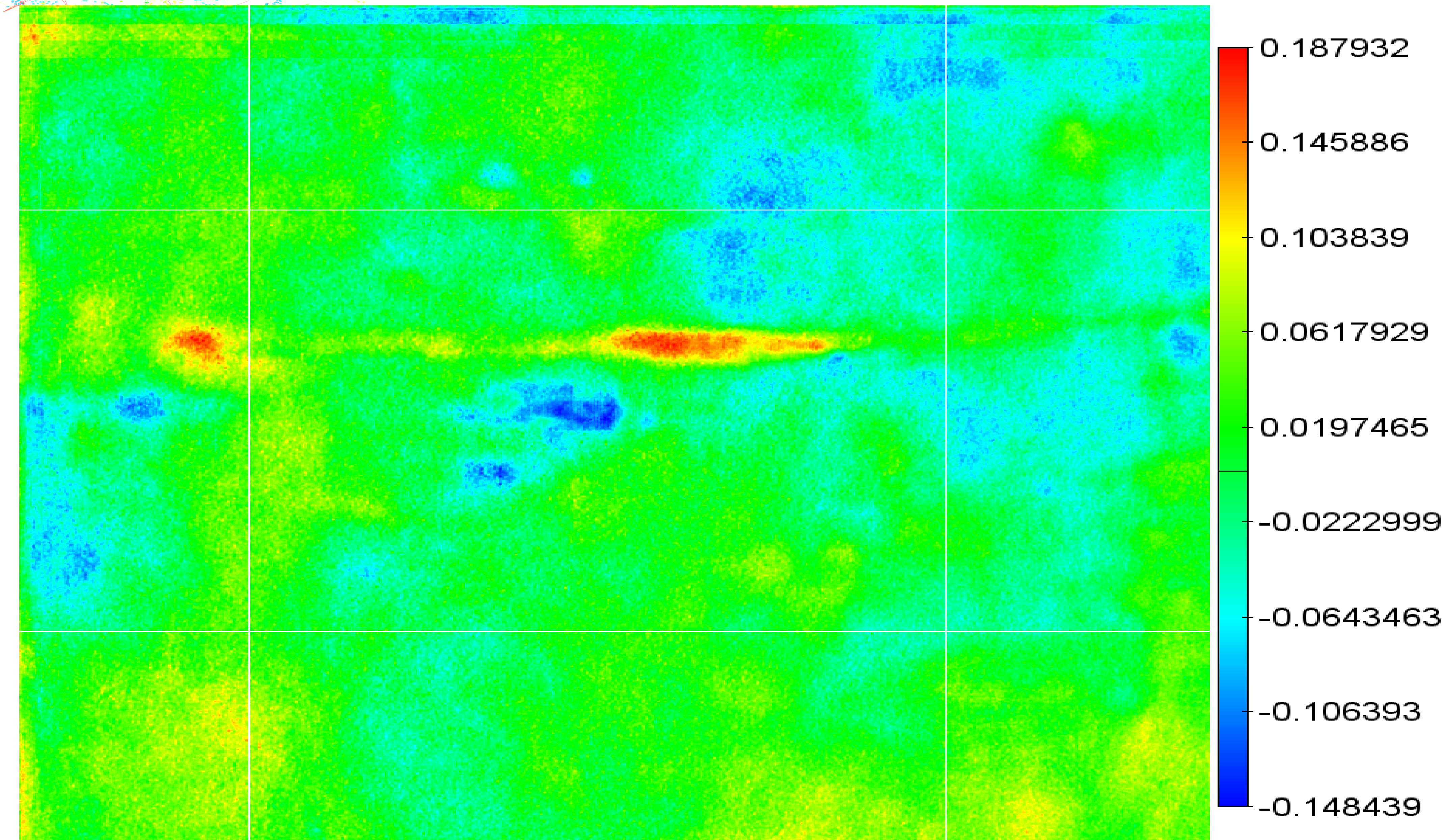
The correlation image can be visualized using different color schemes. In Fig. 1 green indicates positive correlations and brown negative correlations.

The preferred correlation is the robust Spearman rank order correlation (ρ-correlation) [27]. This non-parametric test allows us to ignore the specific distributions of gel intensity levels and external parameters. ρ-correlation requires a ranking of the two participating vectors and then relies on a standard linear Pearson correlation. The ranking process will replace every value in the input vector by its specific rank. When ties occur (the same value occurring more than once) their rank will by convention be the mean of their ranks as if they all would have had a slightly different value.

Initial Problem

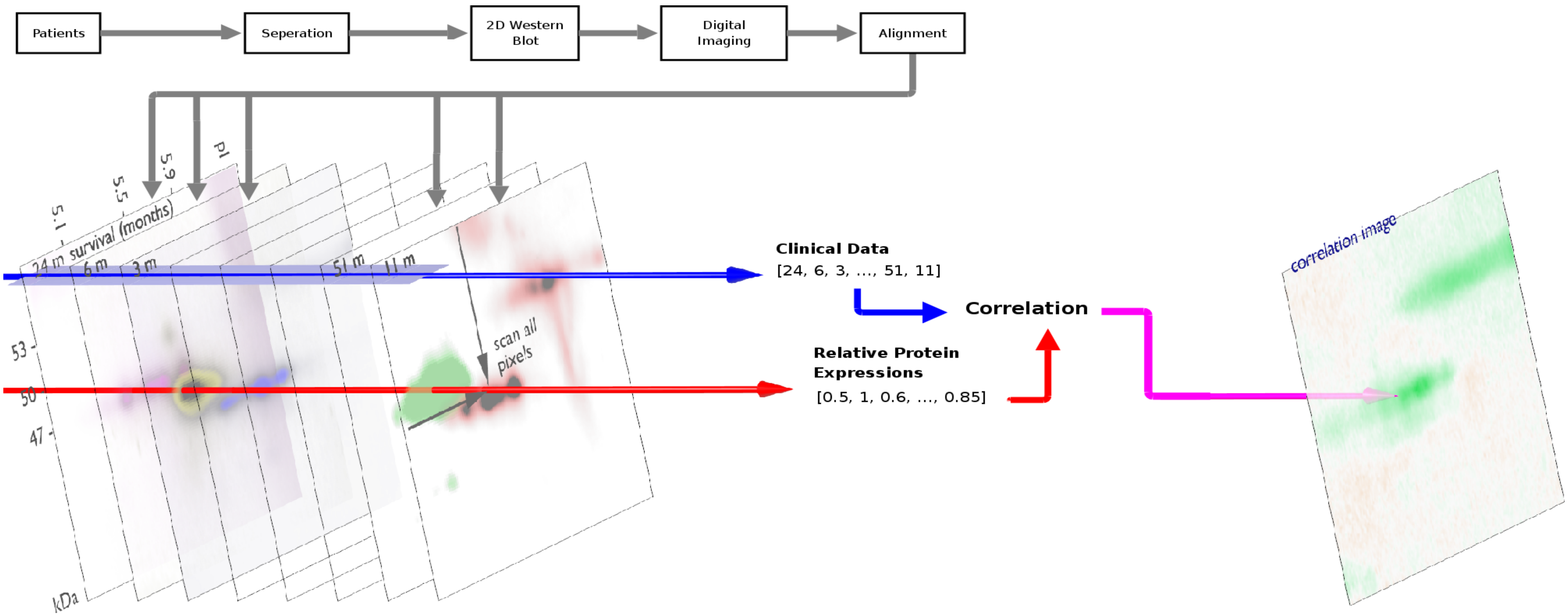
- Is there a relation between various parameters of AML/ALL cancer patients and their P53 isoforms ?

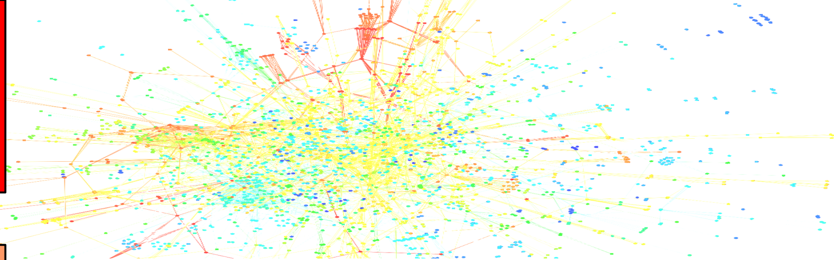
P53 Biosignatures vs Age



2D Gel Analysis

- Step 4: Mask

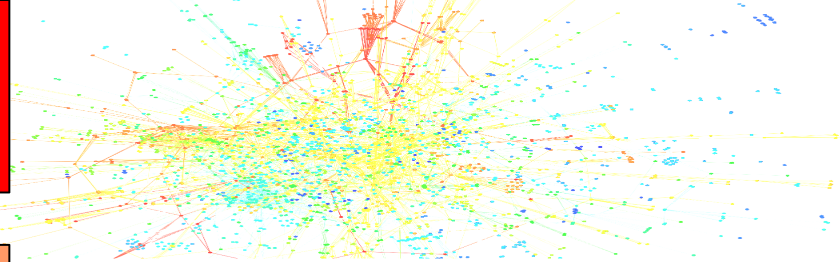




Step 4: Masking

Step 4: Masking

Correlation does not necessarily imply a causal, significant, or useful relationship. To filter out some possibly useless relations, a number of masks limit the visible correlations. The first mask removes correlations that might be occurring by coincidence: some data sets easily correlate with any other data set (significance). The second mask removes correlations that offer little useful information (E.g: a data set containing all zero's).



Step 4a: Significance

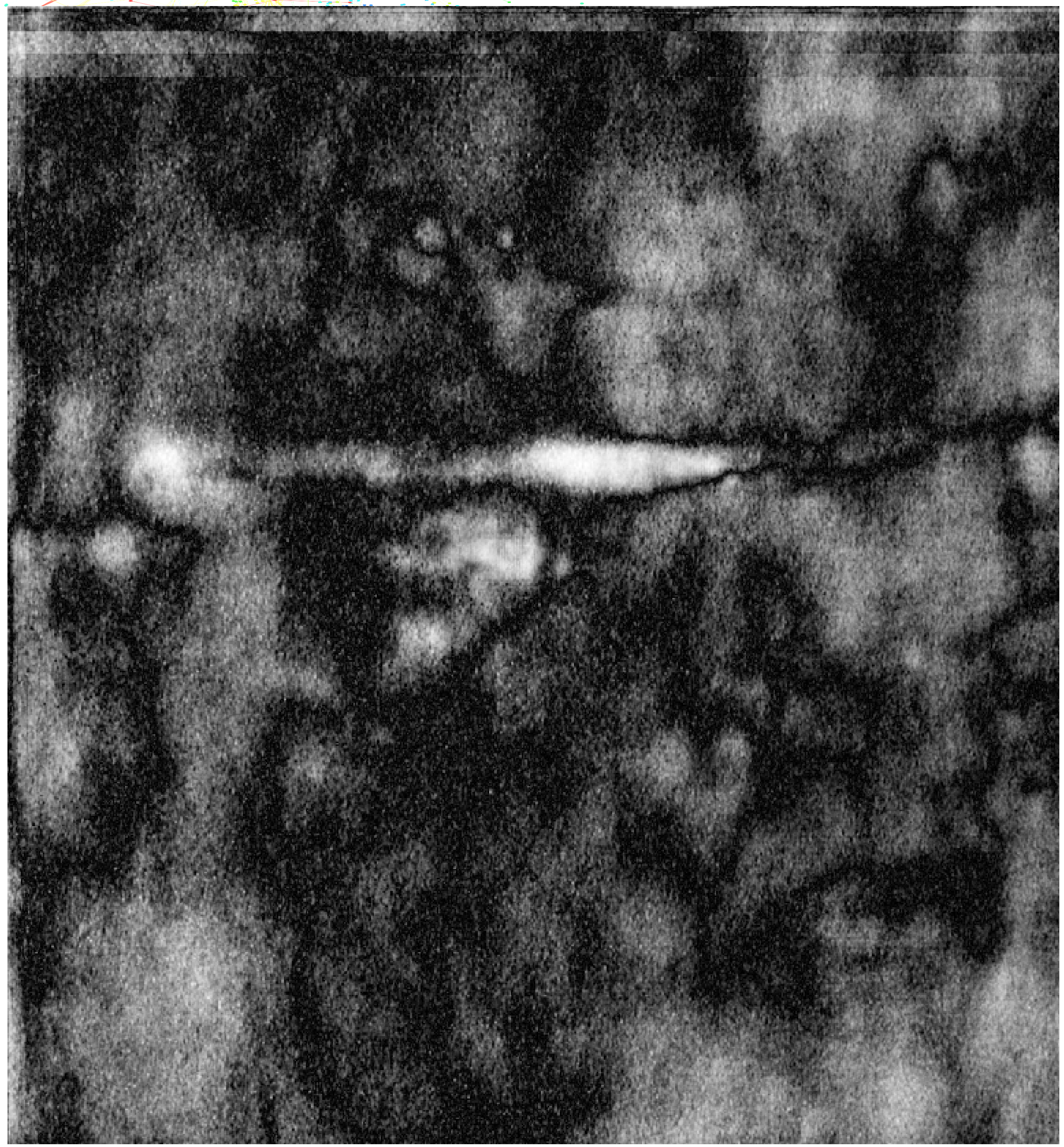
Step 4a: Significance

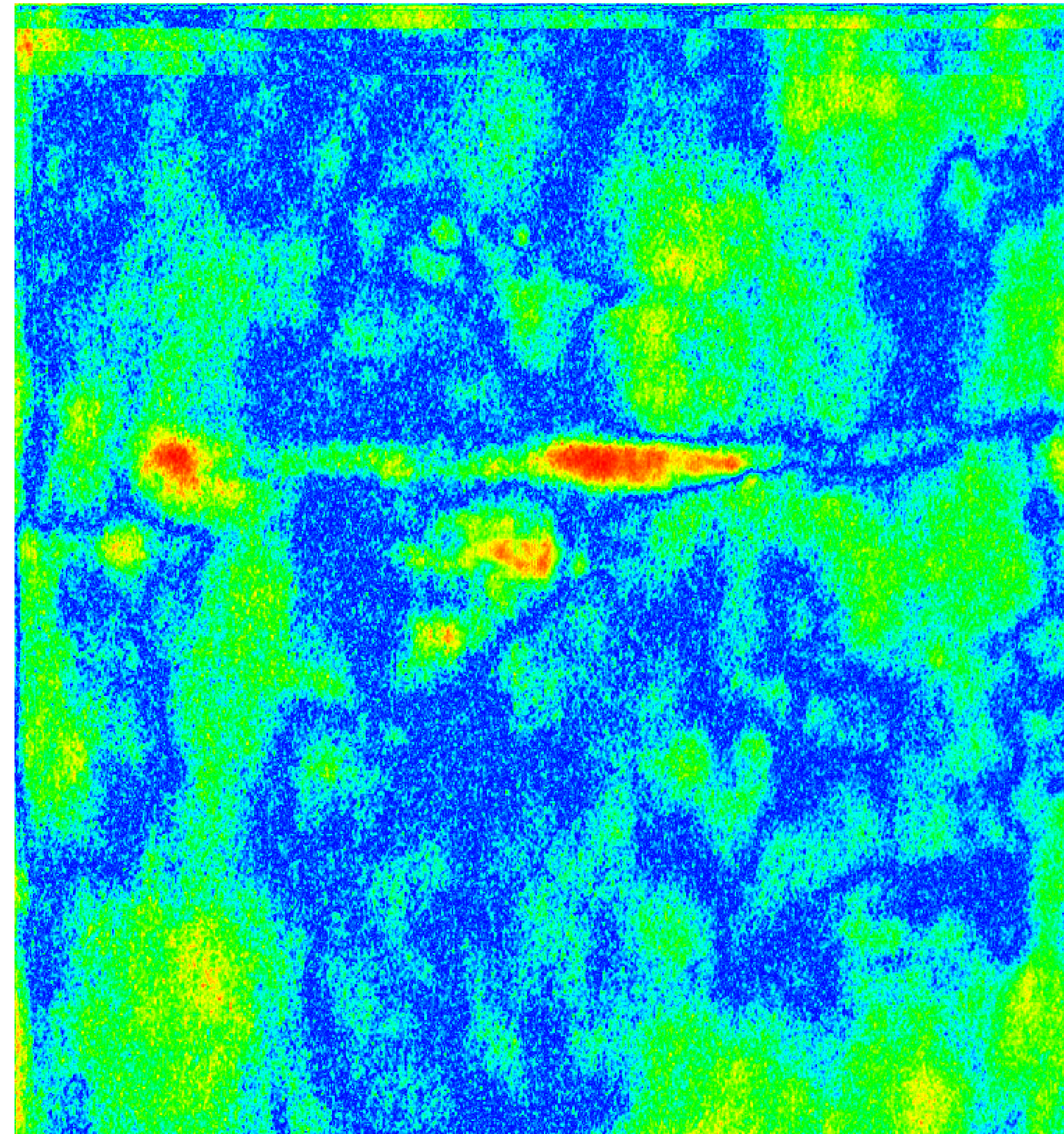
To remove correlations that have a high probability of occurring, the significance test typically associated with the Spearman correlation test was used. In this context, it is defined as

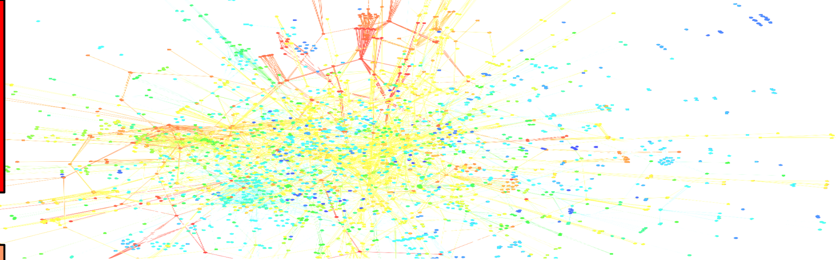
$$S_{x,y} = 1 - C_{x,y}\sqrt{\frac{n-2}{1-C_{x,y}^2}} \qquad (2)$$

If this number is close to 1 then there exists a low probability that some random data would happen to correlate with the given result set. Likewise, if this number is 0 then there exists a high probability that the correlation is coincidental.

Significance Mask







Step 4b: Variance

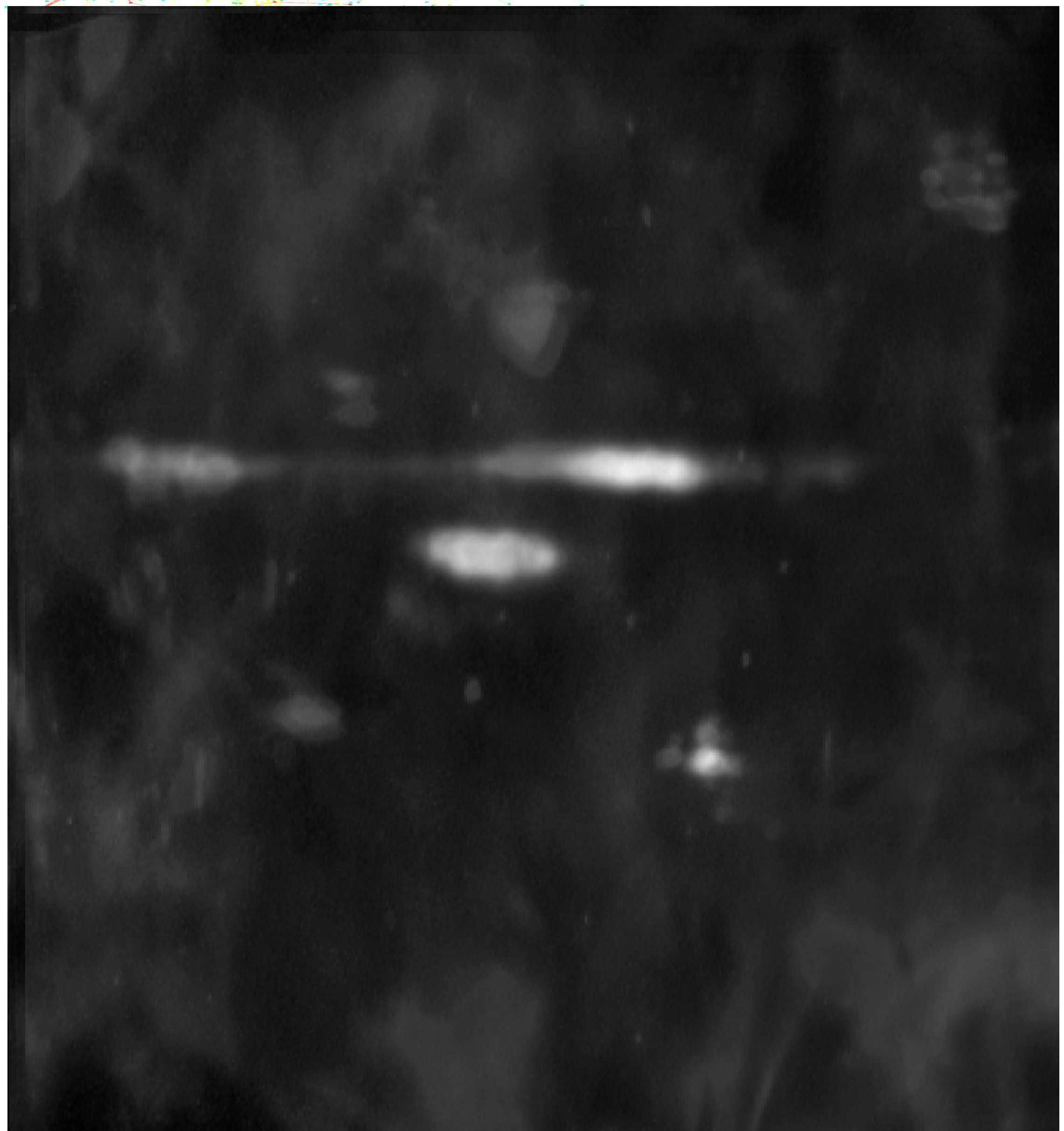
Step 4b: Variance

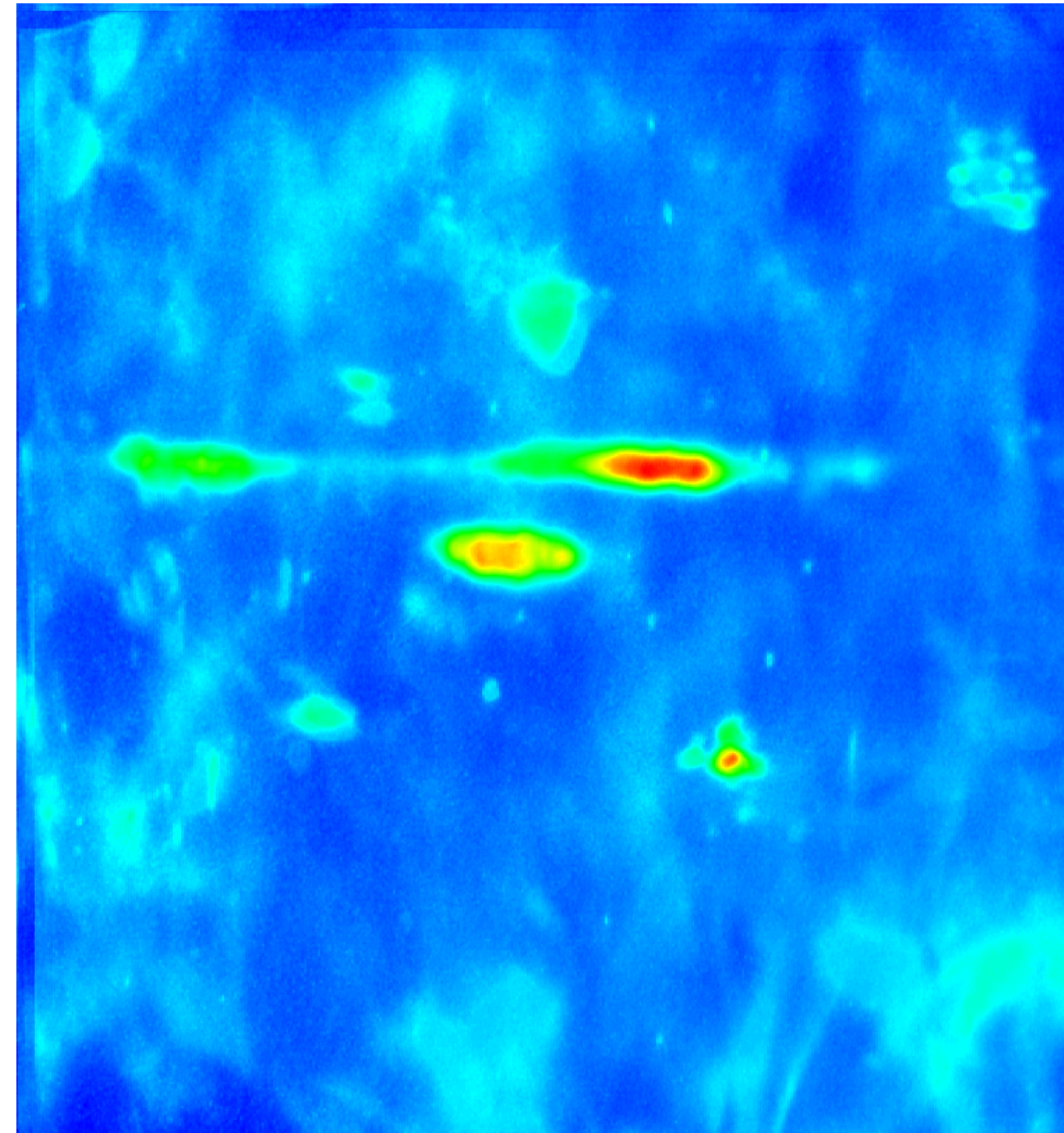
The second mask avoids strong and significant correlations that have a low biological significance because the gel intensities do not change enough. It relies on the standard deviation [54] measured on the relative, non-ranked, gel intensities

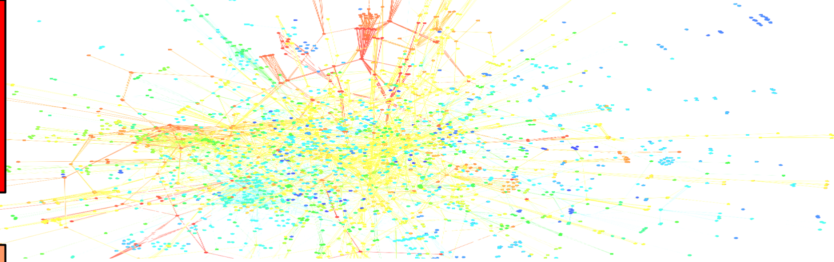
$$D_{x,y} = \frac{\sqrt{\sum_{z=0}^{n-1}\left(\frac{A^{"}_{x,y,z}}{A^{"}_{x,y,*}} - 1\right)^2}}{N} \quad (3)$$

The standard variance (or RMS) of the mean divided gels will have a large value where there is a varying gel expression. At places where the gel expression is constant this value will be zero.

Variance Mask







Step 4c: Overall Mask

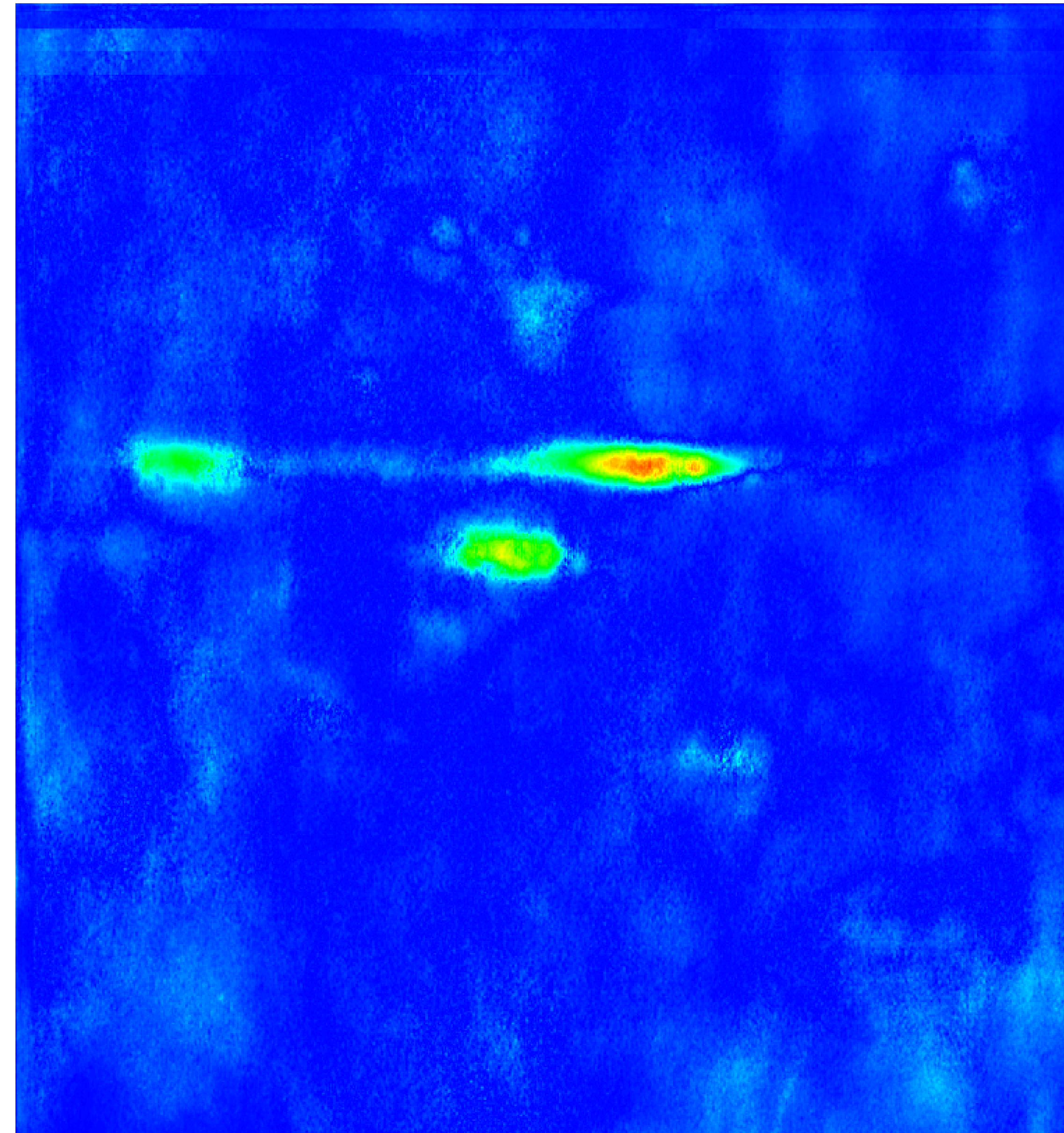
Step 4c: The masked correlation image

Multiplying the standard deviation mask (Eq. 3) with the significance mask (Eq. 2) gives a new mask that can be superimposed over the correlation image (Eq. 1).

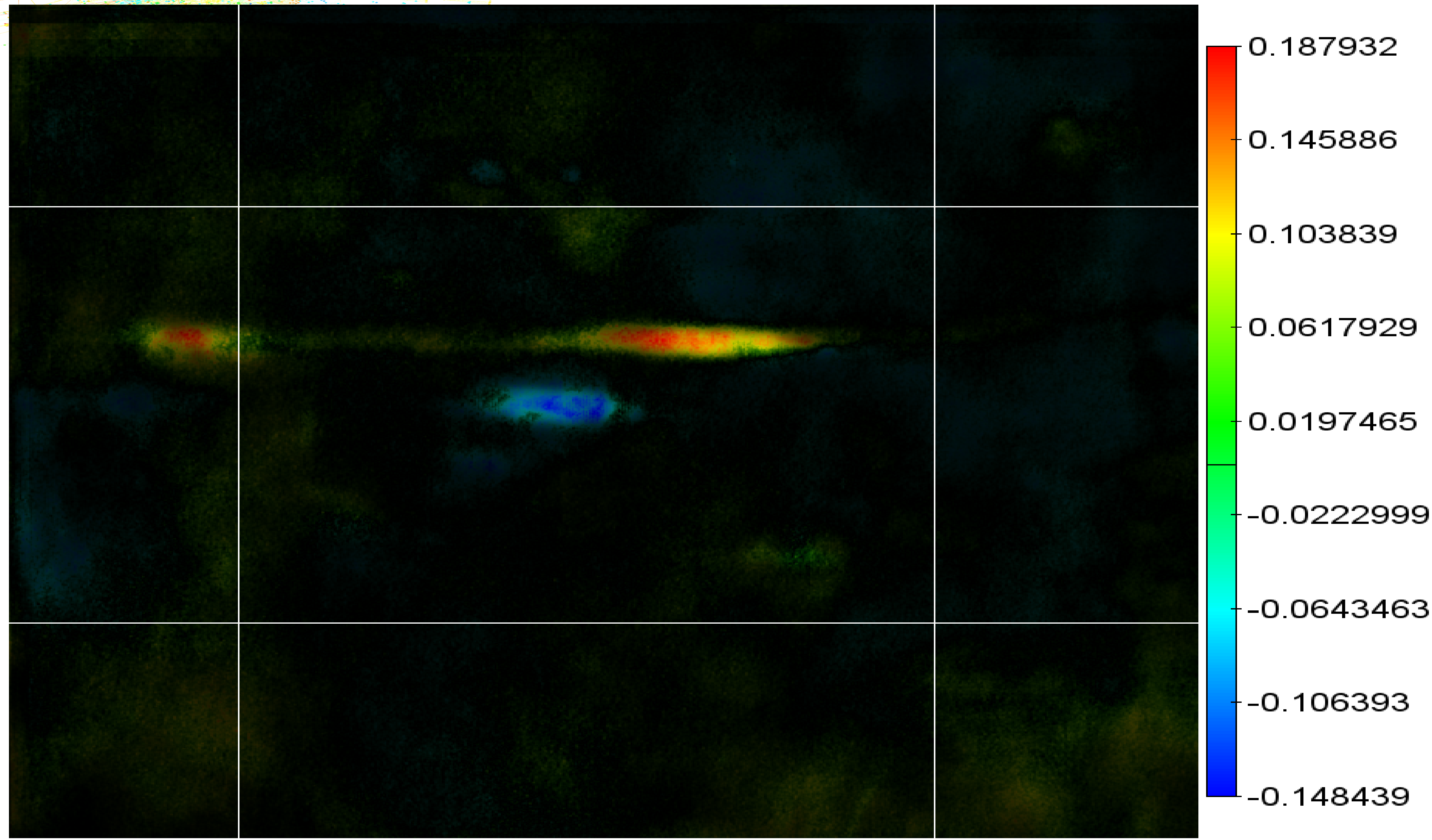
$$R = C \times S \times D$$

The pixel values of R no longer relates to the correct correlation measure. Therefore, R forms an indicator, showing position of possible interest.

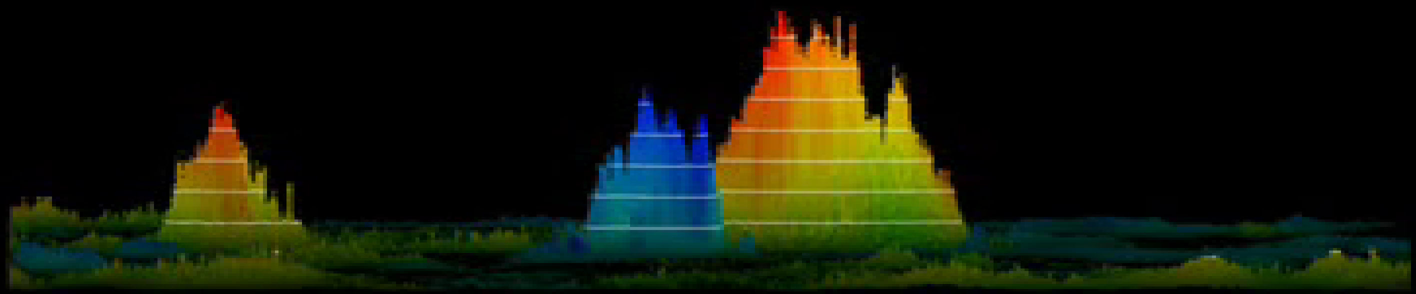
Overall Mask



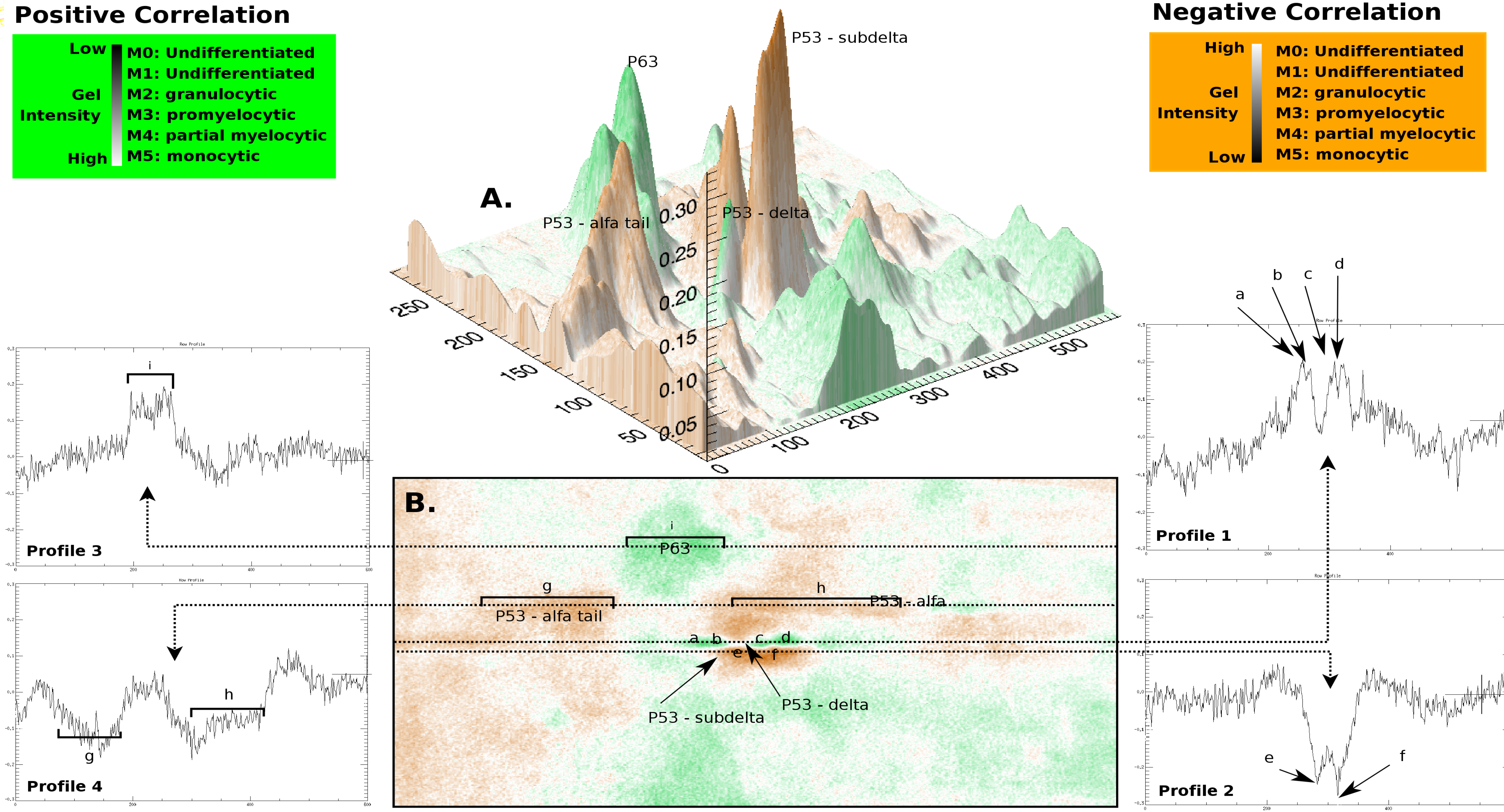
P53 Biosignature vs Age



Step 5: 3D Visualization

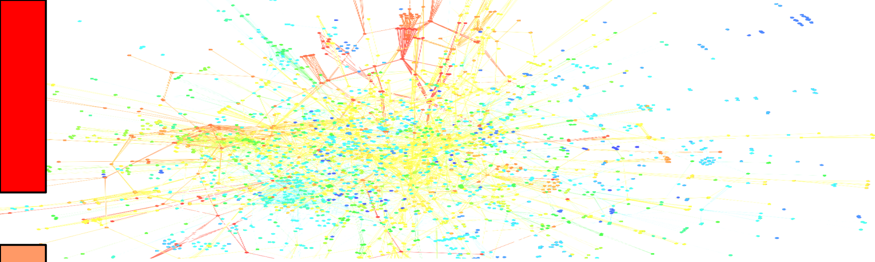


Step 5: 3D Visualization



Resource Usage

- 132 Parameters, 13 correlation sets, 128 images
- Creating the fine-tuned overlay alignment: 72h
- Computing all the correlations: 85.55h, which produced 5.8 Gb of raw data.
- Rendering of the movies: 5 hours per movie, with 1416 images: 7080h



Part 2. Maldi-TOF Artefacts

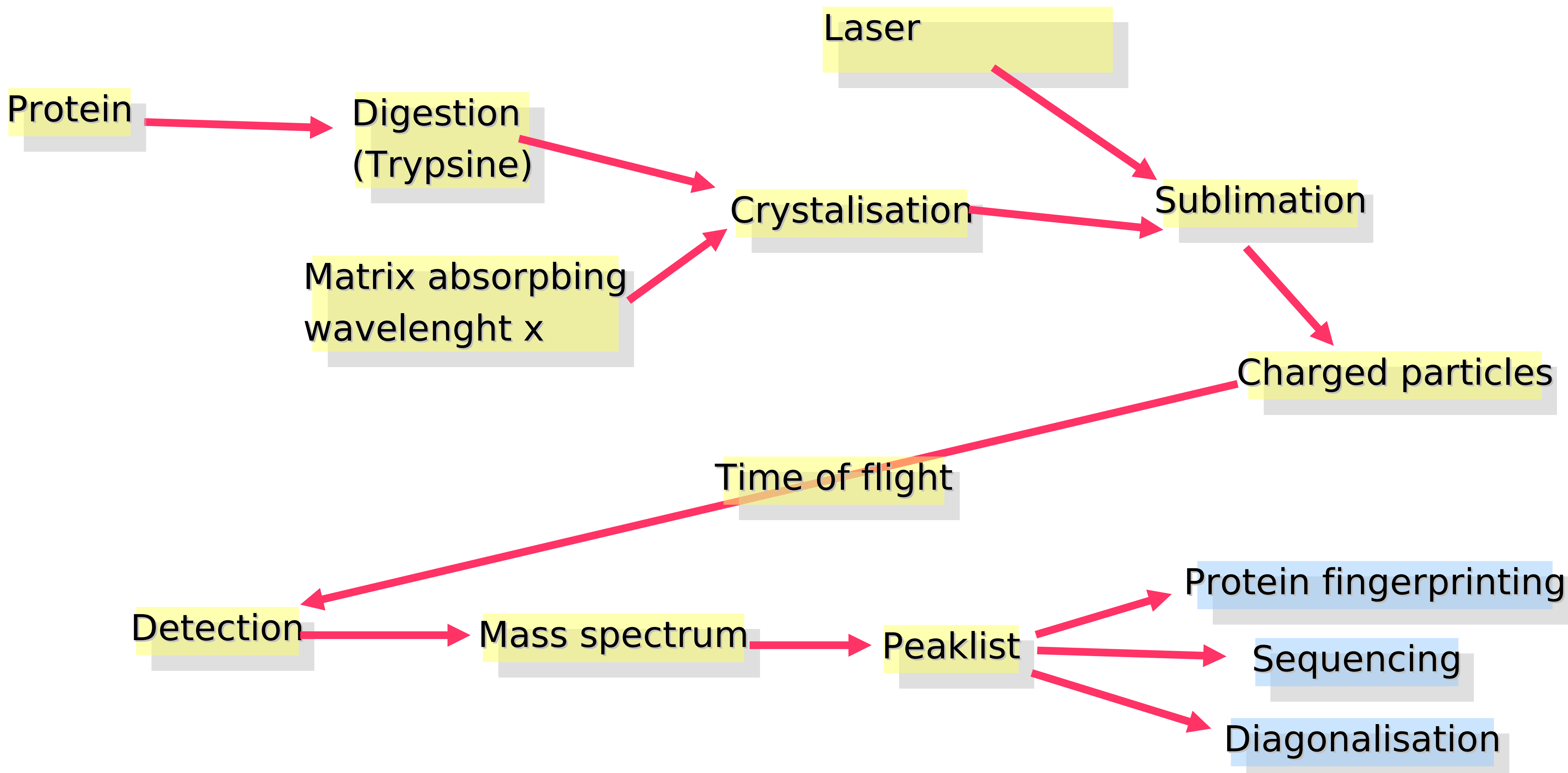
Werner Van Belle

werner.van.belle@gmail.com, werner@onlinux.be

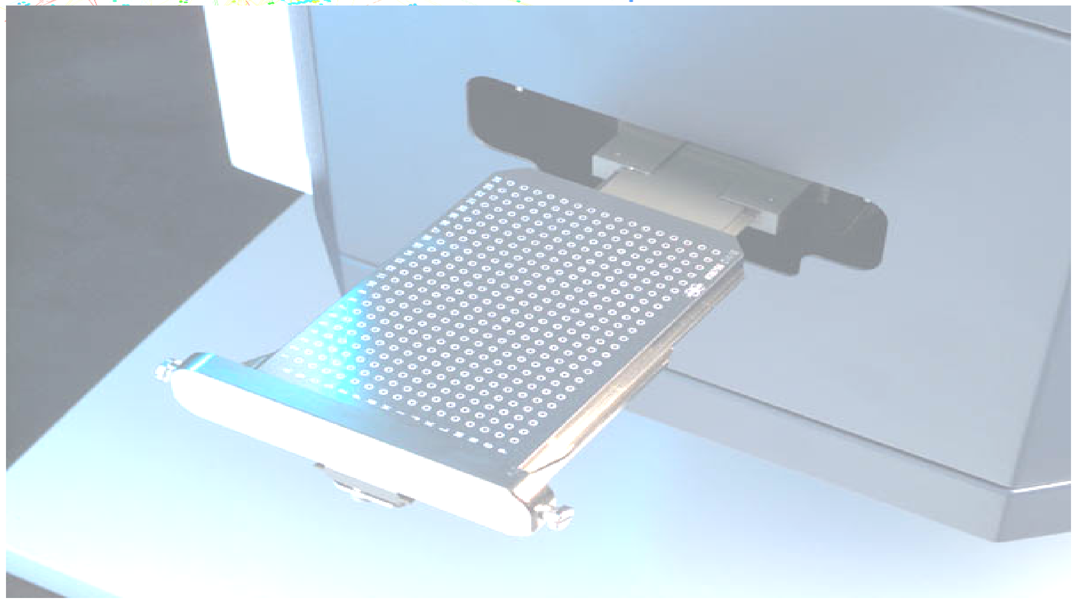
In cooperation with:Olav Mjaavatten, Kari Espolin Fladmark

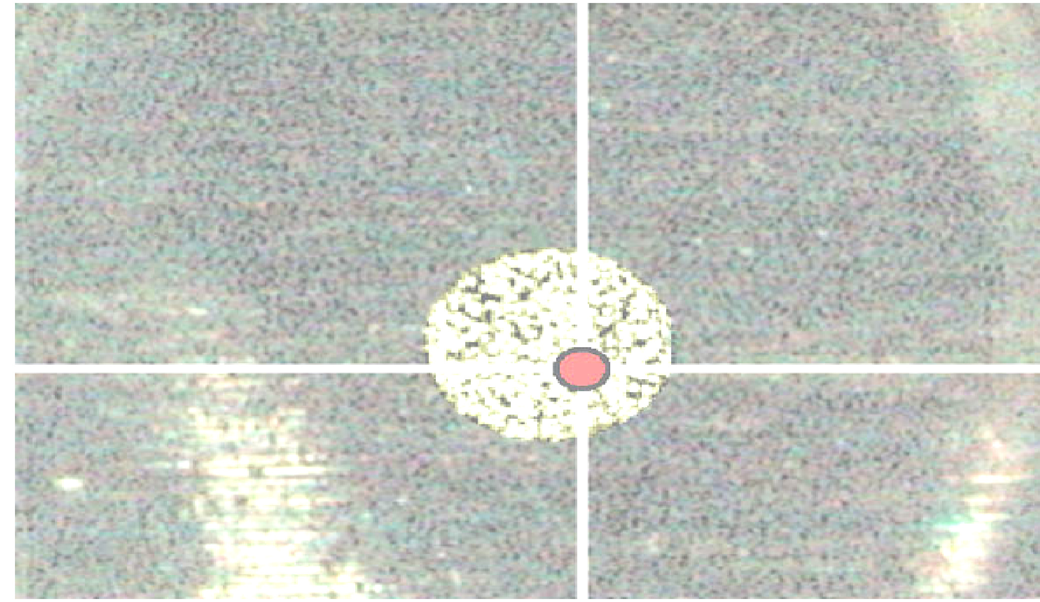
Stijn Ove Døskeland

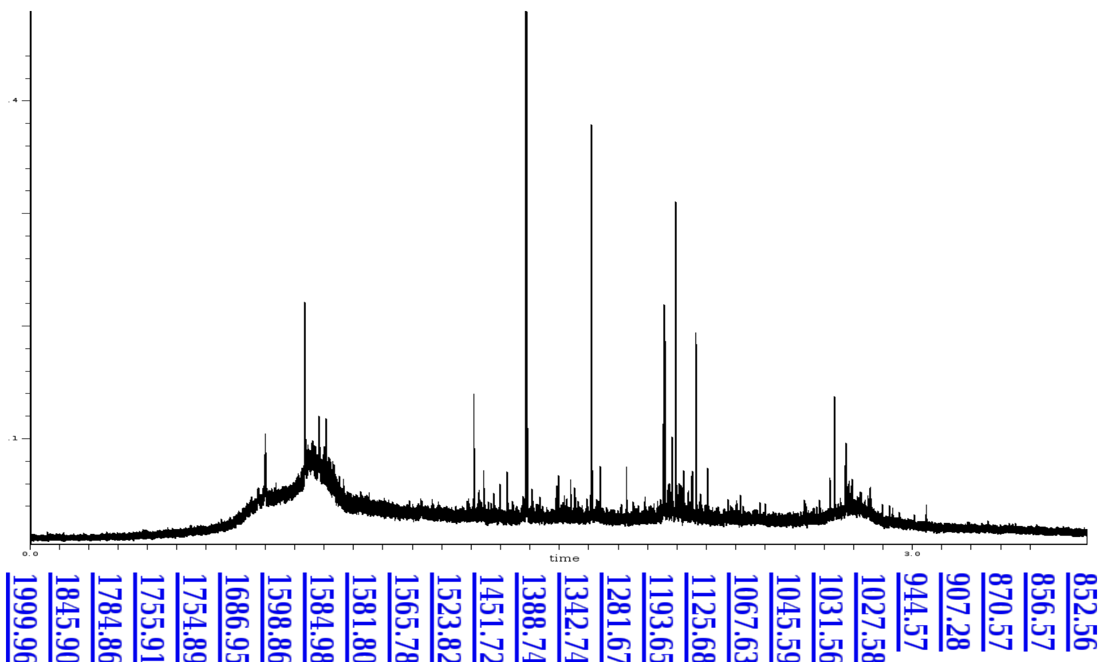
MALDI


Laser
Protein
Digestion
(Trypsine)
Crystalisation
Sublimation
Matrix absorpbing
wavelenght x
Charged particles
Time of flight
Protein fingerprinting
Detection
Mass spectrum
Peaklist
Sequencing
Diagonalisation

MALDI







gi|39932549 **Mass:** 132419 **Score:** 86 **Expect:** 0.005 **Queries matched:** 13

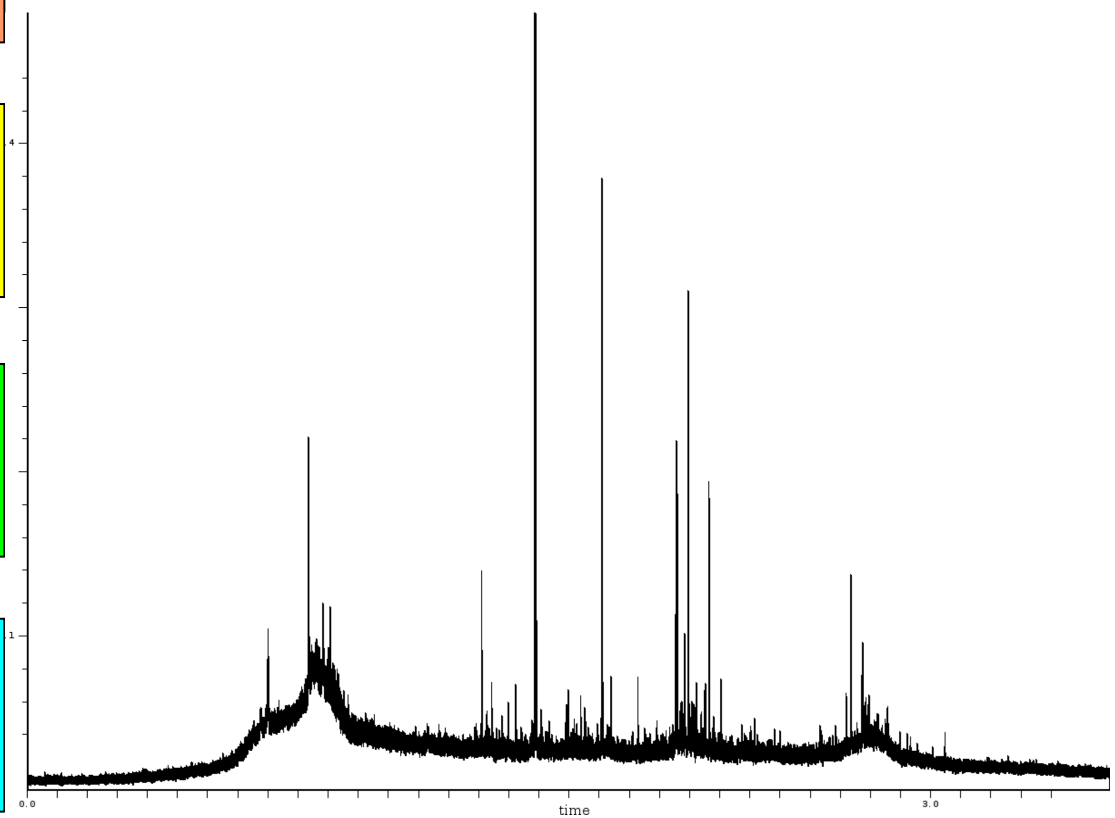
[Segment 2 of 2] Neuroblast differentiation associated protein AHNAK (Desmoyokin)

Observed	Mr(expt)	Mr(calc)	Delta	Start		End	Miss	Peptide
856.57	855.57	855.49	0.08	1155	-	1161	1	FSLFKSK
1067.63	1066.63	1066.58	0.05	267	-	276	1	GPEVDLKGPR
1125.68	1124.67	1124.58	0.09	412	-	421	1	MPKIHMSGPK
1388.74	1387.74	1387.68	0.06	886	-	899	0	SSGCDVNLPGVNVK
1523.82	1522.82	1522.86	-0.04	233	-	246	1	APKISIPDVDLDLK
1565.78	1564.77	1564.81	-0.04	444	-	459	0	APDVDVNIAGPDAALK
1584.98	1583.98	1583.84	0.14	124	-	138	1	GDVDVTLPKVEGDLK
1598.86	1597.86	1597.89	-0.04	1261	-	1275	0	VGIQLPEVELSVSTK
1686.95	1685.94	1685.89	0.06	1000	-	1016	1	FAGGLHFSGPKVEGGVK
1754.89	1753.88	1753.96	-0.08	286	-	302	1	LSGPSLKMPSLEISAPK
1755.91	1754.91	1754.91	-0.00	532	-	547	1	MPDVDISVPKIEGDLK
1784.86	1783.85	1783.91	-0.06	754	-	771	1	GPSLQGDLAVSGDIKCPK
1999.96	1998.96	1998.91	0.05	470	-	485	1	TMFGKMYFPDVEFDIK + 2 Oxidation (M)

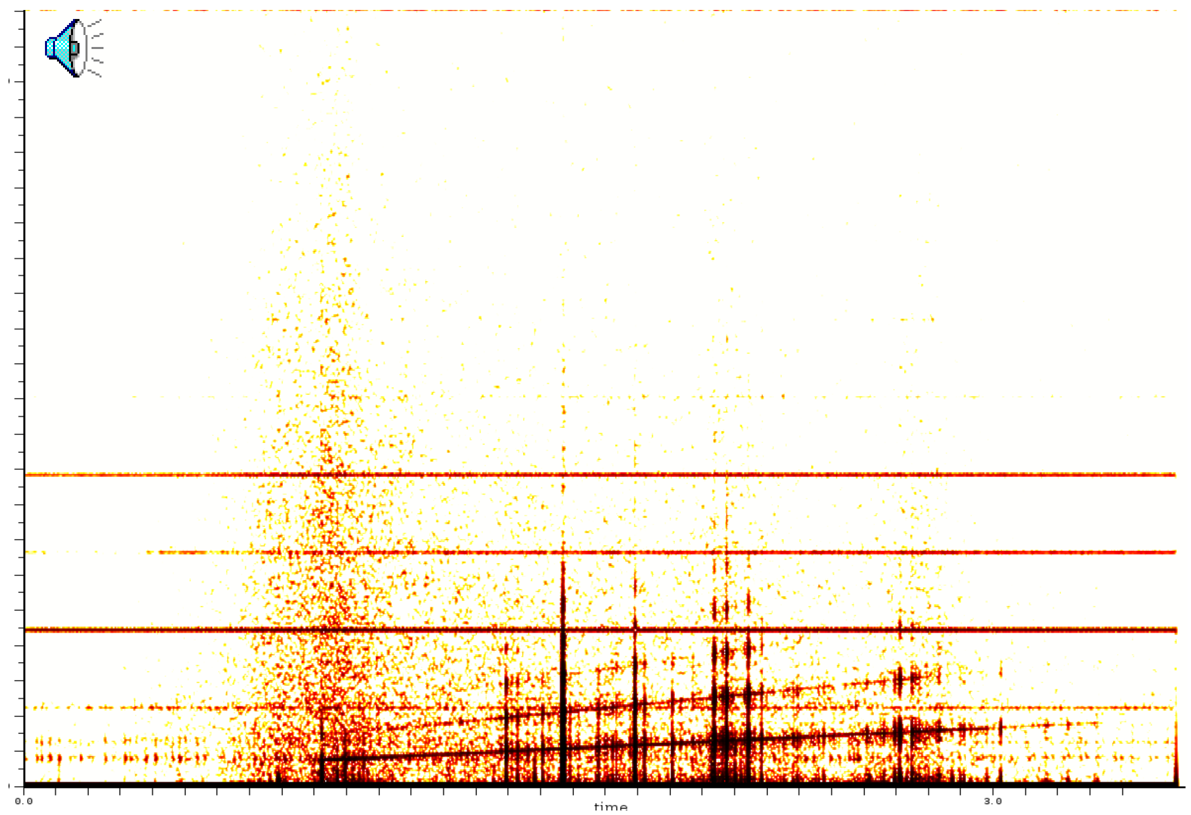
No match to: 852.56, 870.57, 907.28, 944.57, 1027.58, 1031.56, 1045.59, 1193.65, 1281.67,

Artefacts I

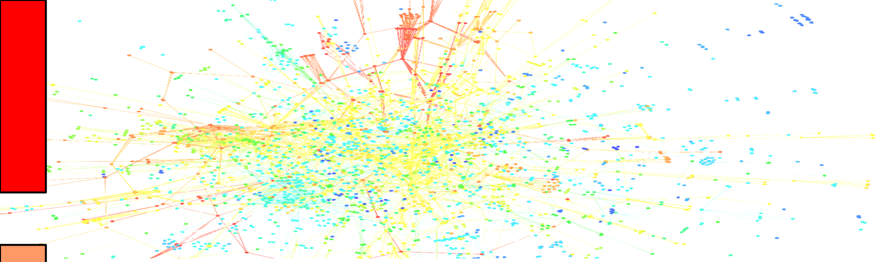
- Static Noise



Mass spectrum output

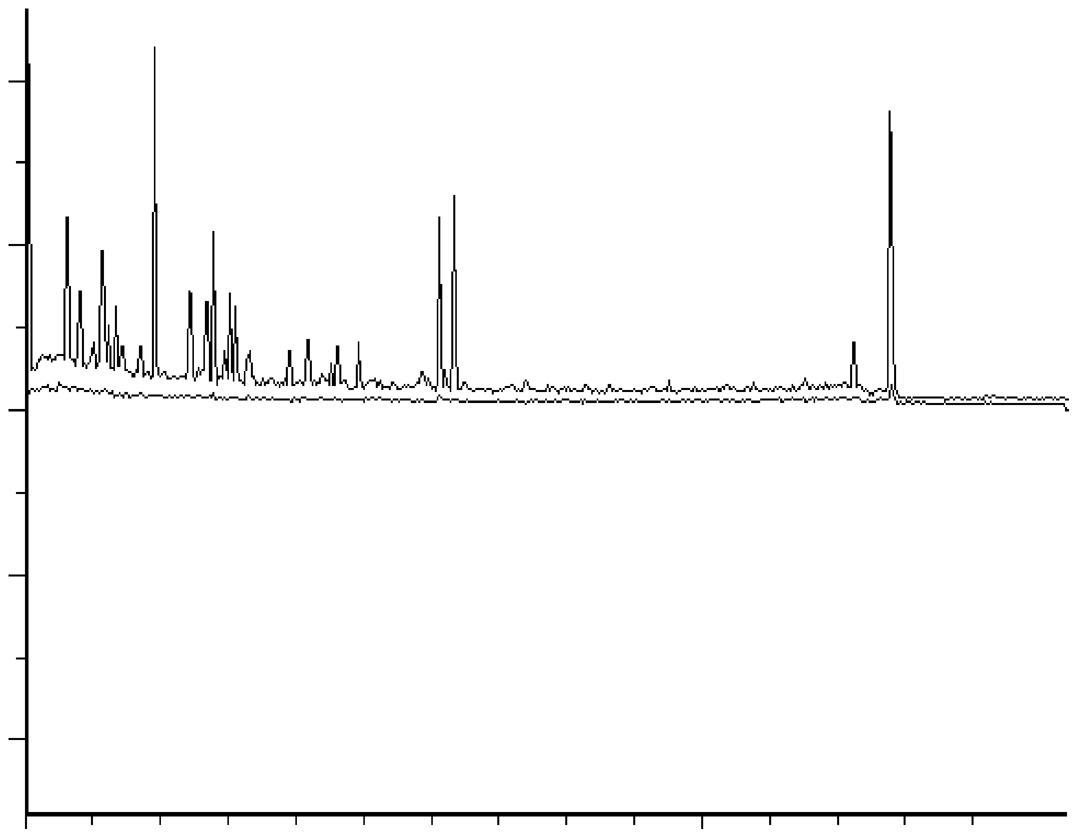


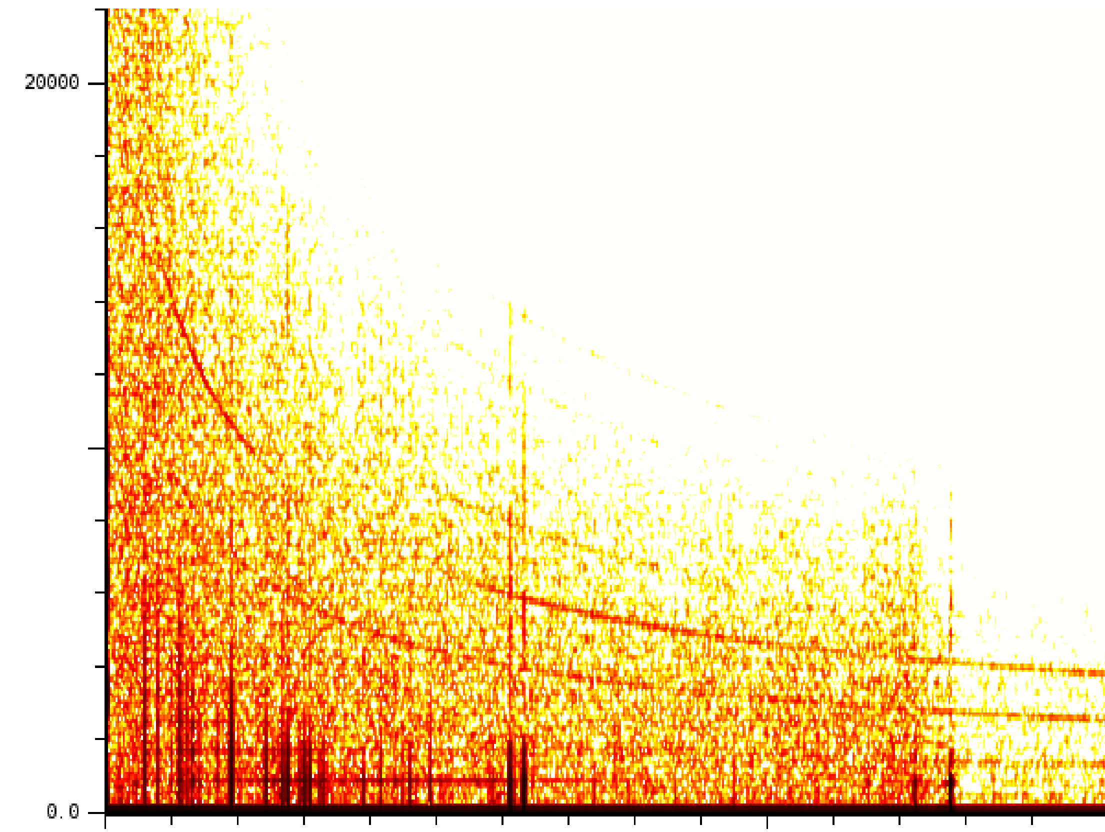
Frequency Analysis



Artefacts II

- Sweeping Tones in LIFT





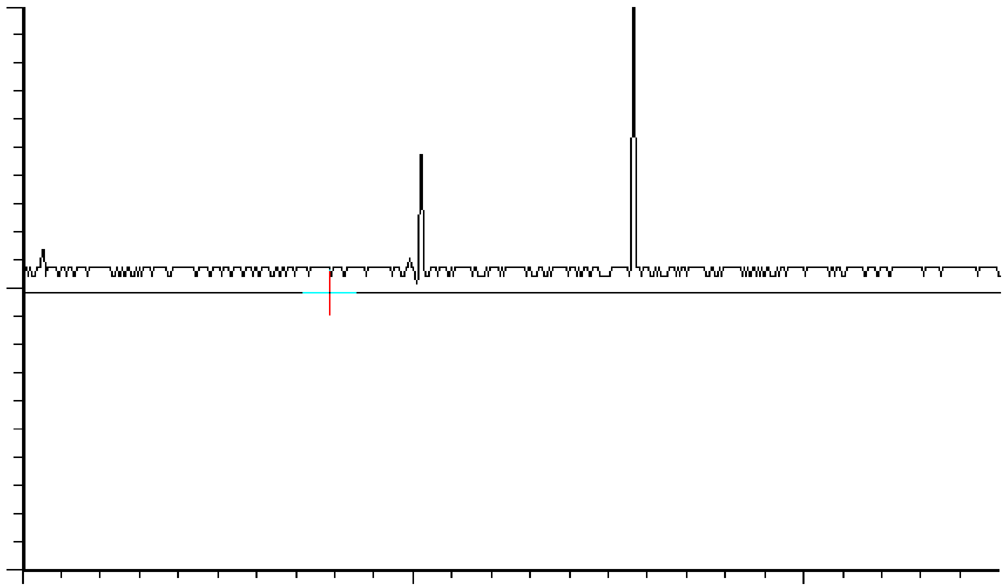
Mass spectrum output

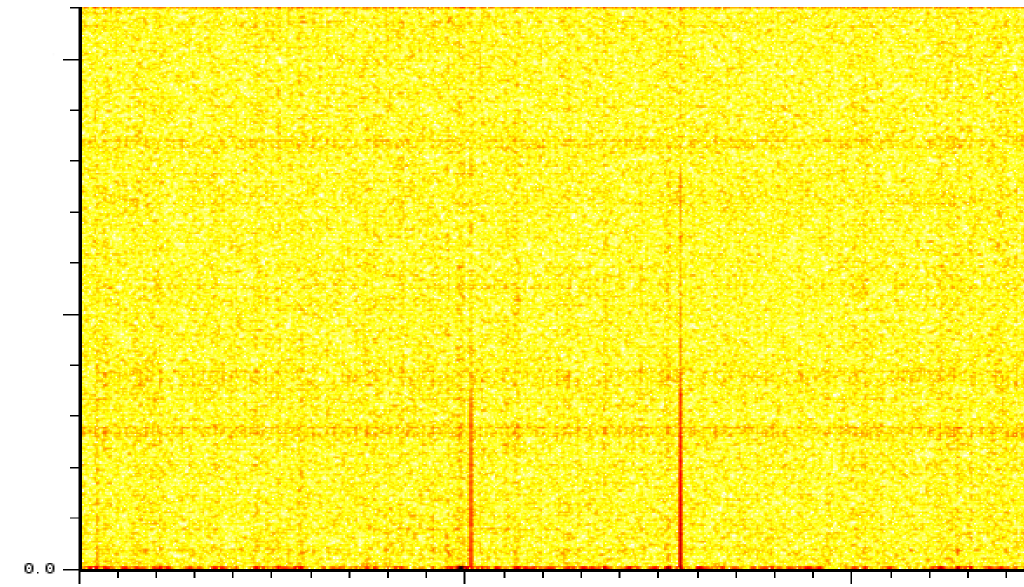
Frequency Analysis

Artefacts III

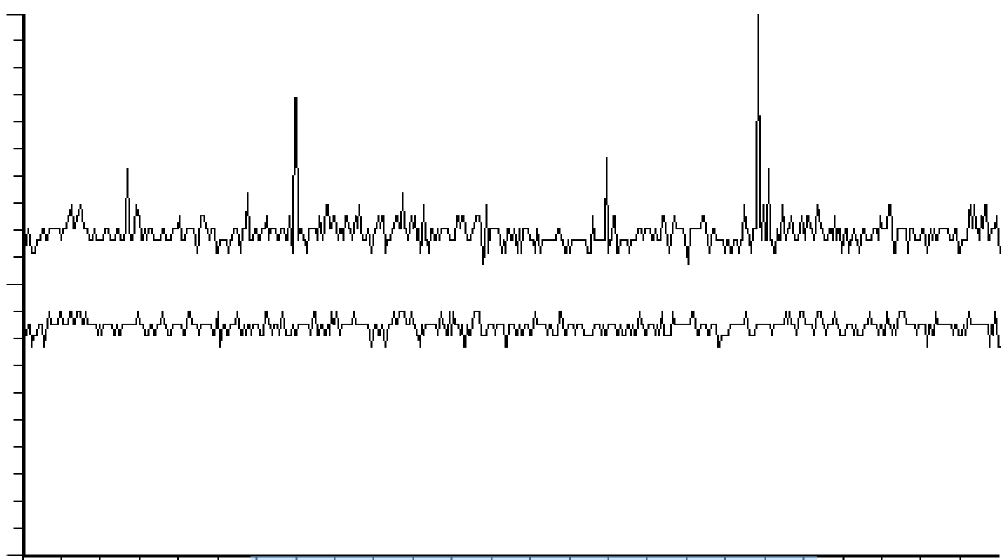
- Coherent Pulses

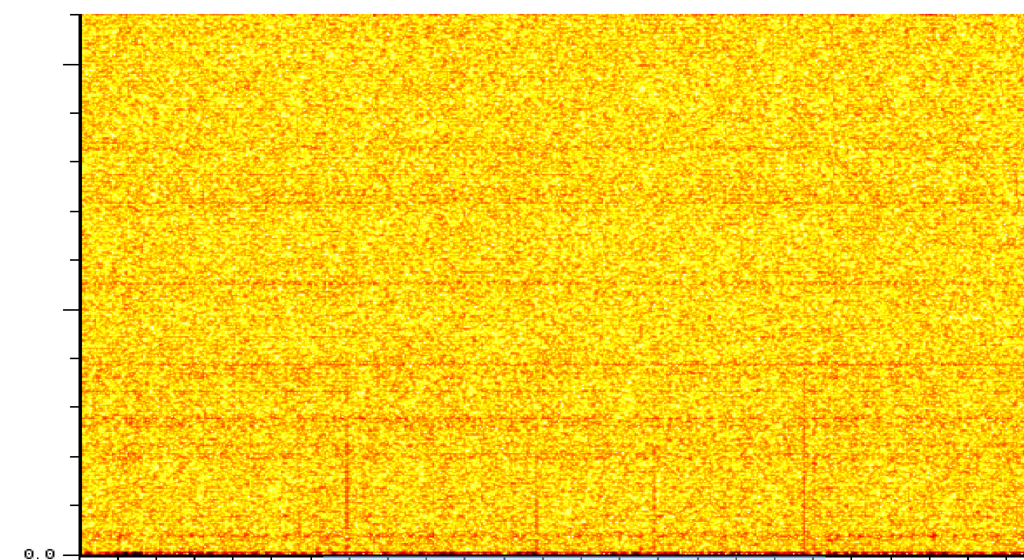
1 Shot





10 Shots





Mass spectrum output

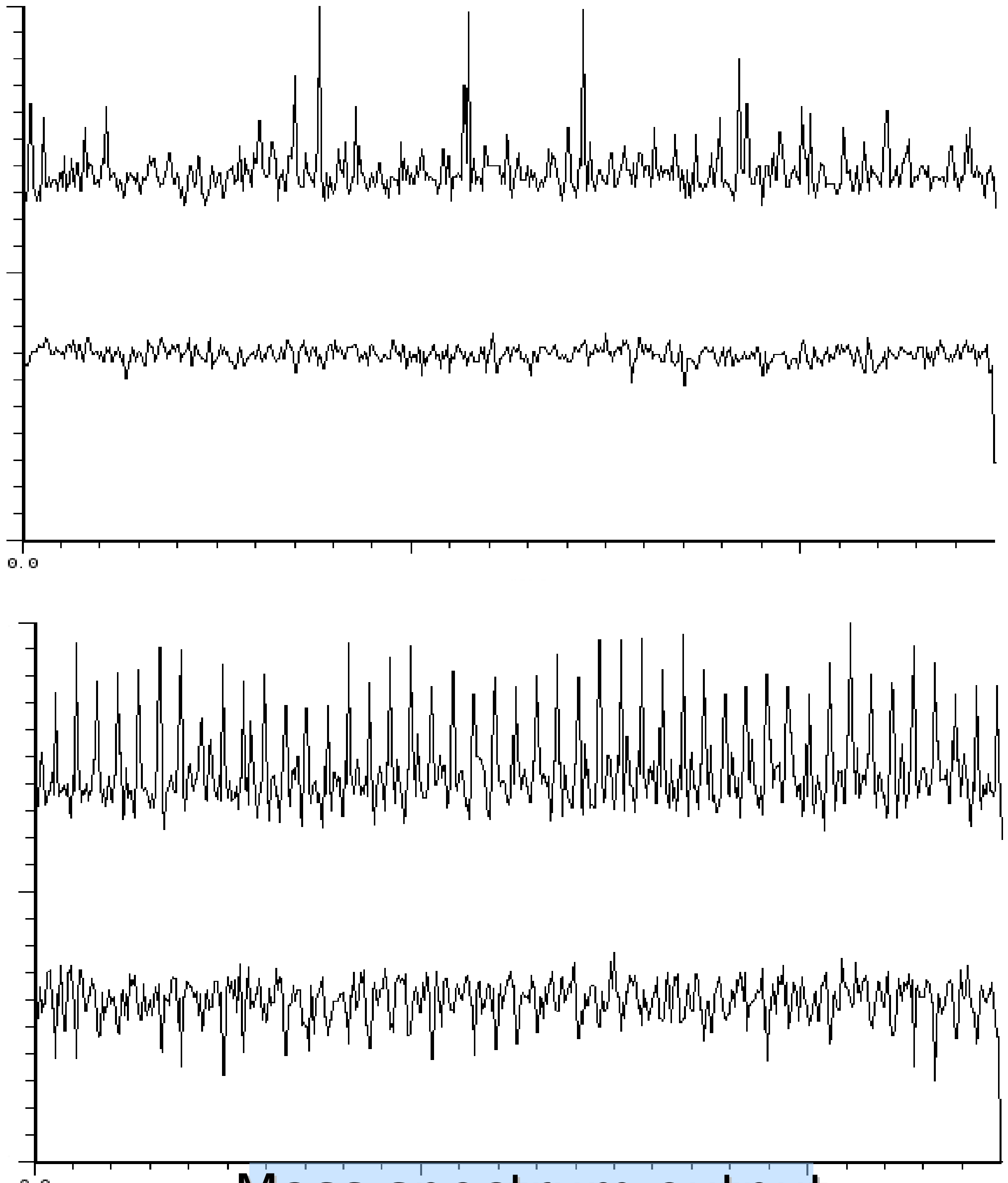
Frequency Analysis

Artefacts III

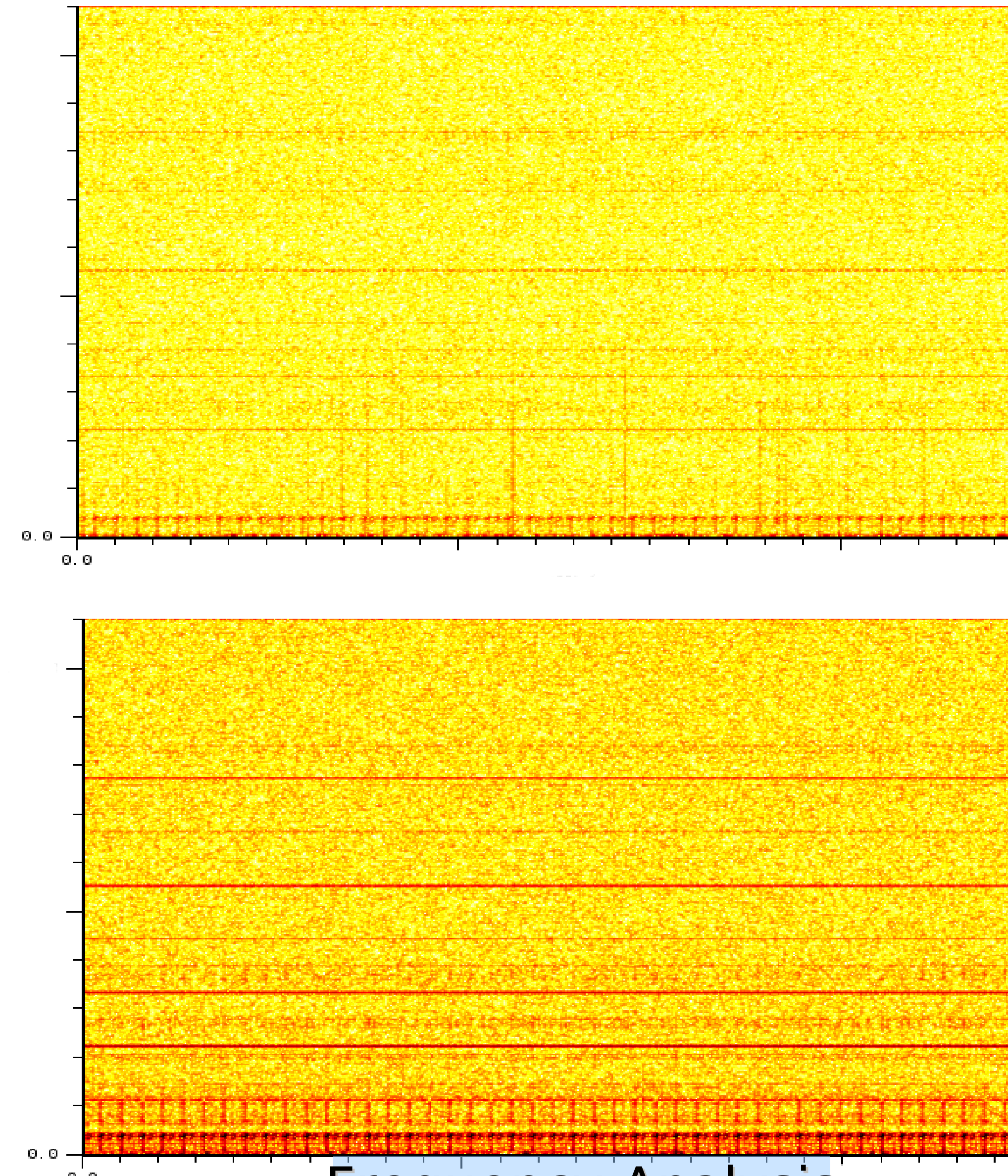
- Coherent Pulses

100 Shot

1000 Shots



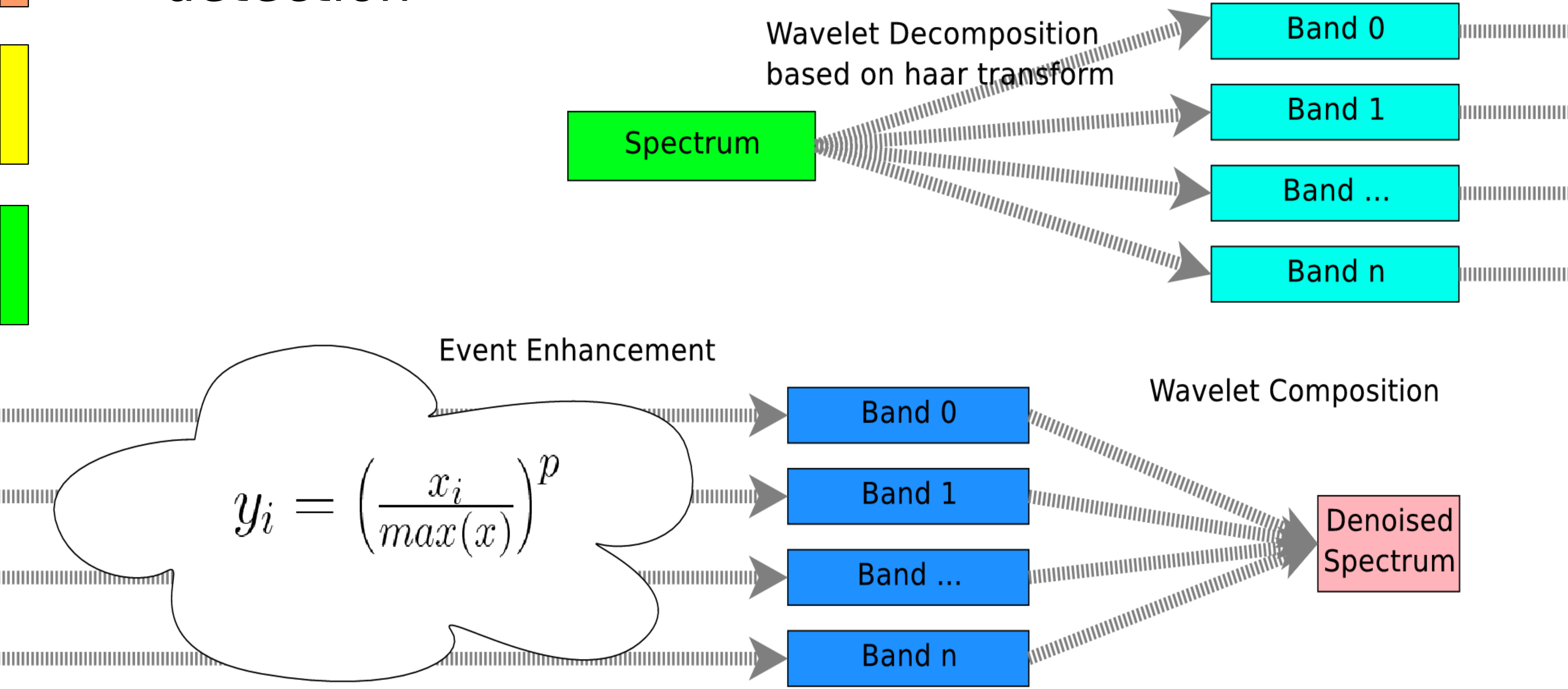
Mass spectrum output

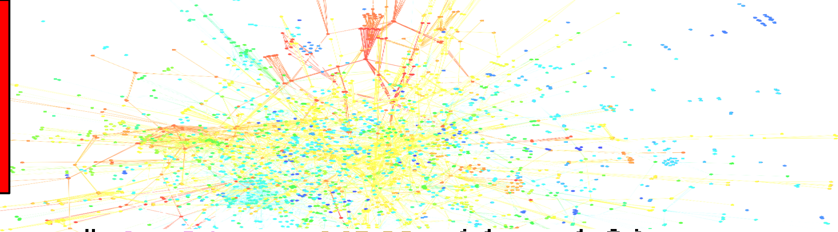


Frequency Analysis

Denoising method

- Multi-rate spectral analysis excelent tool for event detection





Denoising method

```
#define MAX (1<<18)

int main(int argc, char* argv[])
{
   double pos[MAX],data[MAX], m[MAX], d[MAX];
   int i,b,t;
   double a, mean, diff, dev;
   for(i = 0 ; i < MAX ; i ++) data[i]=0;
   FILE * f = fopen(argv[1],"rb");
   for(t=0;!feof(f);t++)
      fscanf(f,"%g %g\n",pos+t,data+t);
   fclose(f);
   int max = MAX;
   while(max >=2)
      {
         for(i = 0; i < max / 2 ; i ++)
            {
               diff = data[ i * 2 + 1 ] - data[ i * 2 ] ;
               mean = data[ i * 2 + 1 ] + data[ i * 2 ] ;
               diff/=sqrt(2.0);
               mean/=sqrt(2.0);
               m [ i ] = mean ;
               d [ i ] = diff ;
            }
         for ( i = 0 ; i < max / 2; i ++ )
            {
               data [ i ] = m [ i ] ;
               data [ i + max / 2 ] = d [ i ] ;
            }
         max /= 2 ;
      }
```

Initialisation

Haar wavelet Decomposition

Denoising method

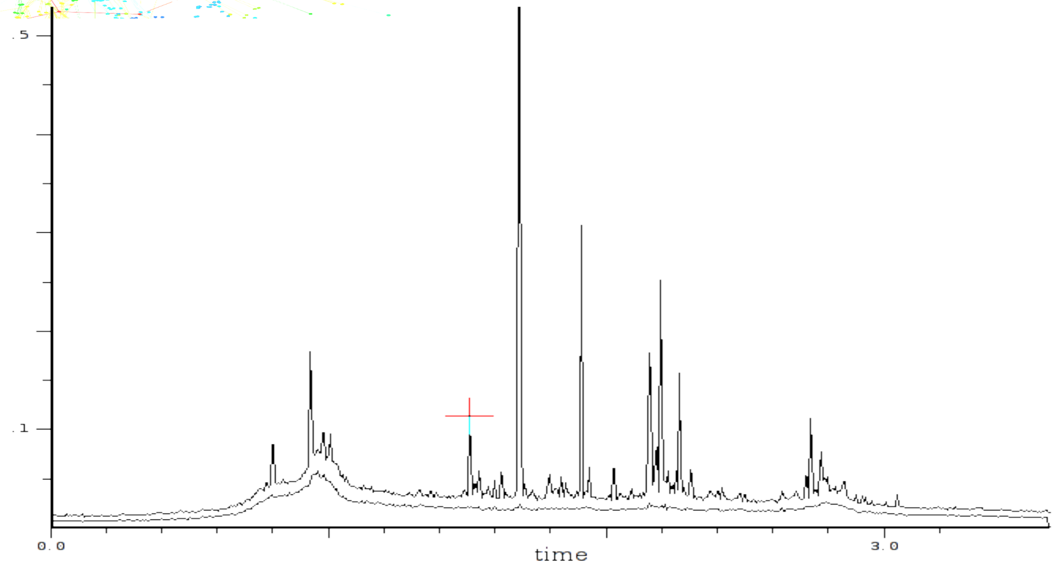
```
   for(mean = i = 0 ; i < MAX ; i ++)
     mean+=fabs(data[i]);
   mean/=MAX;
   for(dev = i = 0 ; i < MAX ; i ++)
     dev+=(data[i]-mean)*(data[i]-mean);
   dev=sqrt(dev/MAX);
   double clip = mean+0.90*dev;
   for(int i = 0 ; i < MAX ; i ++)
     if (fabs(data[i])<clip)
       {
         data[i]/=clip;
         data[i]*=data[i]*data[i];
         data[i]*=clip;
       }
   while(max <= MAX)
     {
       for ( i = 0 ; i < max / 2; i ++ )
         {
           m [ i ] = data [ i ] * sqrt(2);
           d [ i ] = data [ i + max / 2 ] * sqrt(2);
         }
       for( i = 0; i < max / 2 ; i ++ )
         {
           data[ i * 2 + 1 ] = ( m [ i ] + d [ i ] ) / 2 ;
           data[ i * 2 + 0 ] = ( m [ i ] - d [ i ] ) / 2 ;
         }
       max *= 2;
     }
   f = fopen(argv[2],"wb");
   for(i=0;i<t;i++)
     fprintf(f,"%g %d\n",pos[i],data[i]);
   fclose(f);
}
```

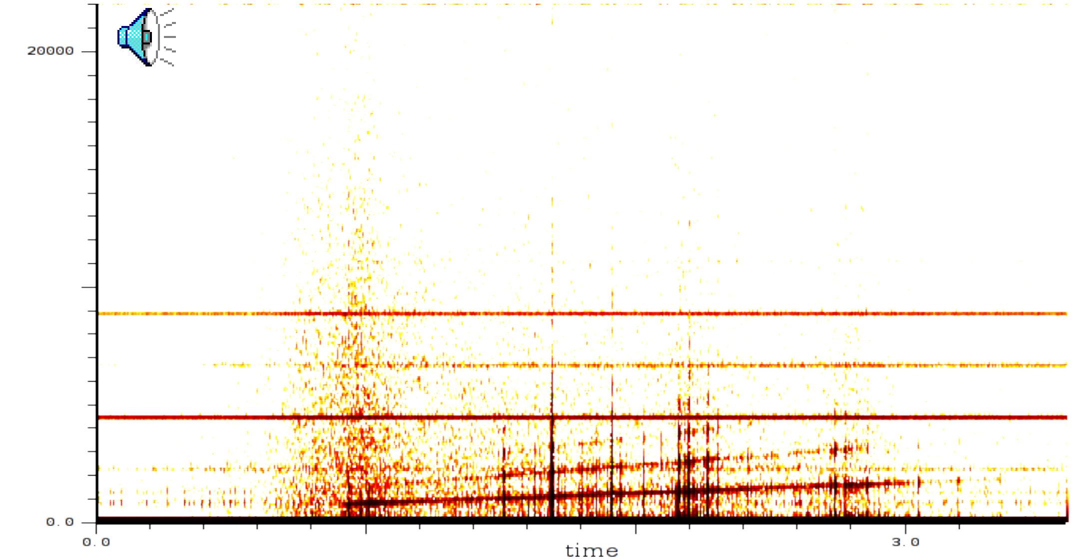
Normalisation & Event Enhancement

Wavelet Composition & Saving

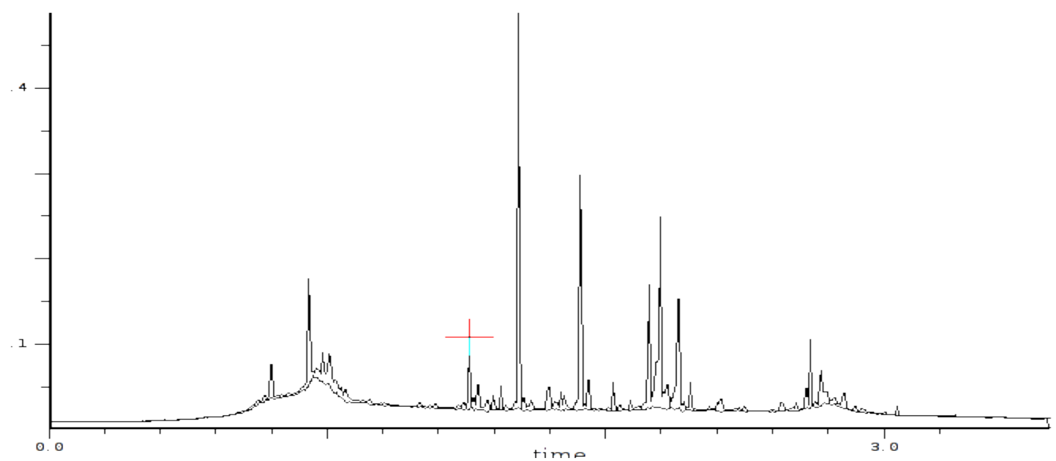
Wavelet Enhancement – Global

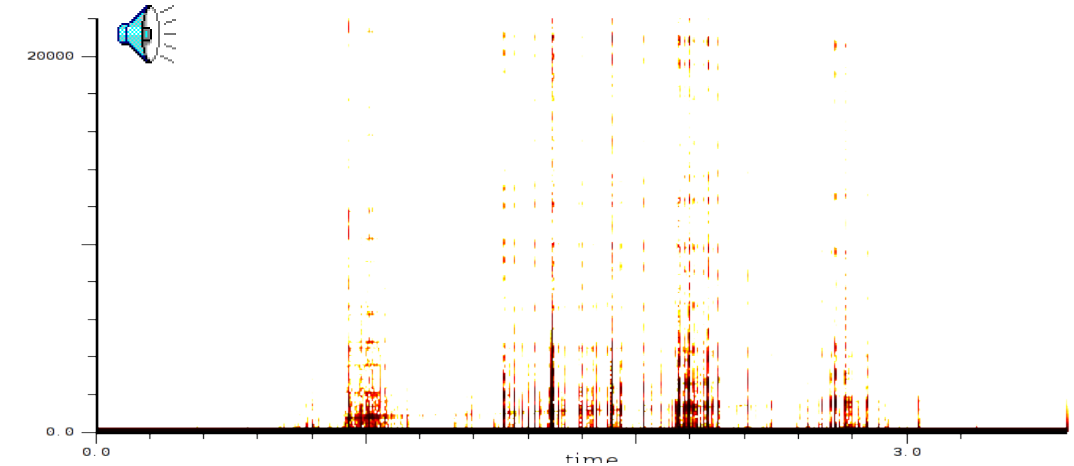
Original





Denoised





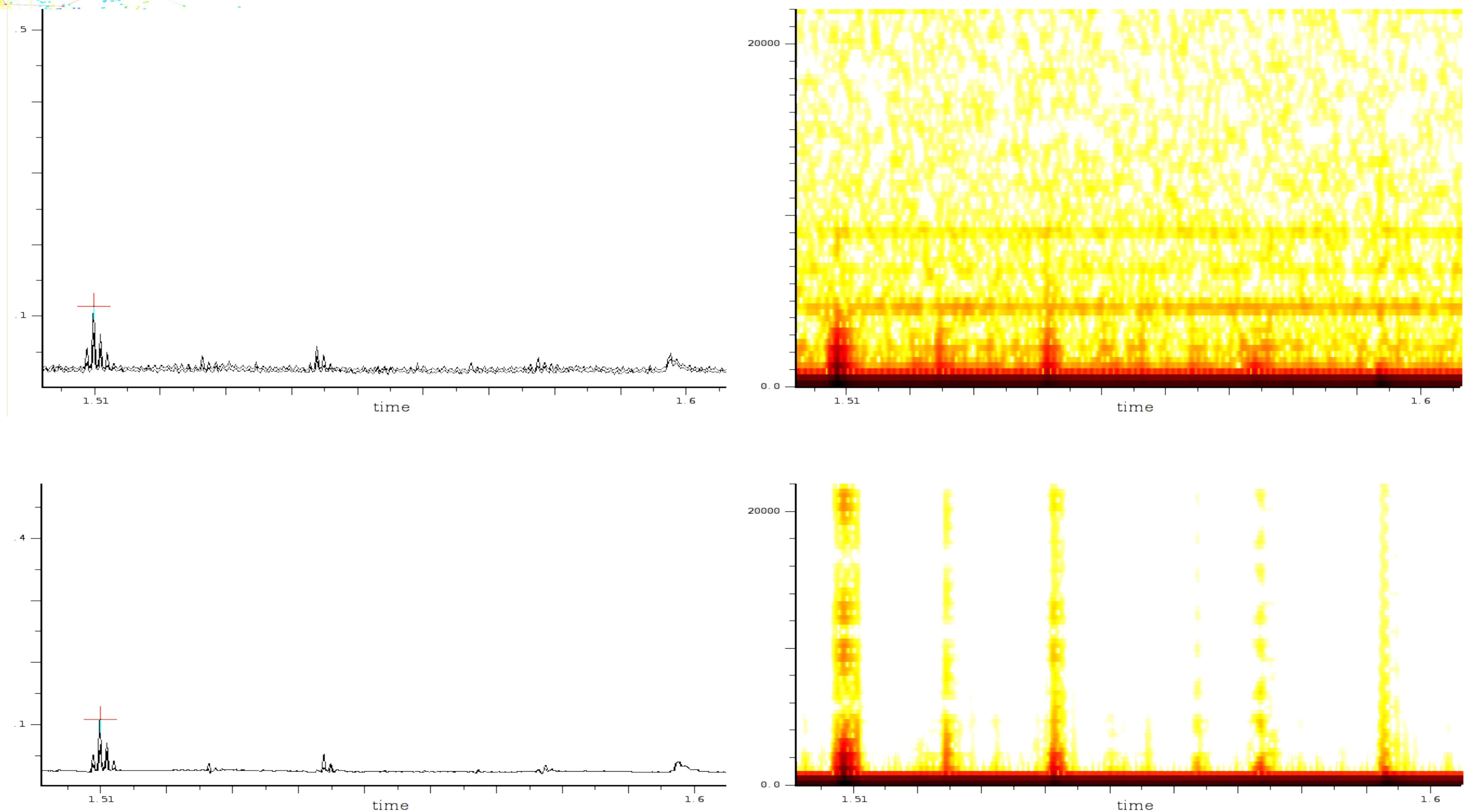
Mass Spectrum

Frequency Analysis

Wavelet Enhancement – Local

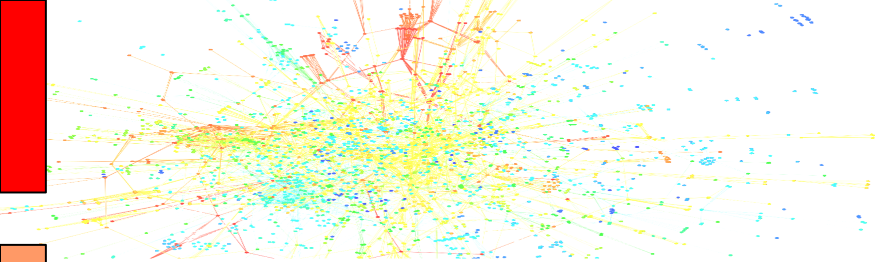
Original

Denoised



Mass Spectrum

Frequency Analysis



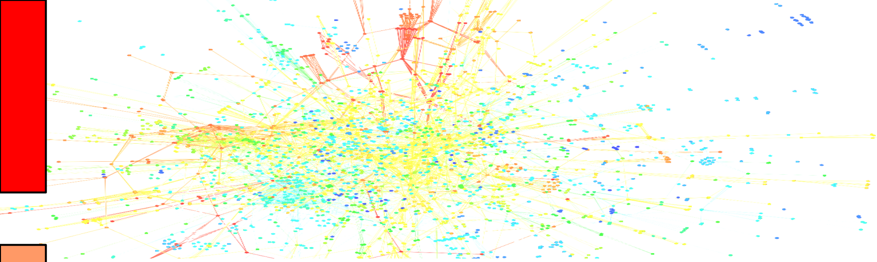
Part 3. Micro-Array Accuracy Analysis

Werner Van Belle

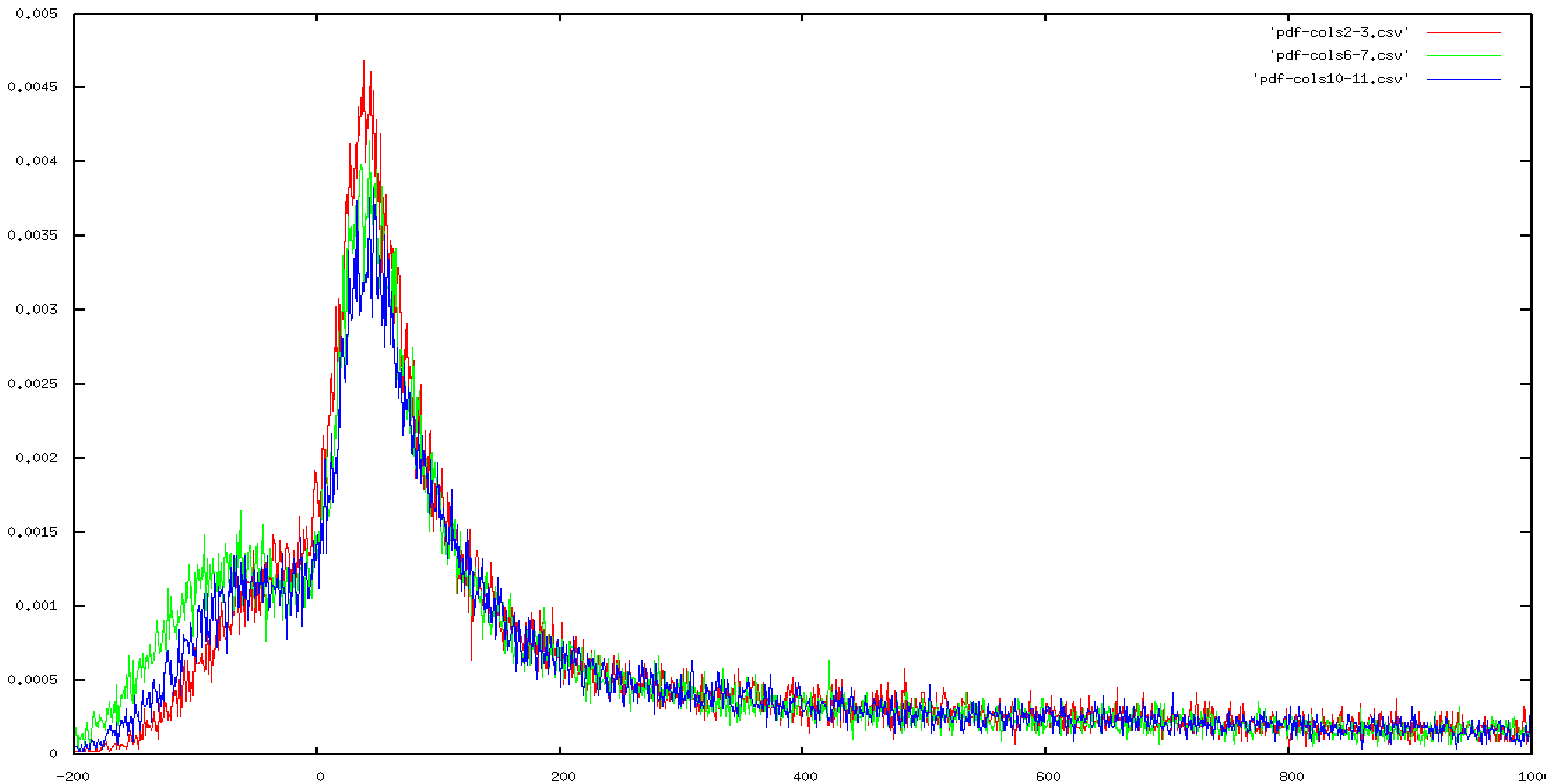
werner.van.belle @ gmail.com, werner @ onlinux.be

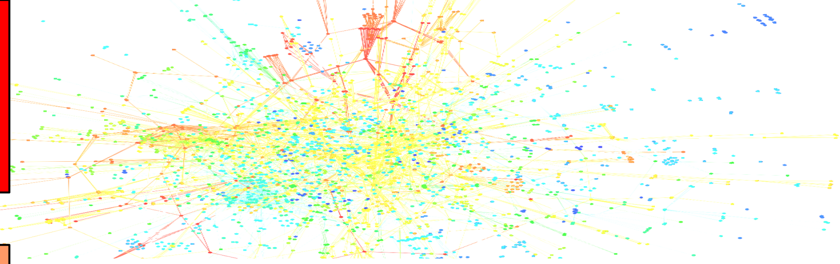
In cooperation with: Nancy Gerits, Ugo Moens,

Halvor Grønaas, Lotte Olsen, Ruth Paulssen

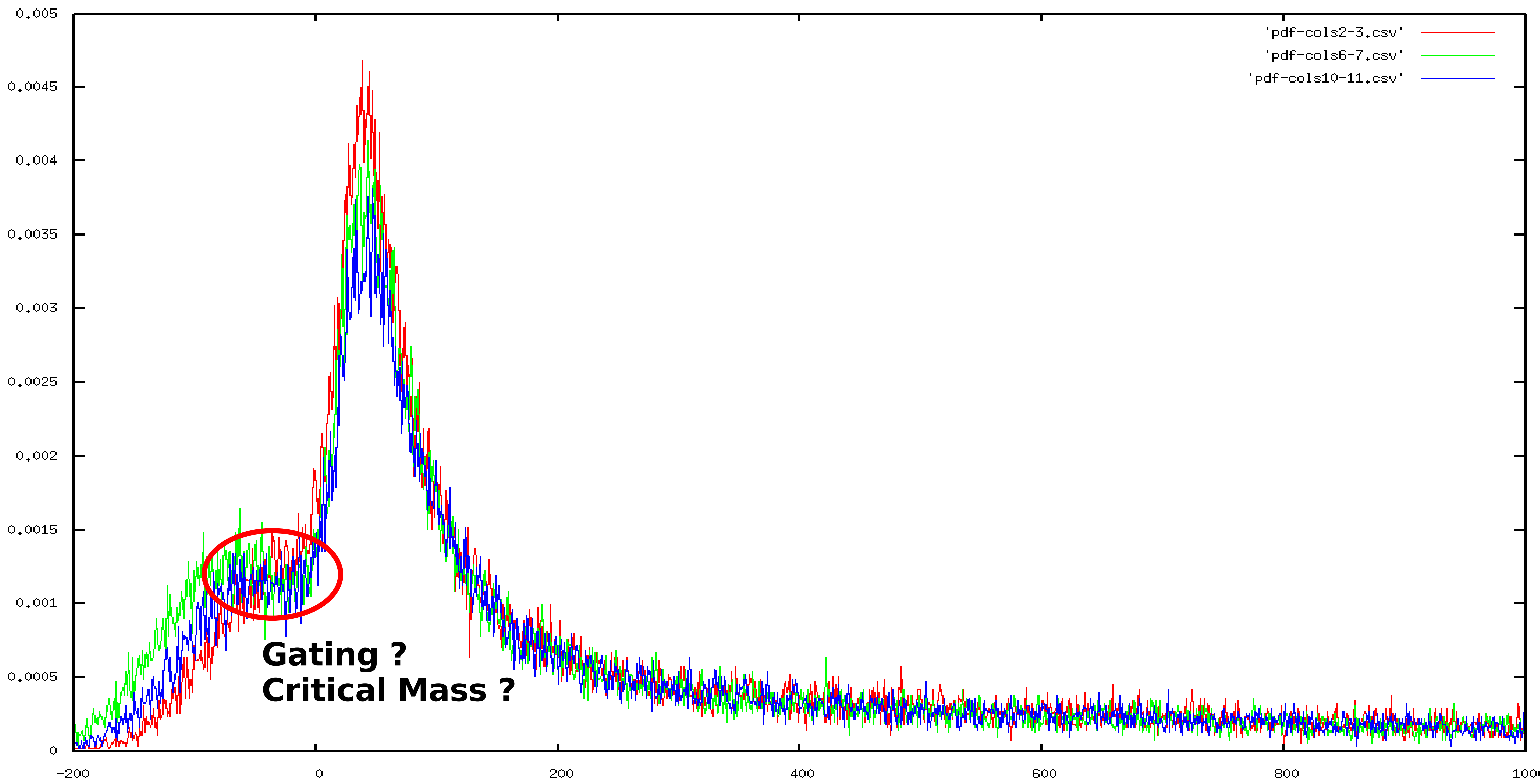


Intensity Distribution

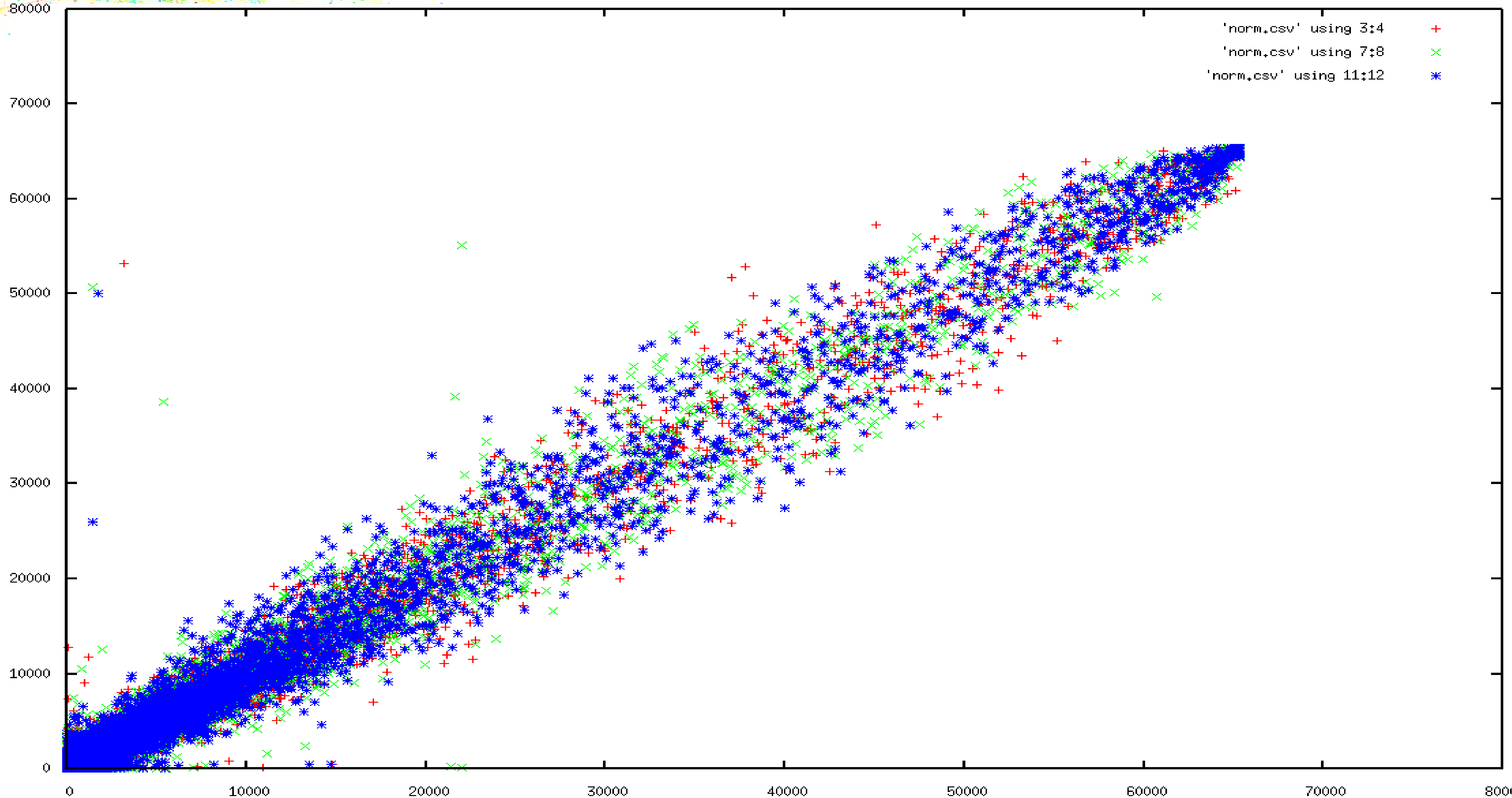




Intensity Distribution

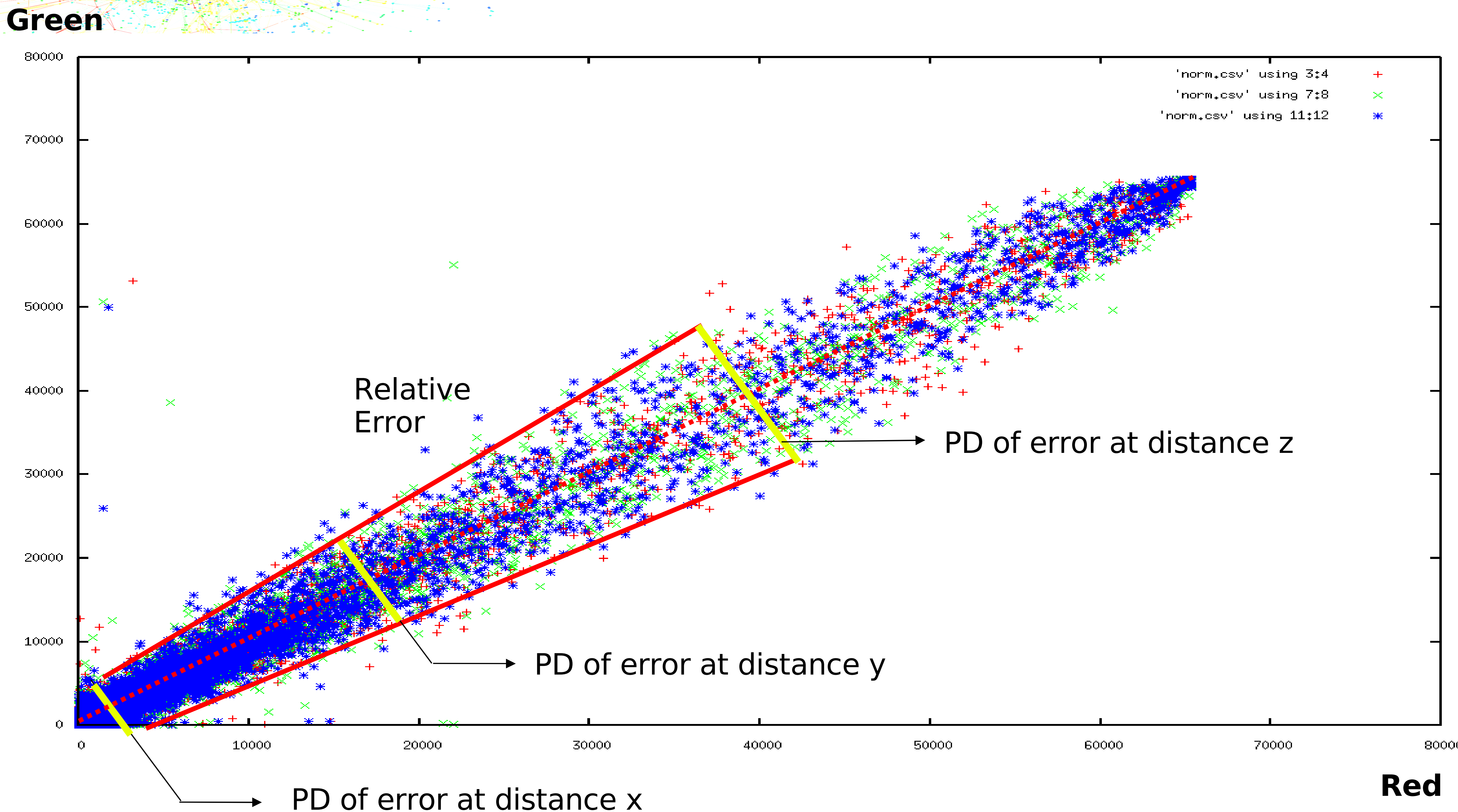


Cy5 ?= Cy3

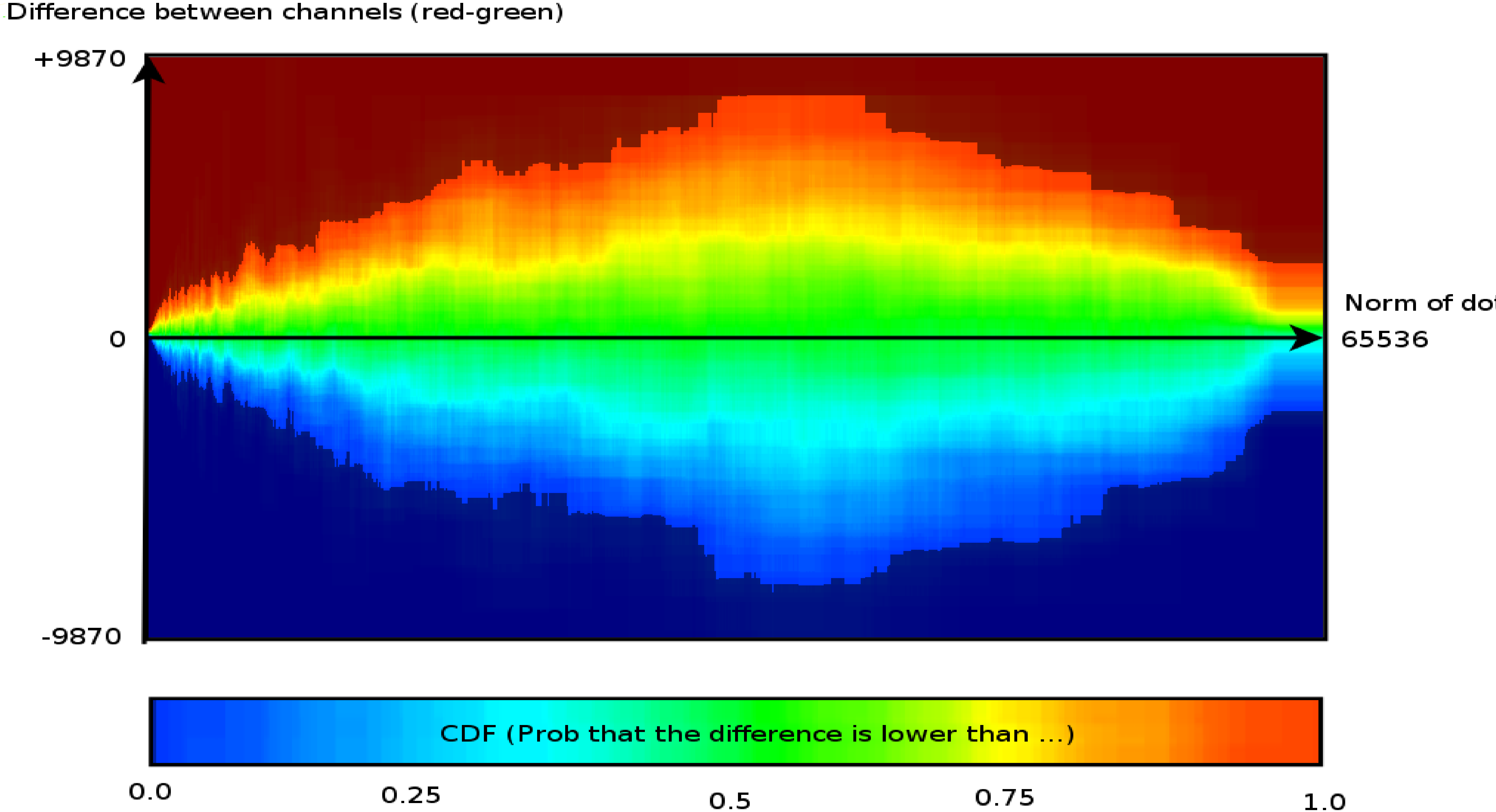


'norm.csv' using 3:4
'norm.csv' using 7:8
'norm.csv' using 11:12
0
10000
20000
30000
40000
50000
60000
70000
80000

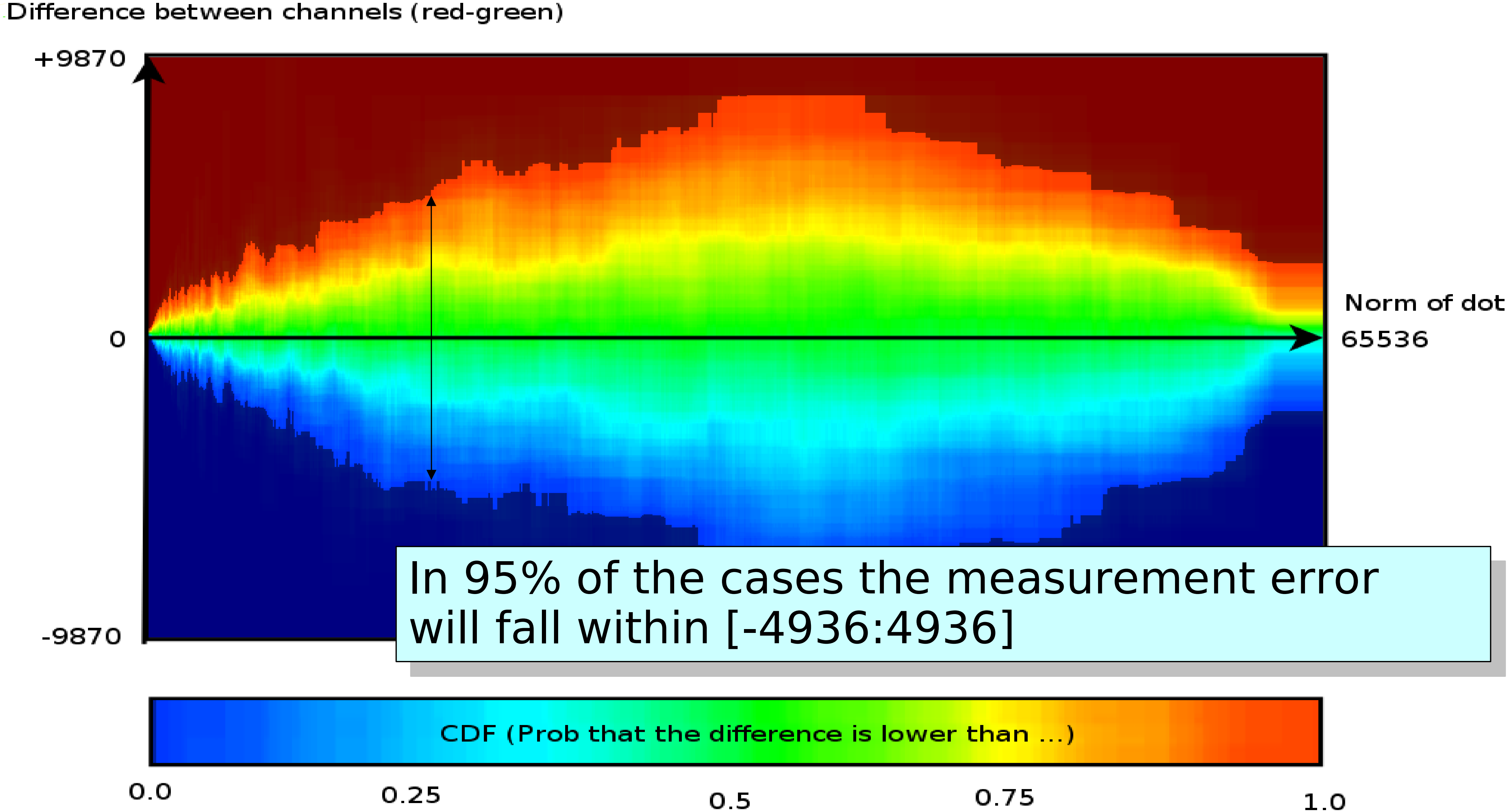
Measurement Accuracy



Intensity Dependent Error Distribution

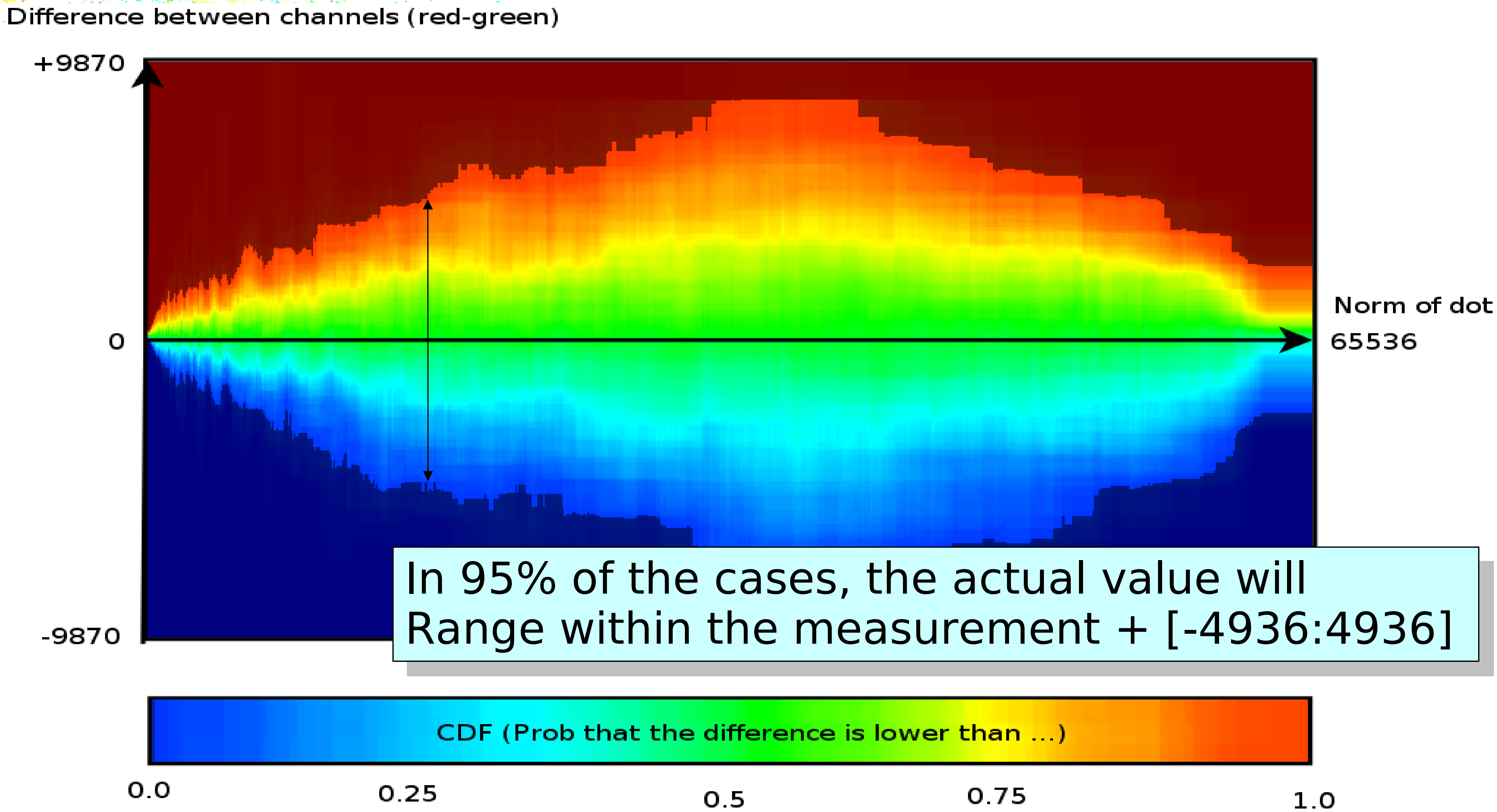


Intensity Dependent Error Distribution



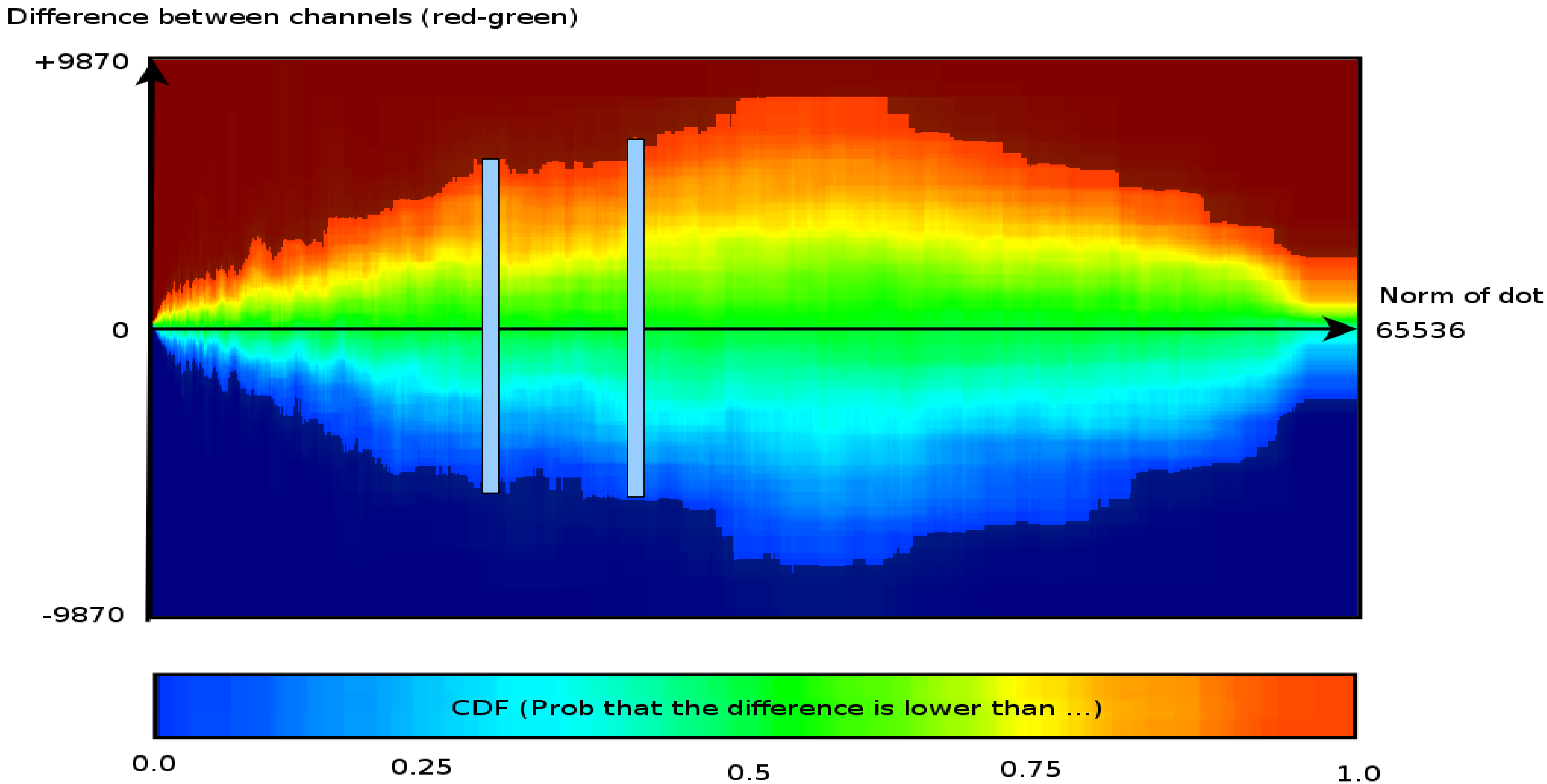
In 95% of the cases the measurement error will fall within [-4936:4936]

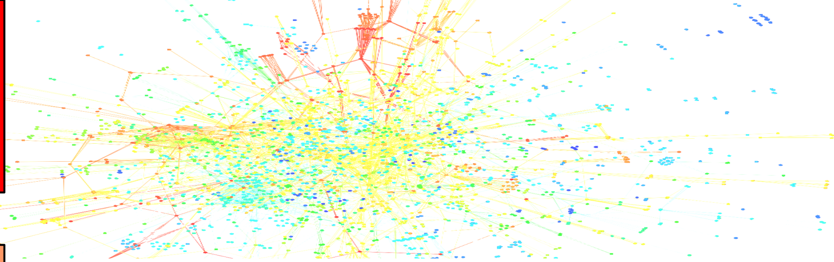
Confidence Interval for 1 Spot



Multiple Spots

- Multiple measurements lead to better estimates / smaller confidence intervals



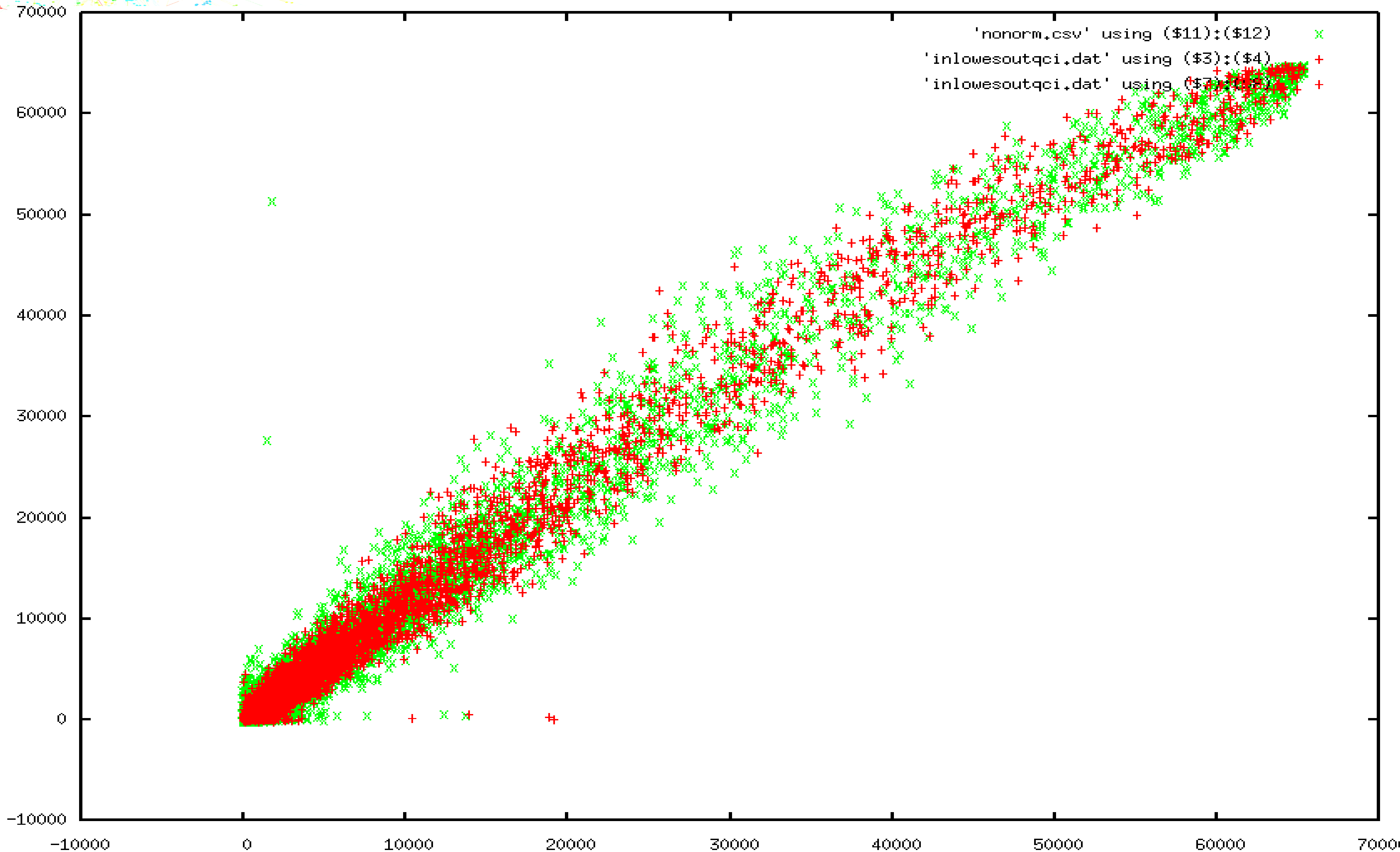


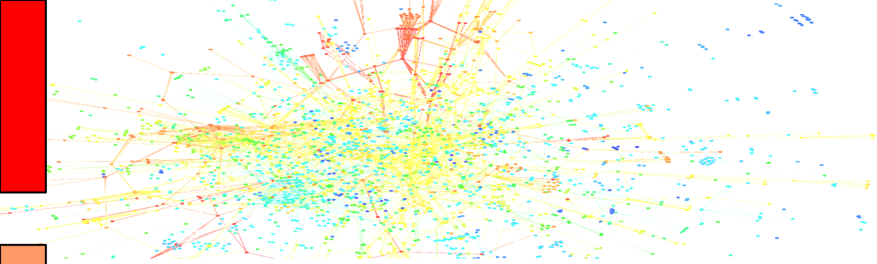
Reported Regulations

		Values			Difference				Factor
ID	**C.I.**	**Red**	**Green**	**#**	**Lo**	**Normal**	**Hi**	**Regu**	**Lowest**
R000008_01	[-1515.52:1812.48]	2502	81	1	-905.48	-2421	-4233.48	down	2.08
R000068_01	[-4894.72:5621.76]	18661	12833	2	-933.28	-5828	-11449.8	down	1.06
R000088_01	[-3194.88:2938.88]	7637	2963	2	-1479.12	-4674	-7612.88	down	1.32
R000137_01	[-307.2:307.2]	39	-294	2	-25.8	-333	-640.2	up	2.32
R000141_01	[-8273.92:7905.28]	24161	34097	2	18209.9	9936	2030.72	up	1.07
R000177_01	[-993.28:972.8]	1154	72	1	-88.72	-1082	-2054.8	down	1.16
R000186_01	[-204.8:215.04]	86	-155	1	-36.2	-241	-456.04	up	1.36
R000248_01	[-204.8:215.04]	115	-125	1	-35.2	-240	-455.04	up	1.04
R000293_01	[-3573.76:3737.6]	10913	7012	2	-327.24	-3901	-7638.6	down	1.04
R000310_01	[-3328:3665.92]	12767	2304	2	-7135	-10463	-14128.9	down	2.8
R000490_01	[-665.6:655.36]	826	149	2	-11.4	-677	-1332.36	down	1.02
R000504_01	[-10506.2:10537]	57663	69019	2	21862.2	11356	819.04	up	1.01
R000668_01	[-307.2:317.44]	340	-157	2	-189.8	-497	-814.44	down	1.58
R000711_01	[-665.6:655.36]	866	-206	2	-406.4	-1072	-1727.36	down	4.21

...

Omitted spots: too close to error



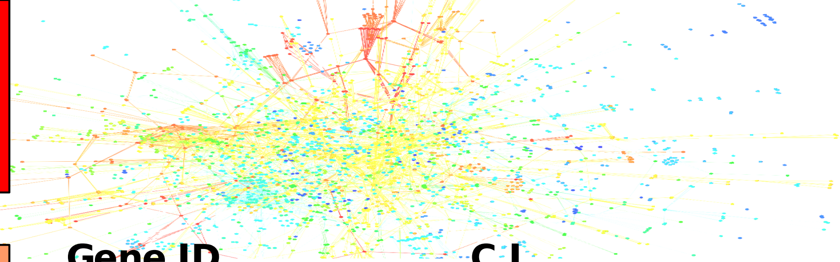


Part 4. Protein Interaction Map Integration

Werner Van Belle

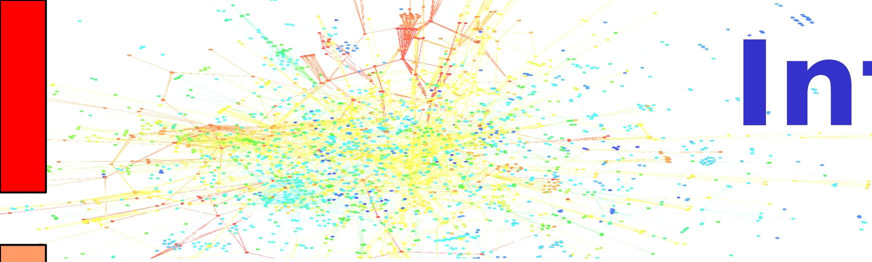
werner.van.belle@gmail.com, werner@onlinux.be

In cooperation with: Nancy Gerits, Ugo Moens



Gene Expression

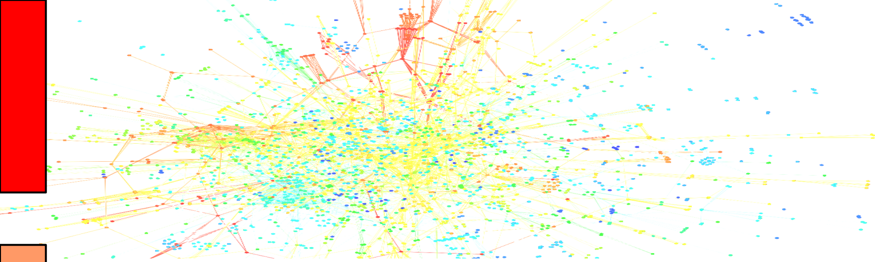
Gene ID	C.I.	Difference Low	Difference Norm	Difference Hi			Regulation Factor	At least	At most	Biological Process
ENSG00000072121	[-5017.6:5212.16]	14288.4	19306	24518.2	6	up	3.86	4.87	inf	-
ENSG00000104894	[-3399.68:4085.76]	-13862.7	-10463	-6377.24	2	down	3.21	4.62	inf	-
ENSG00000156006	[-2058.24:1966.08]	-5562.24	-3504	-1537.92	2	down	2.62	4.69	inf	metabolism
ENSG00000183762	[-2826.24:2682.88]	2817.76	5644	8326.88	2	up	2.45	3.9	inf	cell communication
ENSG00000132334	[-2416.64:2529.28]	1905.36	4322	6851.28	2	up	2.34	4.04	inf	macromolecule metabolism
ENSG00000140988	[-6891.52:7372.8]	-40614.5	-33723	-26350.2	4	down	2.14	2.45	3.49	protein metabolism
ENSG00000137198	[-2109.44:2426.88]	625.56	2735	5161.88	2	up	1.78	4.41	inf	metabolism
ENSG00000176105	[-6348.8:7229.44]	10030.2	16379	23608.4	2	up	1.74	2.21	4.73	cell communication
ENSG00000100219	[-2621.44:2949.12]	-6333.44	-3712	-762.88	2	down	1.7	4.41	inf	transcription, DNA-dependent
ENSG00000177272	[-2058.24:1914.88]	447.76	2506	4420.88	2	up	1.66	4.68	inf	monovalent inorganic cation transport
ENSG00000169992	[-3461.12:3645.44]	2536.88	5998	9643.44	10	up	1.65	2.55	41.77	cell communication
ENSG00000152684	[-2426.88:2887.68]	-6224.88	-3798	-910.32	2	down	1.65	3.72	inf	signal transduction
ENSG00000175279	[-2775.04:2754.56]	1531.96	4307	7061.56	6	up	1.61	2.7	inf	-
ENSG00000171428	[-2437.12:2447.36]	-5919.12	-3482	-1034.64	4	down	1.6	3.01	inf	metabolism



Influenced by/Influences

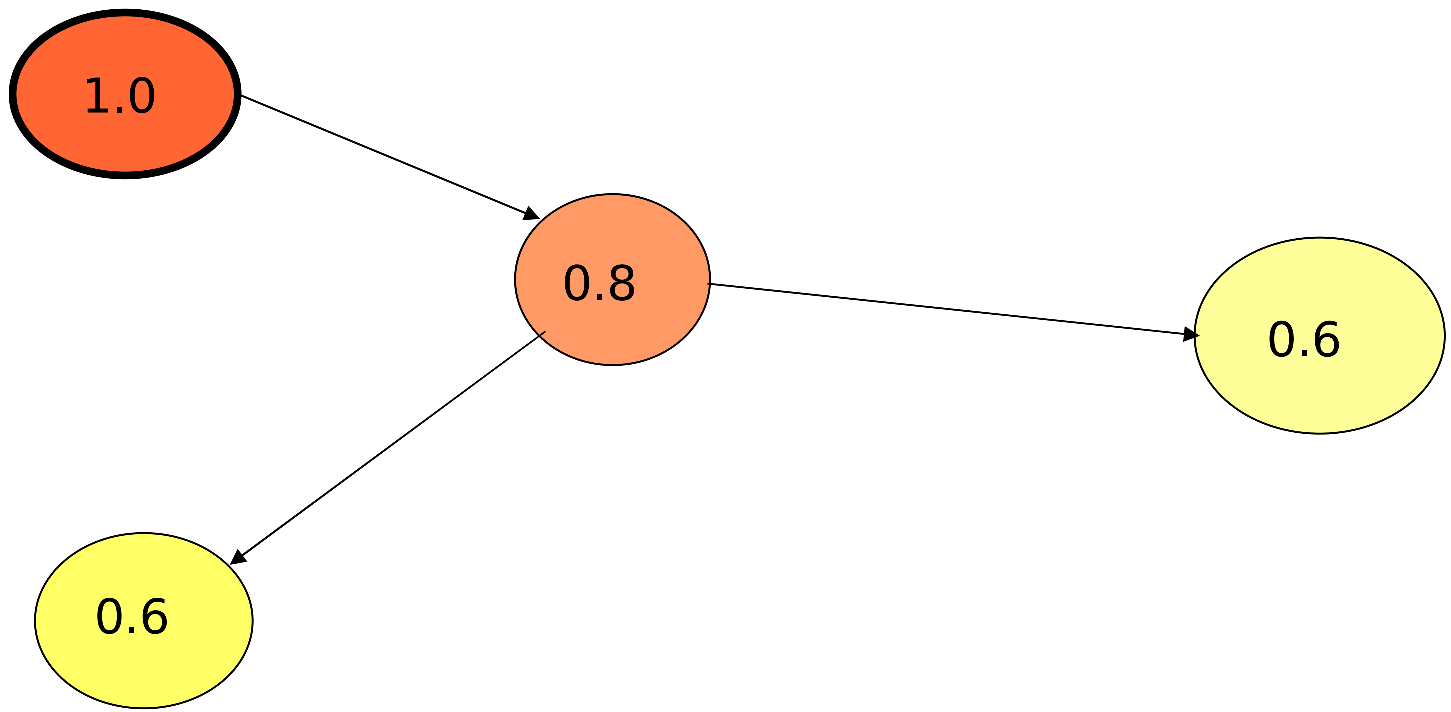
- MK5 -> Multiple changes in gene expression
- 27000 gene expressions measured
- Those that change will very likely influence other proteins

Which proteins are likely influenced by our measured up/down regulations ?



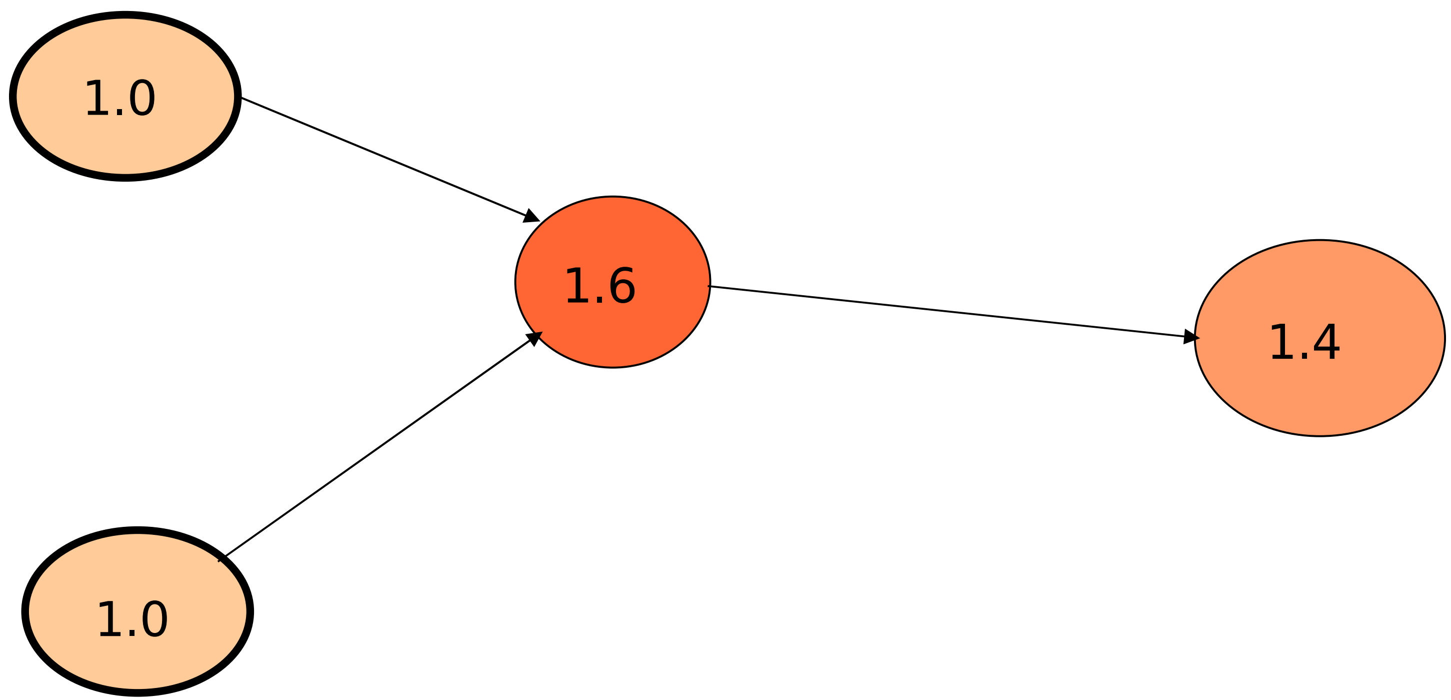
The 'Involved' Game

- Protein change will influence nearby proteins, which in turn ...



The 'Involved' Game

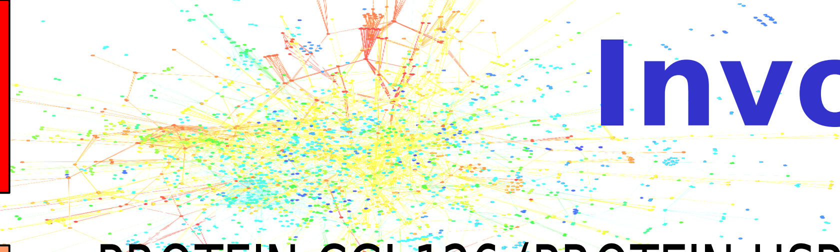
- Multiple proteins changes will all influence their neighbors as well.



The 'Involved' Game

- This network is iterated a number of times to expand the sphere of influence of all the altered gene expressions.
  - affected proteins will have higher numbers
- Protein Interaction key mechanism for signal transduction
  - Protein Interaction Network as published by

> Jean François Rual *et al.* Towards a Proteome Scale Map of the Human Protein Protein Interaction Network – Nature 2005 – vol 437, p. 1173-1178



Involved Proteins by Rank

PROTEIN CGI-126 (PROTEIN HSPC155)

RAD50-INTERACTING PROTEIN 1. [Source:RefSeq;Acc:NM_021930]

RHO-RELATED BTB DOMAIN-CONTAINING PROTEIN 2 (DELETED IN BREAST CANCER 2 GENE PROTEIN) (P83).

NADH-UBIQUINONE OXIDOREDUCTASE 18 KDA SUBUNIT, MITOCHONDRIAL PRECURSOR (EC 1.6.5.3) (EC 1.6.99.3) (COMPLEX I-18 KDA) (CI-18 KDA) (COMPLEX I- AQDQ) (CI-AQDQ).

CHROMATIN ACCESSIBILITY COMPLEX PROTEIN 1 (CHRAC-1) (CHRAC-15) (HUCHRAC15) (DNA POLYMERASE EPSILON SUBUNIT P15).

ADIPONECTIN RECEPTOR 2. [Source:RefSeq;Acc:NM_024551]

ODD-SKIPPED RELATED 1; ODZ (ODD OZ/TEN-M) RELATED 1.

DNA POLYMERASE EPSILON P12 SUBUNIT (DNA POLYMERASE EPSILON SUBUNIT 4)

PROTEIN X 0004. [Source:RefSeq;Acc:NM_016301]

XPA BINDING PROTEIN 1; MBD2 INTERACTOR PROTEIN; PUTATIVE ATP(GTP)-BINDING PROTEIN

HBS1-LIKE. [Source:RefSeq;Acc:NM_006620]

HOMEOBOX PROTEIN HLX1 (HOMEOBOX PROTEIN HB24).

Involved Proteins by Rank

PROTEIN CGI-126 (PROTEIN HSPC155)

RAD50-INTERACTING PROTEIN 1. [Source:RefSeq;Acc:NM_021930]

RHO-RELATED BTB DOMAIN-CONTAINING PROTEIN 2 (DELETED IN BREAST CANCER 2 GENE PROTEIN) (P83).

NADH-UBIQUINONE OXIDOREDUCTASE 18 KDA SUBUNIT, MITOCHONDRIAL PRECURSOR (EC 1.6.5.3) (EC 1.6.99.3) (COMPLEX I-18 KDA) (CI-18 KDA) (COMPLEX I- AQDQ) (CI-AQDQ).

CHROMATIN ACCESSIBILITY COMPLEX PROTEIN 1 (CHRAC-1) (CHRAC-15) (HUCHRAC15) (DNA POLYMERASE EPSILON SUBUNIT P15).

ADIPONECTIN RECEPTOR 2. [Source:RefSeq;Acc:NM_024551]

ODD-SKIPPED RELATED 1; ODZ (ODD OZ/TEN-M) RELATED 1.

DNA POLYMERASE EPSILON P12 SUBUNIT (DNA POLYMERASE EPSILON SUBUNIT 4)

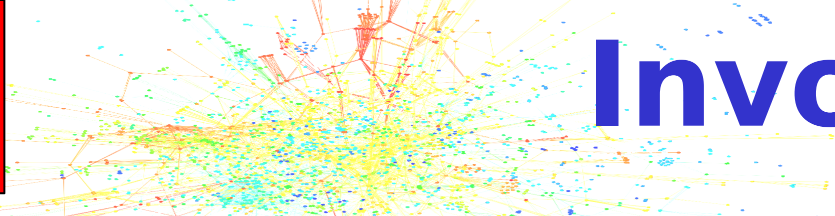
PROTEIN X 0004. [Source:RefSeq;Acc:NM_016301]

XPA BINDING PROTEIN 1; MBD2 INTERACTOR PROTEIN; PUTATIVE ATP(GTP)-BINDING PROTEIN

HBS1-LIKE. [Source:RefSeq;Acc:NM_006620]

HOMEOBOX PROTEIN HLX1 (HOMEOBOX PROTEIN HB24).

NUCLEAR TRANSCRIPTION FACTOR Y SUBUNIT BETA (NF-Y PROTEIN CHAIN B) (NF-YB) (CCAAT-BINDING TRANSCRIPTION FACTOR SUBUNIT A) (CBF-A) (CAAT- BOX DNA BINDING PROTEIN SUBUNIT B).



Involved Proteins by Rank

PROTEIN CGI-126 (PROTEIN HSPC155)

RAD50-INTERACTING PROTEIN 1. [Source:RefSeq;Acc:NM_021930]

RHO-RELATED BTB DOMAIN-CONTAINING PROTEIN 2 (DELETED IN BREAST CANCER 2 GENE PROTEIN) (P83).

NADH-UBIQUINONE OXIDOREDUCTASE 18 KDA SUBUNIT, MITOCHONDRIAL PRECURSOR (EC 1.6.5.3) (EC 1.6.99.3) (COMPLEX I-18 KDA) (CI-18 KDA) (COMPLEX I- AQDQ) (CI-AQDQ).

CHROMATIN ACCESSIBILITY COMPLEX PROTEIN 1 (CHRAC-1) (CHRAC-15) (HUCHRAC15) (DNA POLYMERASE EPSILON SUBUNIT P15).

ADIPONECTIN RECEPTOR 2. [Source:RefSeq;Acc:NM_024551]

ODD-SKIPPED RELATED 1; ODZ (ODD OZ/TEN-M) RELATED 1.

DNA POLYMERASE EPSILON P12 SUBUNIT (DNA POLYMERASE EPSILON SUBUNIT 4)

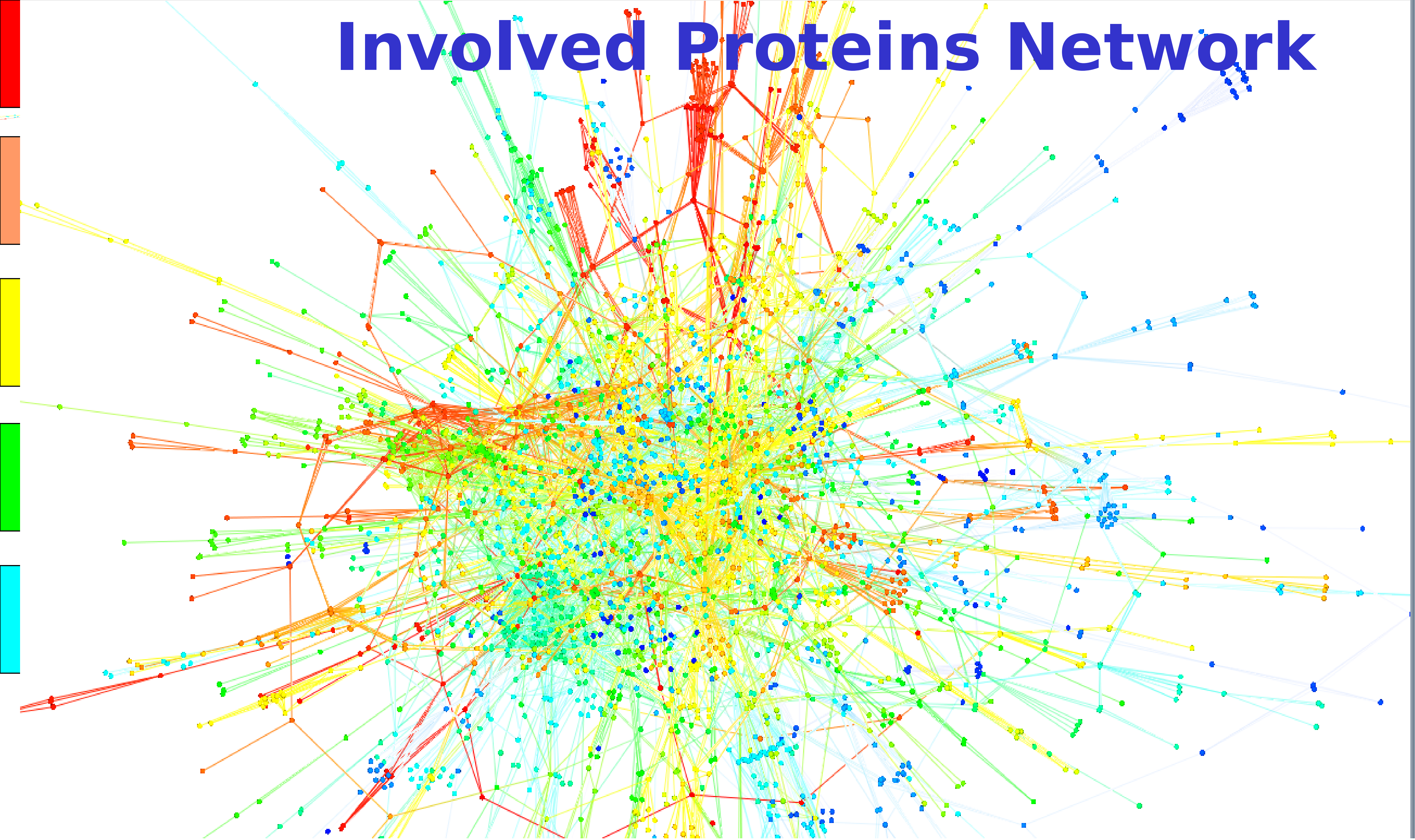
PROTEIN X 0004. [Source:RefSeq;Acc:NM_016301]

XPA BINDING PROTEIN 1; MBD2 INTERACTOR PROTEIN; PUTATIVE ATP(GTP)-BINDING PROTEIN

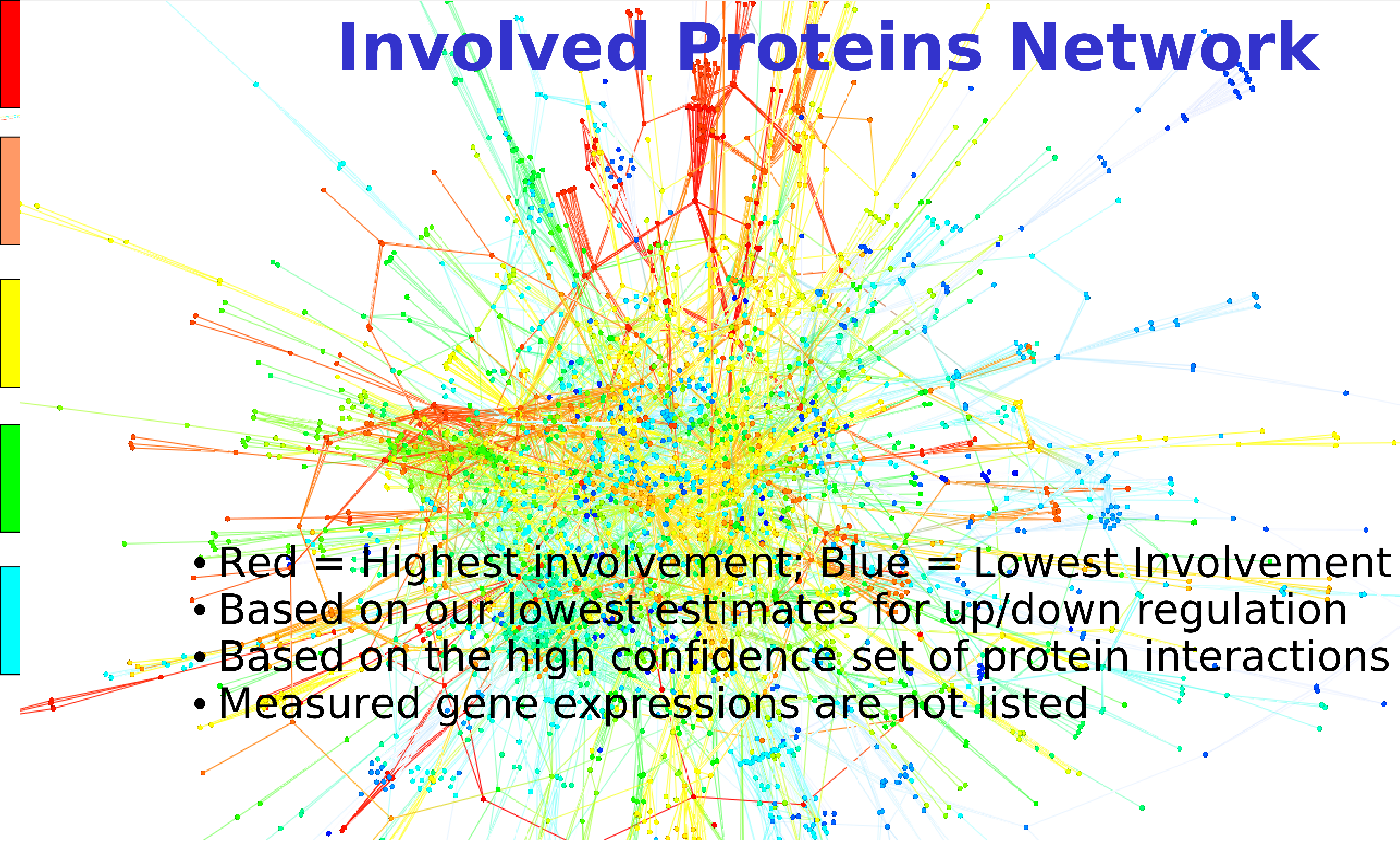
HBS1-LIKE. [Source:RefSeq;Acc:NM_006620]

HOMEOBOX PROTEIN HLX1 (HOMEOBOX PROTEIN HB24).

Involved Proteins Network

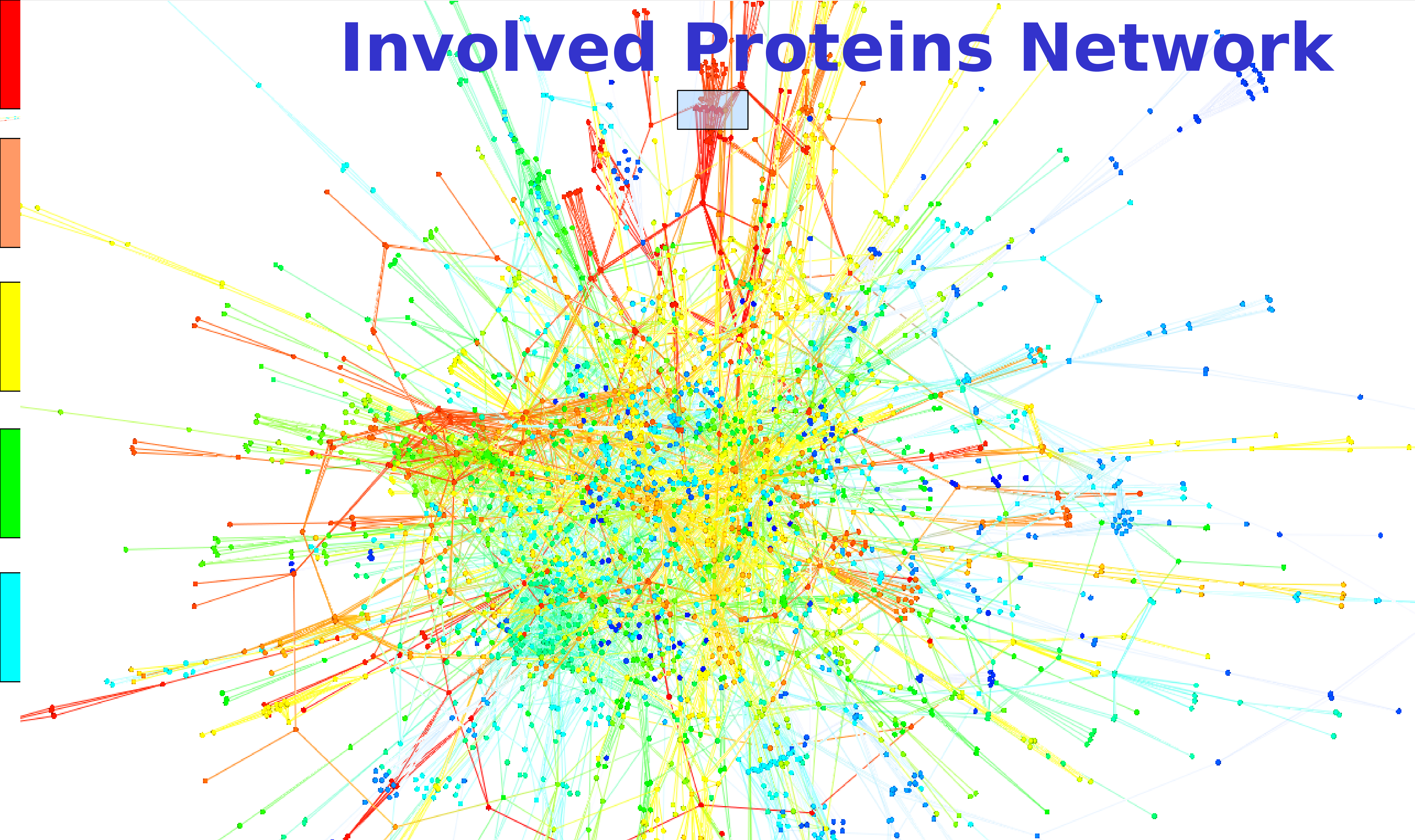


Involved Proteins Network

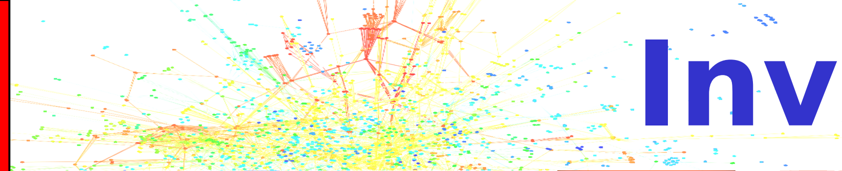


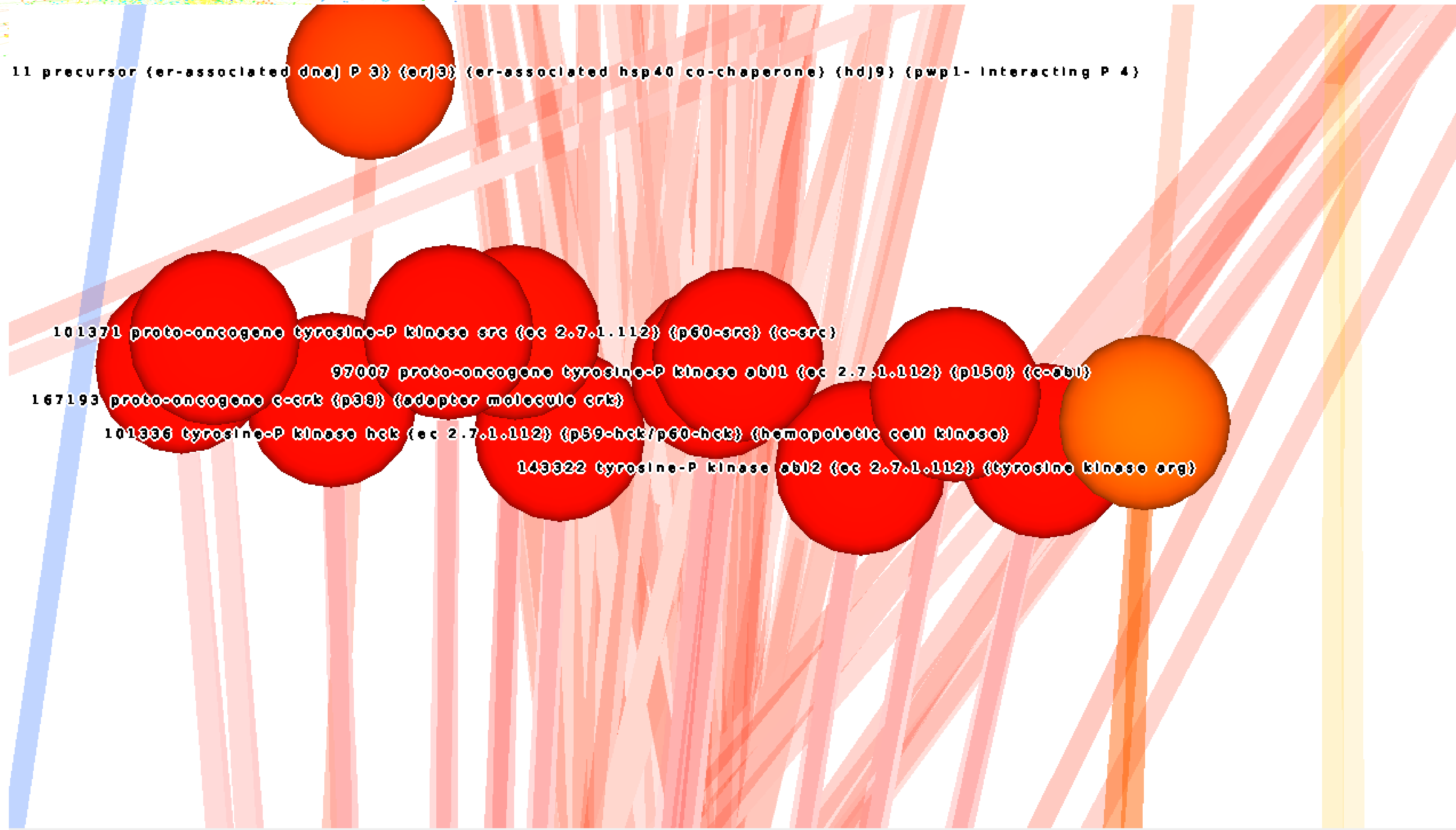
- Red = Highest involvement; Blue = Lowest Involvement
- Based on our lowest estimates for up/down regulation
- Based on the high confidence set of protein interactions
- Measured gene expressions are not listed

Involved Proteins Network

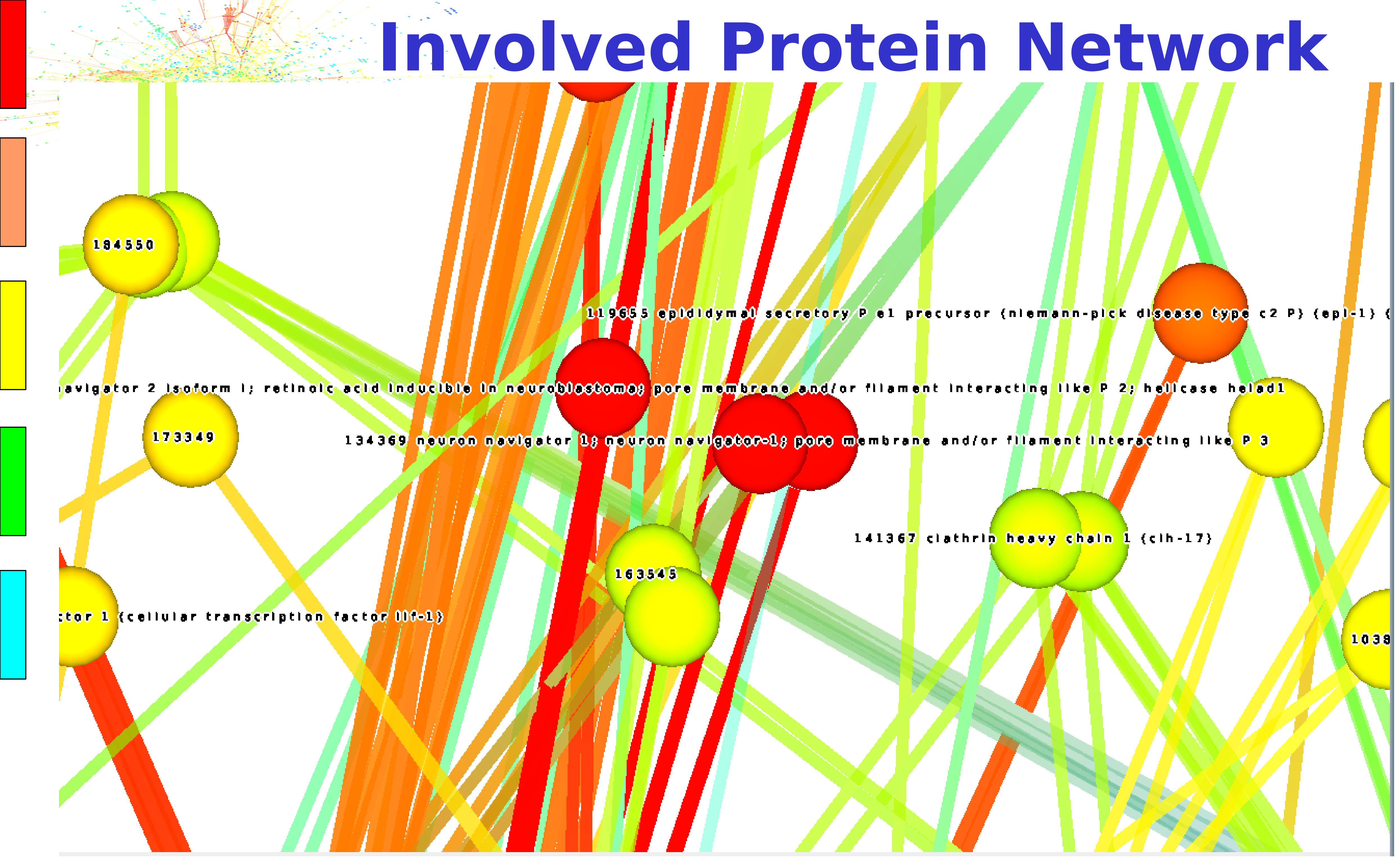


Involved Protein Network

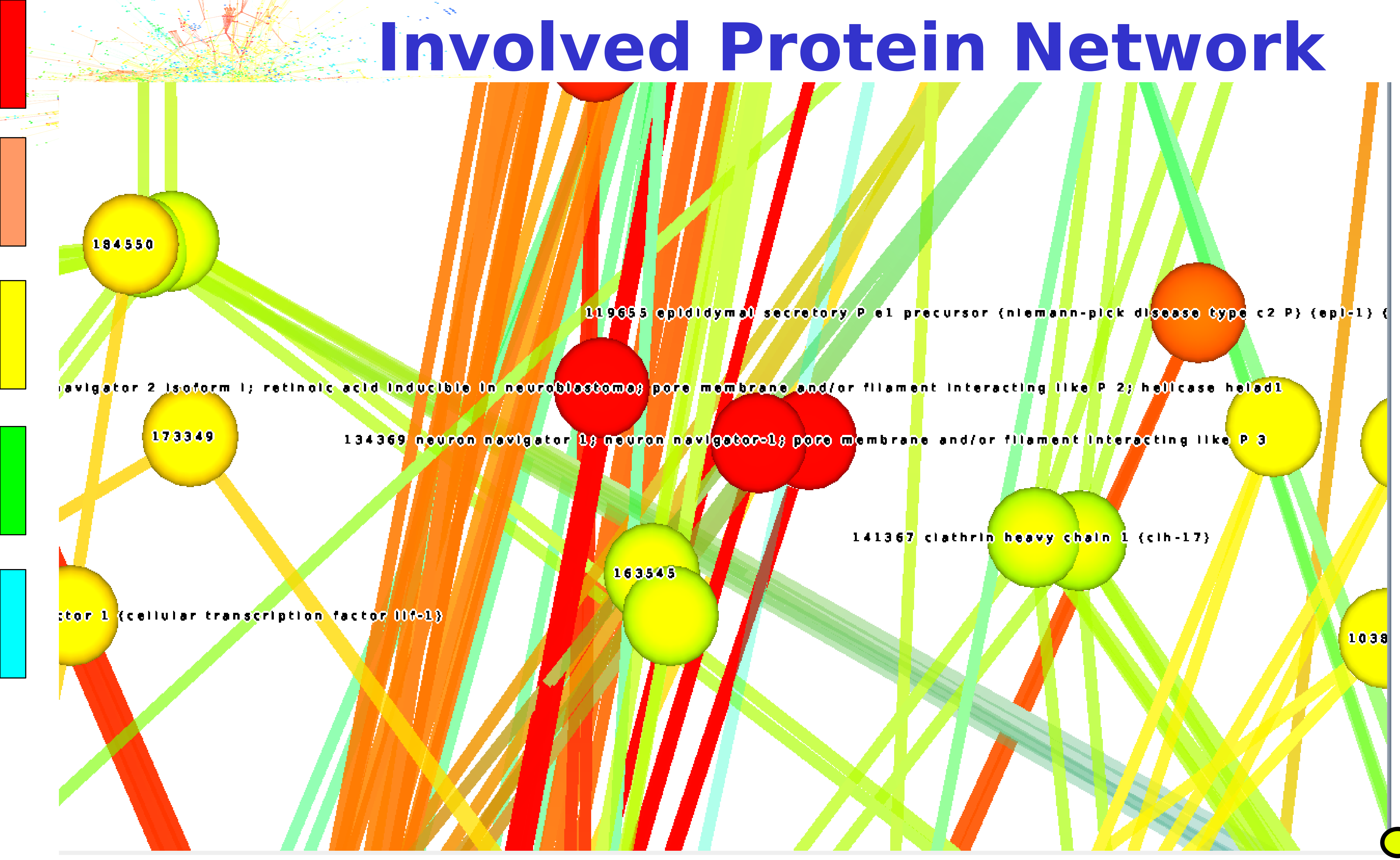


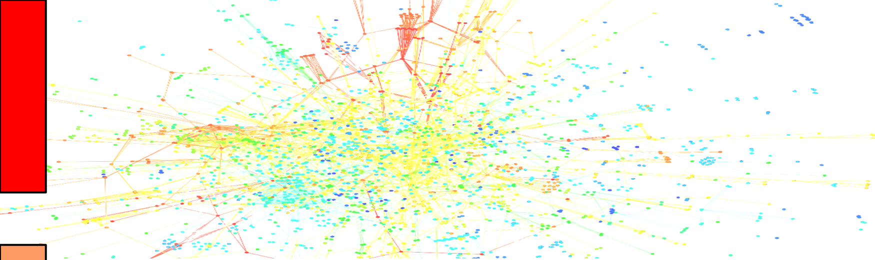


Involved Protein Network



Involved Protein Network





Credits

- 2DE Gel Imaging and Patient sampling
  - Nina Ånensen, Bjørn Tore Gjertsen, Ingvild Haaland, Øystein Bruserud, Gry Sjøholt
- Maldi TOF Mass Spectra
  - Olav Mjaavatten, Kari Espolin Fladmark, Stijn Ove Døskeland
- Micro-arrays
  - Nancy Gerits, Ugo Moens, Halvor Grønaas, Lotte Olsen, Ruth Paulssen