progenesis LCMS

Unique LCMS analysis tool
Identify the real differences in your proteomics data!

Dr Paul N Goulding
Nonlinear Dynamics Ltd, UK

nonlinear
DYNAMICS

Nonlinear Dynamics Ltd.

Who are we?

- Nonlinear Dynamics is dedicated to solving your problems in the fields of Proteomics, Genomics and Biomarker Discovery. This is achieved by applying our core strength in developing innovative, high quality data analysis solutions

What is Progenesis LC-MS?

- Progenesis LC-MS is an advanced proteomics research solution for the label-free quantitative analysis of complex samples

- It uses unique technology to deliver key advantages including speed, protein coverage and statistical power

- It applies simple workflows so you can easily visualise complex analysis data

- It answers biological questions at the protein level.

Presentation plan

- Statistics in protein expression analysis

- Problems with LC-MS protein expression analysis

- How does Progenesis LC-MS solve the probems?

- Benefits of the Progenesis LC-MS solution!

- Brief demonstration of Progenesis LC-MS

What is the aim of protein expression analysis?

- To find the <u>significant</u> expression changes in your data

- What do we mean by significant?

  - we run biological replicate samples,

  - we then apply <u>statistical</u> tests to see if there is a <u>statistically significant</u> difference in protein expression (low p-value)
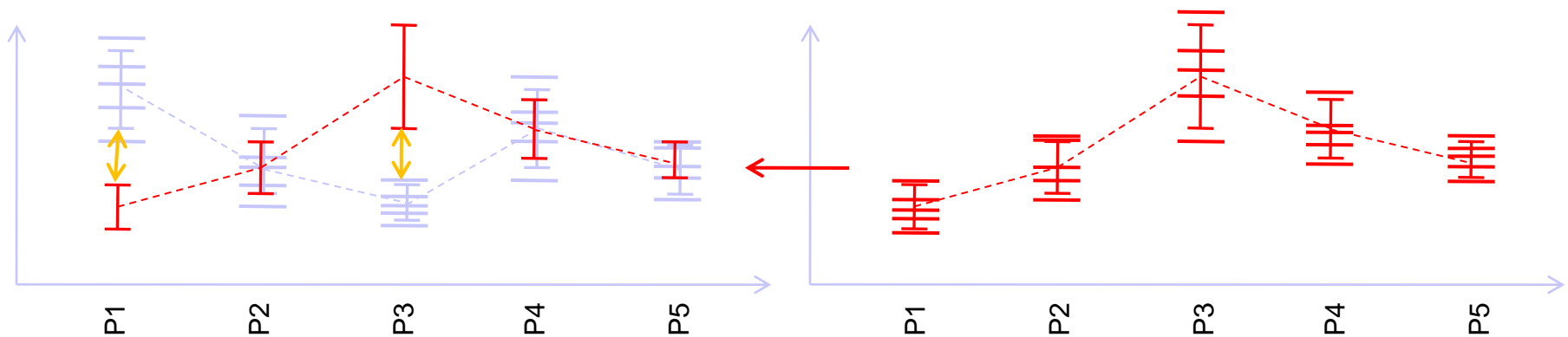
Statistics in protein expression analysis

- Statistical analysis is <u>essential</u> in protein expression analysis!

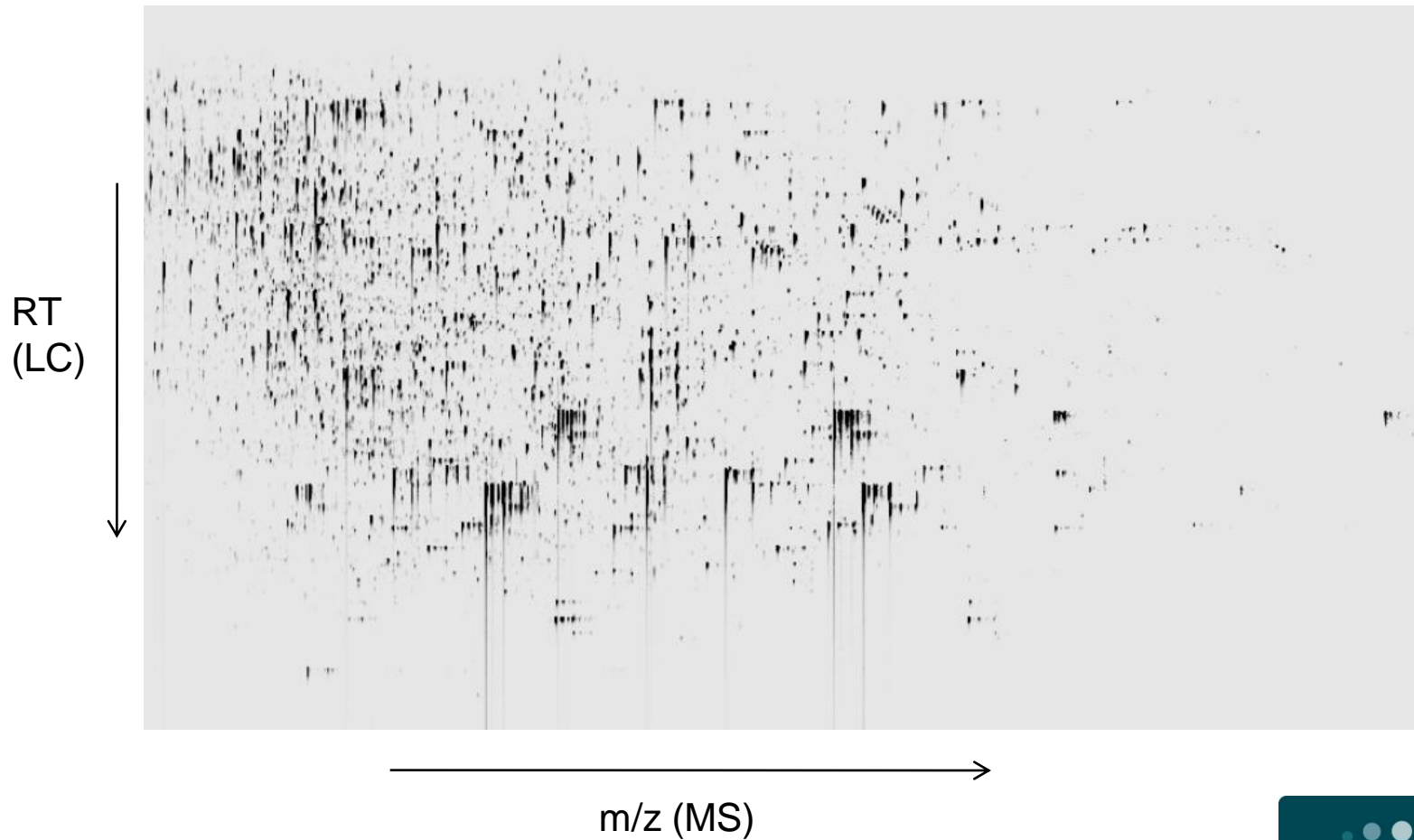- Biological variation in real biological samples means comparing 1 vs 1 to find expression changes is <u>meaningless</u>!
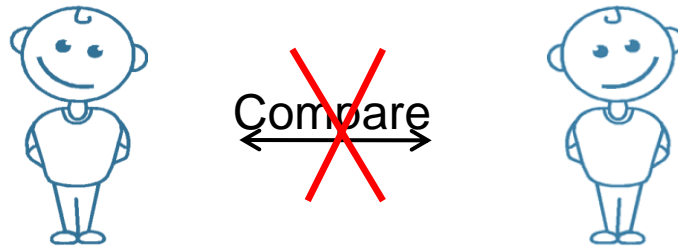
- Comparing insufficient replicates also results in few, or no <u>truly</u> significant results (sufficiently low p-values).

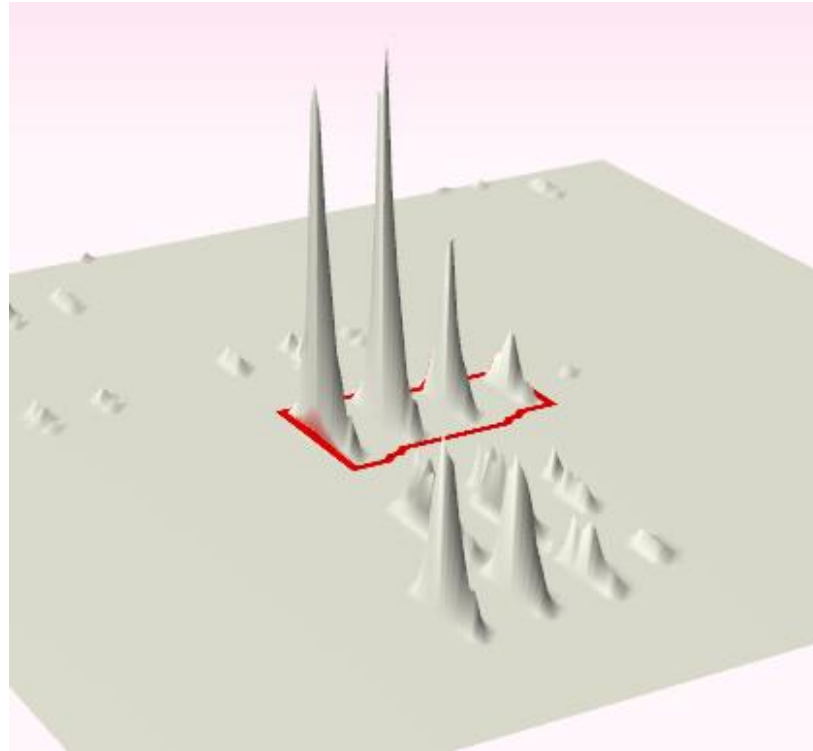# Statistics in protein expression analysis

- Progenesis software uses unique technology to deliver greater **<u>statistical power</u>**

    - ability to find the <u>real</u> expression changes in your data!

- How does it do this?

© Nonlinear Dynamics Group

nonlinear
D Y N A M I C S

LC-MS scans as 2D "feature maps"
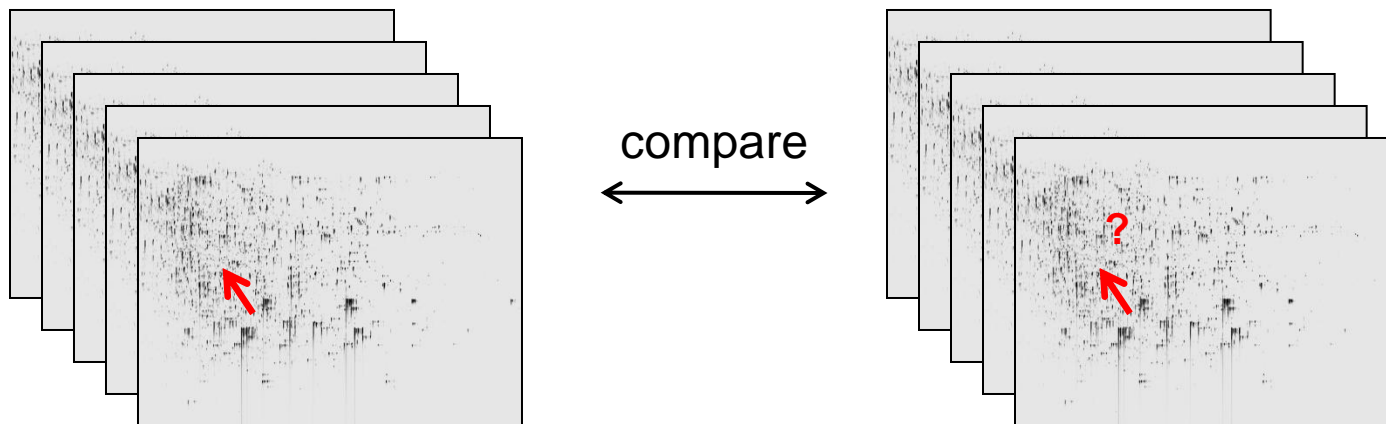

RT
(LC)

m/z (MS)

nonlinear
DYNAMICS

# LC-MS scans as 3D feature maps

Problems with protein expression analysis

- In protein expression analysis, 2D feature maps from <u>biological replicate</u> samples are compared, to find statistically significant differences in the feature quantities.

compare

- The biggest problem with this is "matching" the features.

Problems with LC-MS analysis – unmatched or missing peptides
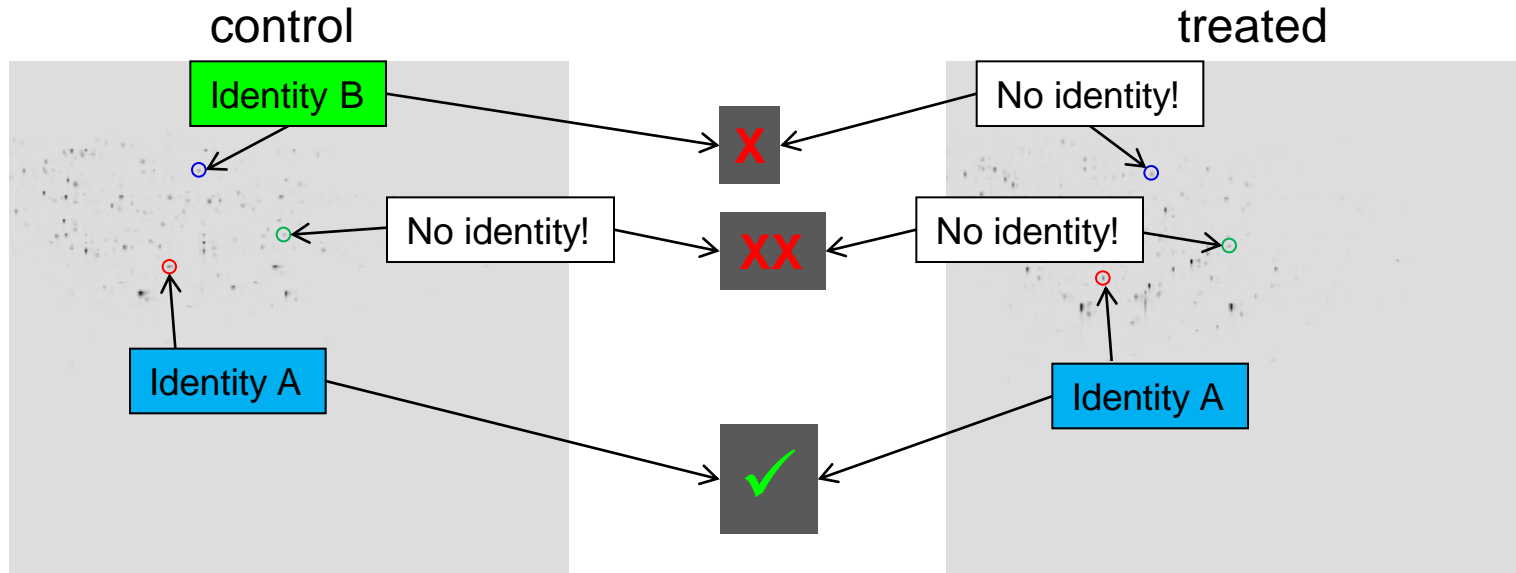
- Most LC-MS analysis methods use peptide IDs from MS/MS spectra for matching.

- You can only <u>quantify</u> what you can <u>identify</u>!

Problems with protein expression analysis – missing values

"Missing Values" in a proteomics data (due to unmatched features)

Control 2a	Control 2b	Control 2c	Treatment 2d	Treatment 2e	Treatment 2f	
Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	t-test (p)
0.018	0.03	0.008	0.026	-	0.008	0.425
0.021	0.026	0.007	0.036	0.007	0.014	0.273
0.012	-	0.003	0.011	0.015	0.016	0.673
0.018	0.026	-	0.035	0.023	-	0.0627
0.049	0.037	0.032	0.009	0.006	0.014	0.0017
0.004	0.002	-	0.12	0.058	0.13	0.032

nonlinear
DYNAMICS

Problems with protein expression analysis – missing values

Control 1	Control 2	Control 3	Control 4	Control 5	Control 6	Treated 1	Treated 2	Treated 3	Treated 4	Treated 5	Treated 6	
Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	Norm. Vol.	t-test (p)
18	12	3	-	8	-	26	6	-	15	-	8	**0.350**
21	-	26	17	7	8	36	-	7	-	14	6	**0.185**
12	20	-	5	3	-	11	22	15	8	16	18	**0.403**
18	-	26	30	-	15	35	-	23	20	-	-	**0.042**

- Increasing the replicate numbers, increases the "missing values"!

nonlinear
DYNAMICS

How Does Progenesis LC-MS eliminate missing values?

- Data <u>alignment</u>, followed by <u>co-detection</u> of features!

- Results in <u>100%</u> feature matching, <u>no missing values</u> and <u>totally reliable</u> statistical analysis!

nonlinear
D Y N A M I C S

Alignment and co-detection in Progenesis LC-MS

reference scan
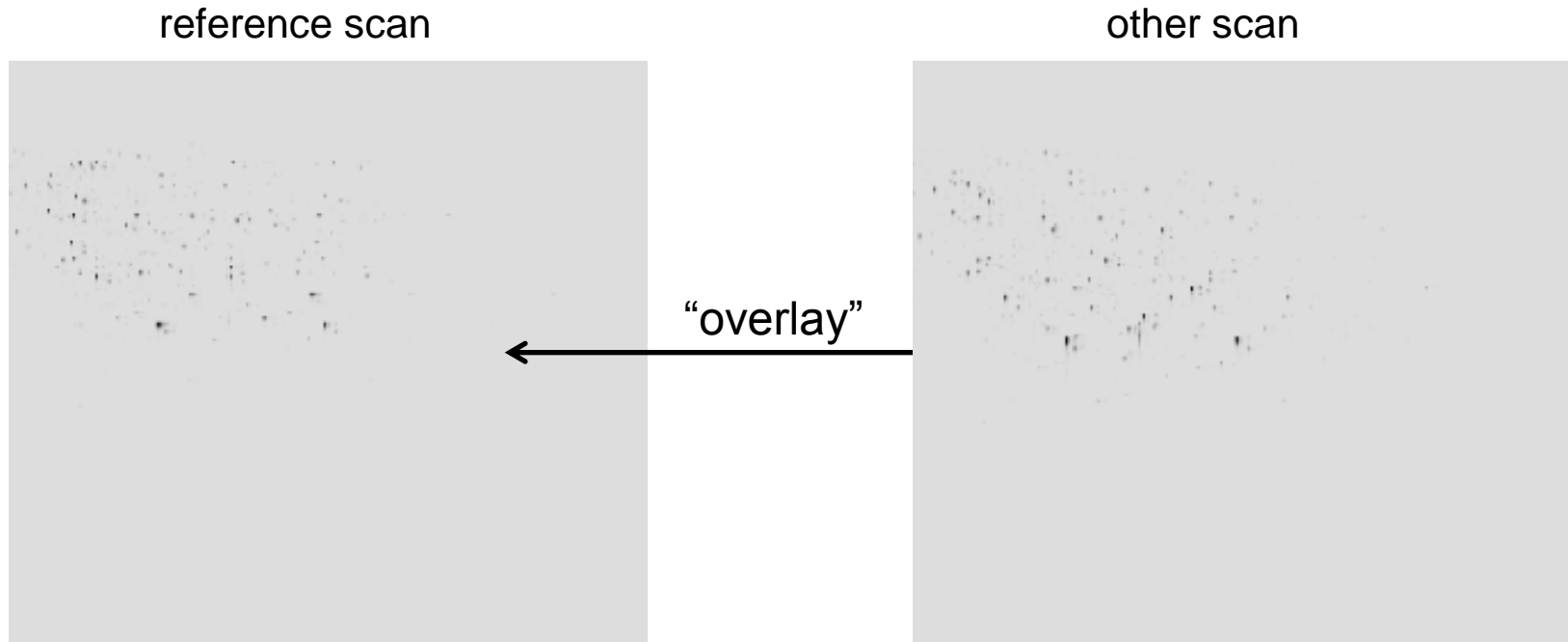
other scan

"overlay"

- One sample is chosen as the alignment "reference".

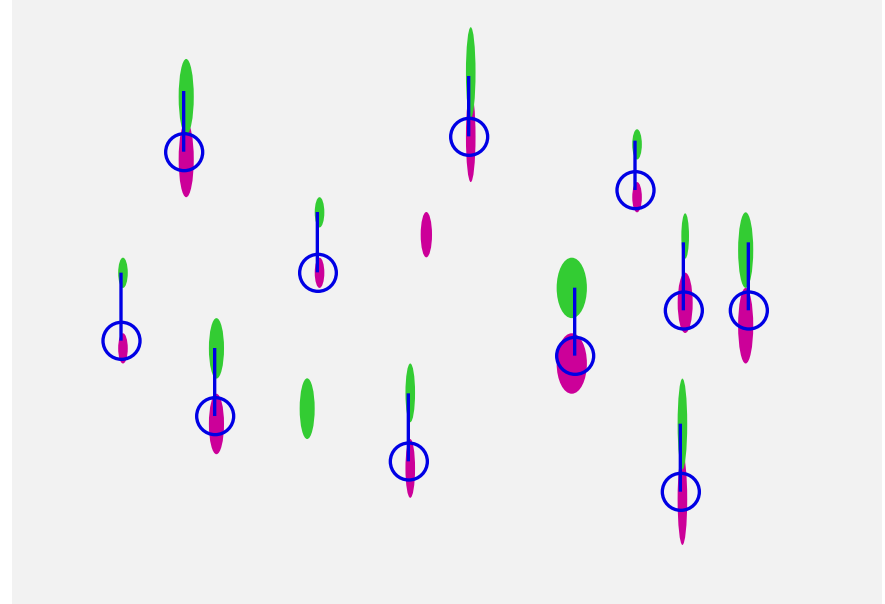- The other samples are aligned against the reference.

- In this way, all are aligned together.

nonlinear
D Y N A M I C S

Alignment and co-detection in Progenesis LC-MS

reference scan

peptide signals

other scan
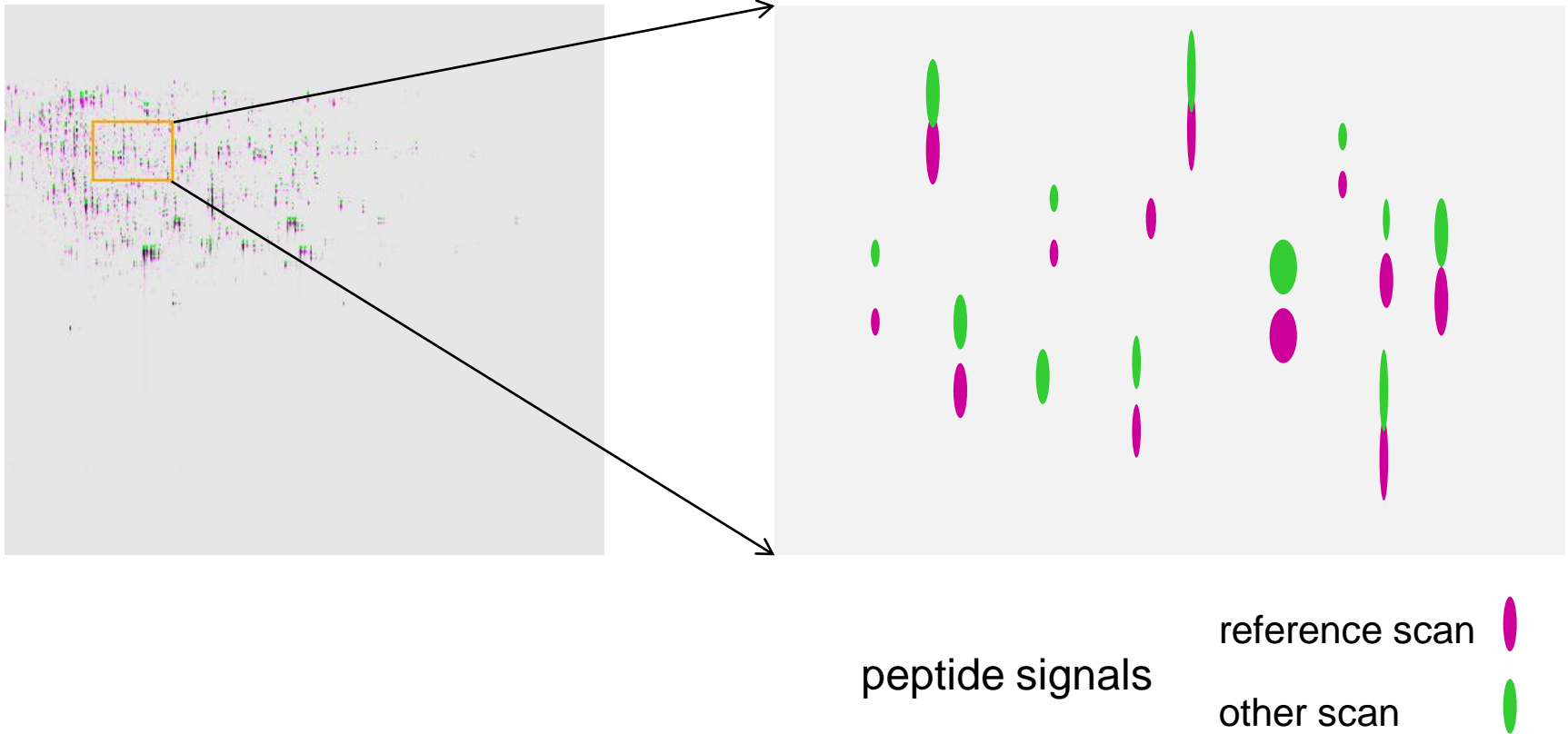
- We can zoom in on the overlaid scans to see the signals in more detail.

Alignment and co-detection in Progenesis LC-MS

Automatic alignment vectors

- A series of "alignment vectors" are automatically created by the software, connecting the alignment points of the "same signal" from each sample.

- The signals are aligned, using the alignment vectors

nonlinear
DYNAMICS

Alignment and co-detection in Progenesis LC-MS
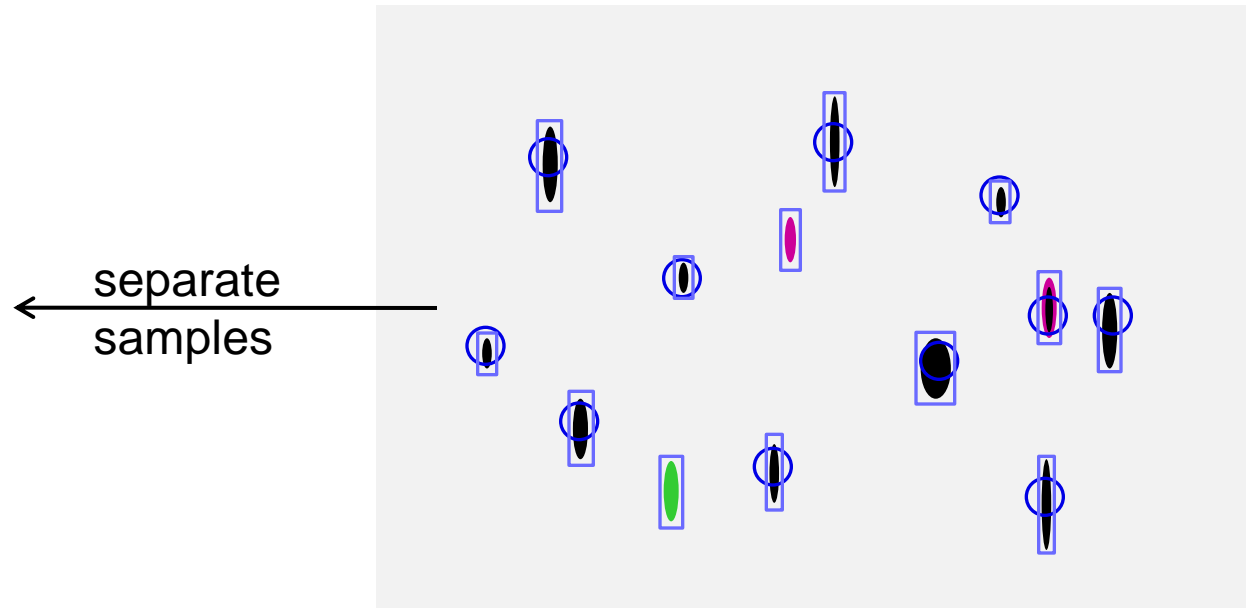
separate
samples

- After alignment, <u>co-detection</u> of the peptides takes place to produce a **single detection pattern** for all the samples!

- Now, the samples can be separated again so you can see where the expression changes are.
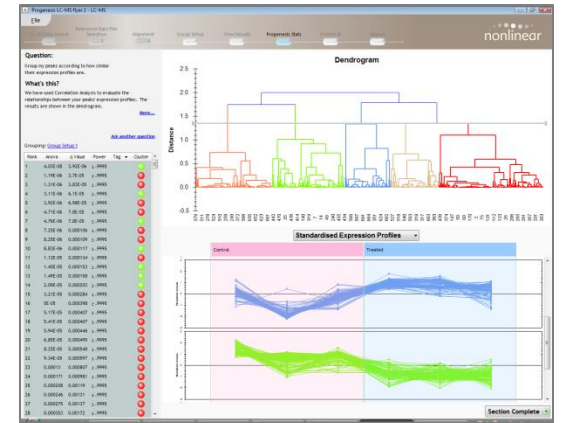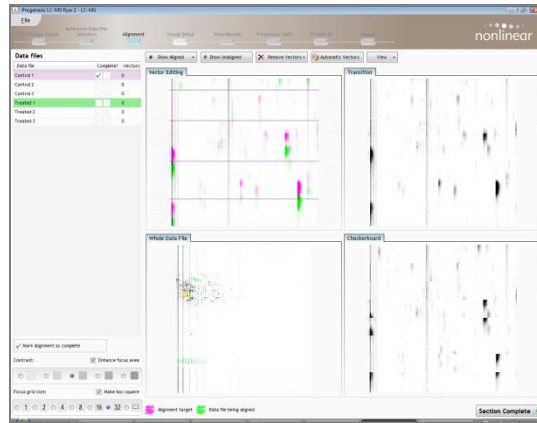
nonlinear
D Y N A M I C S

Alignment and co-detection in Progenesis LC-MS

Same expression change / Large expression change
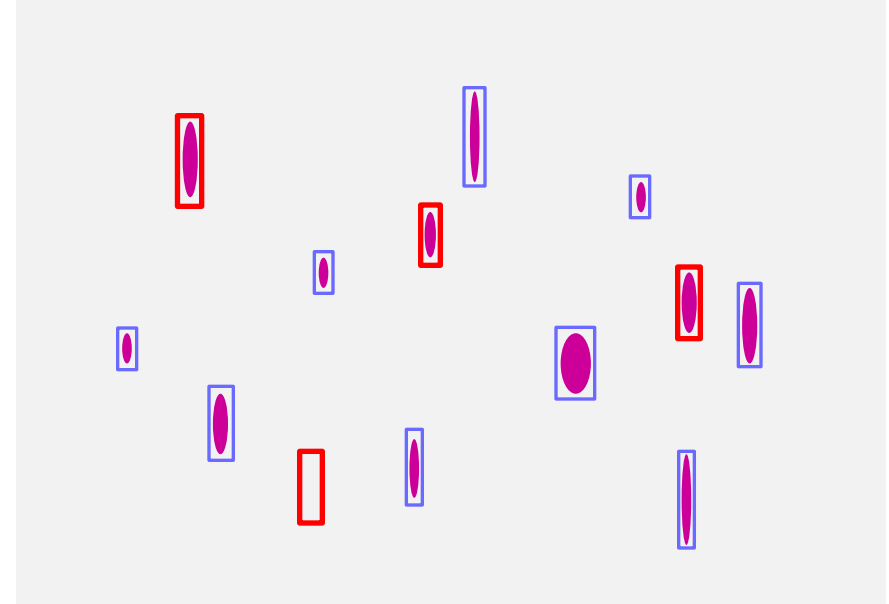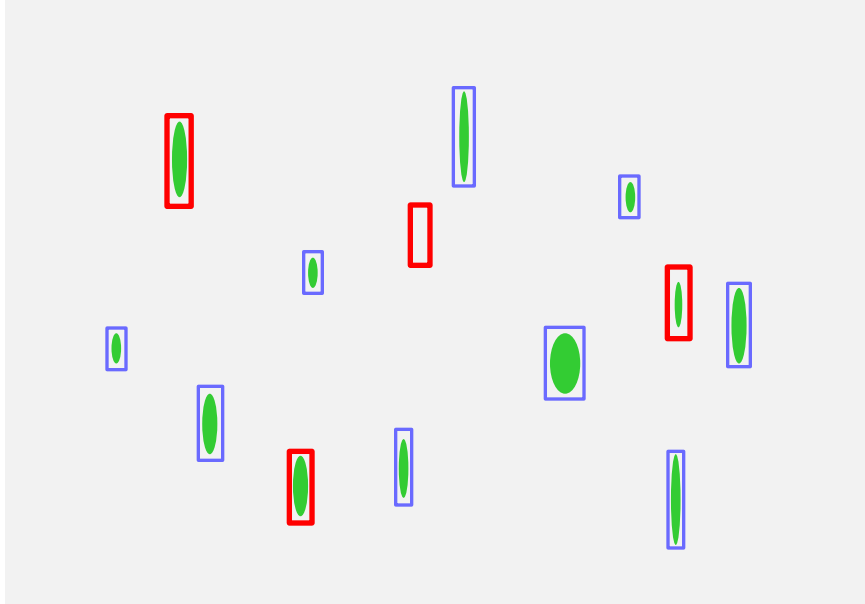
- This same principle can be applied to 2 samples, or to 200 samples!

nonlinear
D Y N A M I C S

A workflow for the differential analysis of complex samples

| Data Import and QC | → | Peak Modelling | → | Retention Time Alignment | → | Data aggregation | → | Detection & Quantification | → | Explore data and visualise results | → |
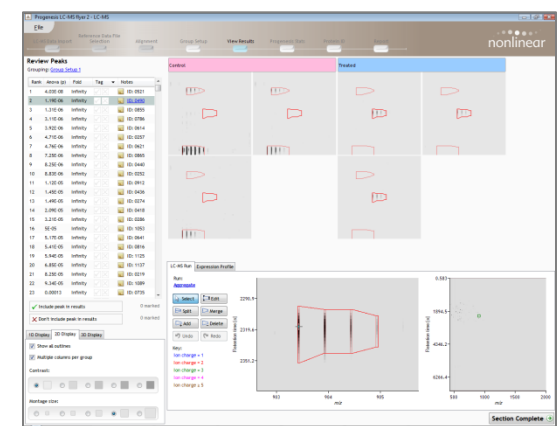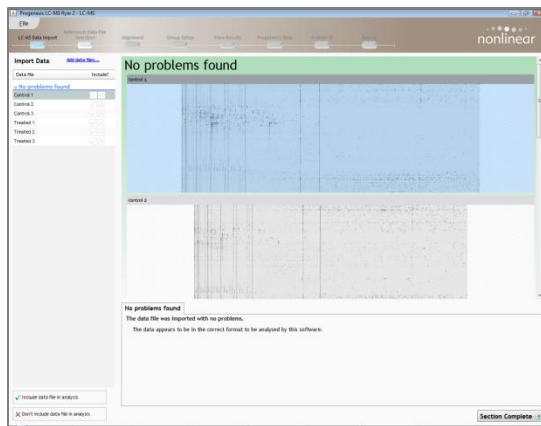
A workflow for the differential analysis of complex samples
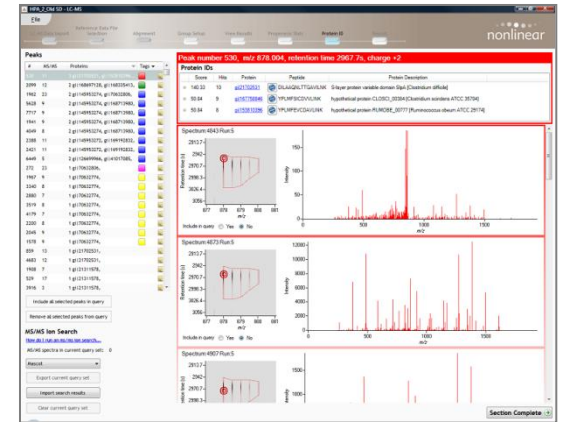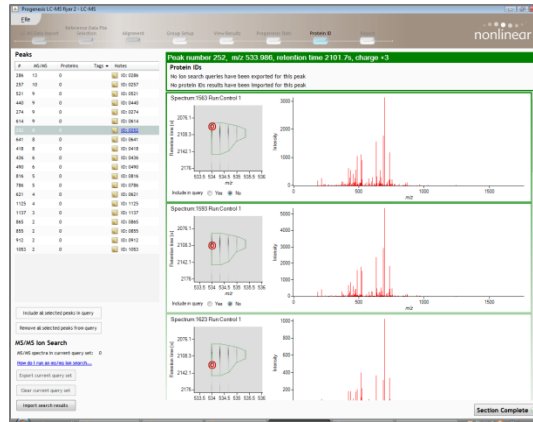
Benefits of the Progenesis approach

- Finds <u>more expression changes</u> with fewer replicates – <u>Statistical power!</u>

- <u>No missed expression changes</u> due to no MS/MS!

- <u>Increases protein coverage</u> from alignment and inclusion lists

- <u>Fast</u>, <u>objective</u> and <u>easy</u> to use!

- <u>No labelling</u> required – suitable for any type of samples.

- <u>Scalable</u> – can handle any experiment size and complexity.

- <u>Flexible</u> – different workflows possible.

- <u>Versatile</u> – compatible with all major equipment and protein search software.

Benefits of the Progenesis approach

- <u>Visual validation</u> - increased confidence in results

nonlinear D Y N A M I C S
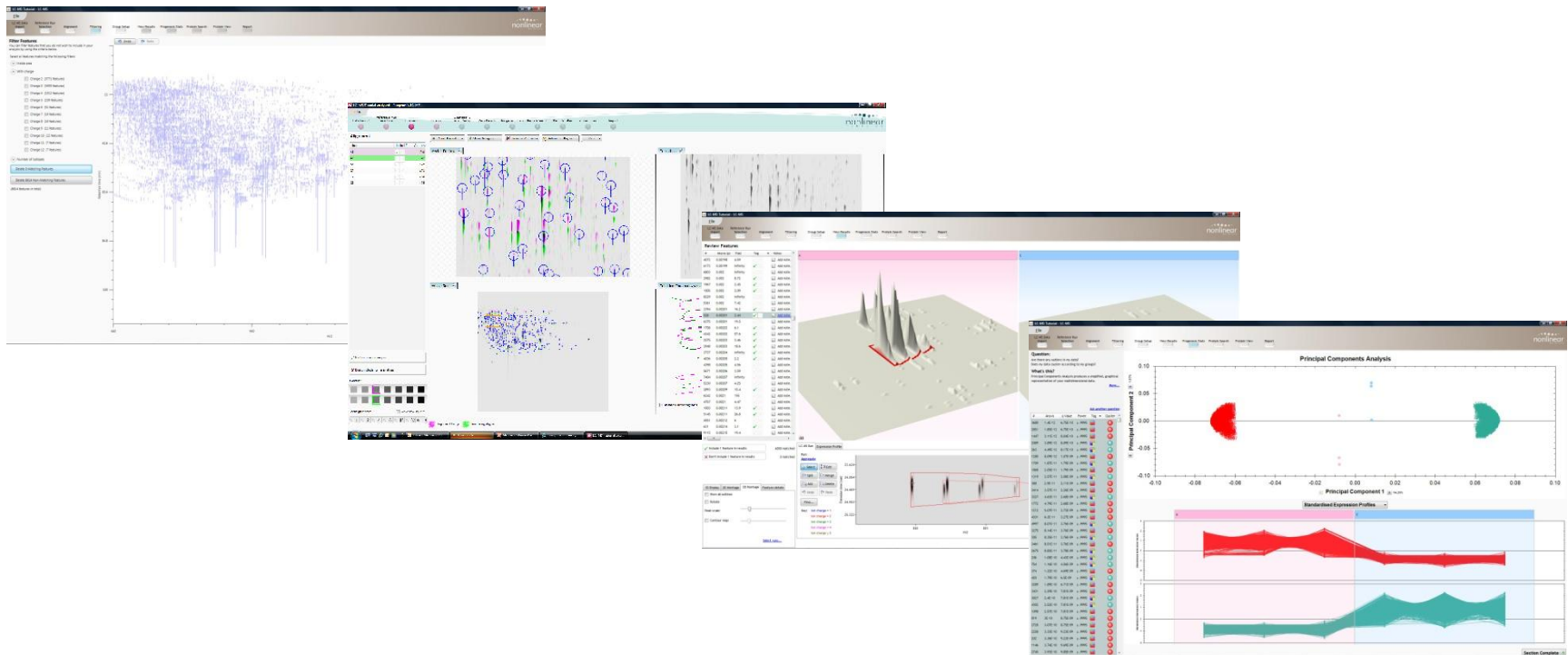
Summary

- Statistics is key to protein expression analysis

- Must interpret statistics properly to find the <u>real</u> expression changes - select only <u>very</u> low p-values (<<0.05).

- Alignment and co-detection technology of Progenesis LCMS maximises statistical power - finds many more expression changes.

- Highly automated technology enables fast, easy and objective analysis!

nonlinear
D Y N A M I C S

Thankyou!

- Progenesis LC-MS can be downloaded from:

www.nonlinear.com