Cube

Computational Systems Biology

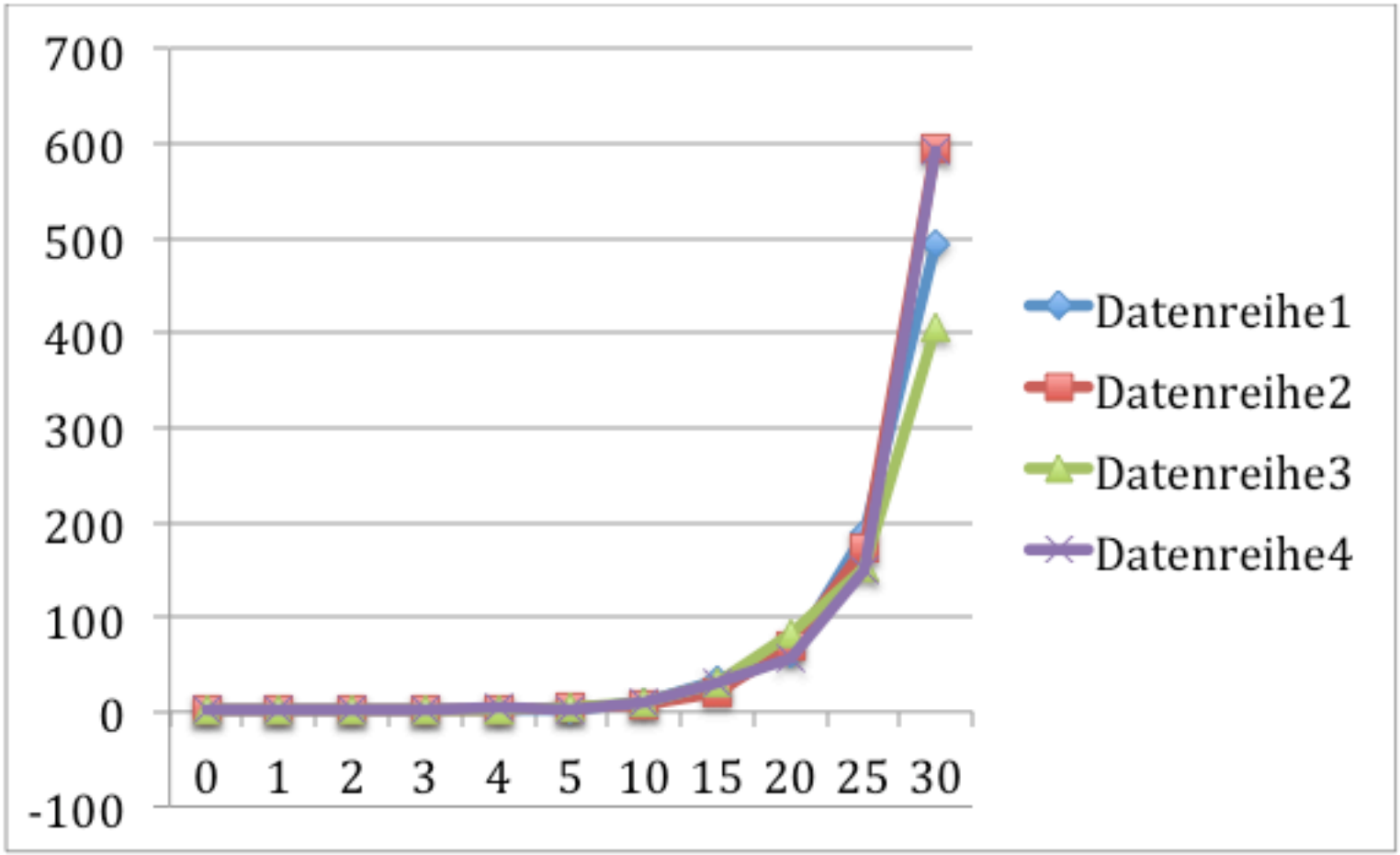
Quantitative Methoden in der Molekularbiologie

4. Protein quantification (2)

WS 2015/16

1. The number of cells in a culture was determined over 30 hours in four parallel experiments. The results have been visualized in the diagram depicted below.

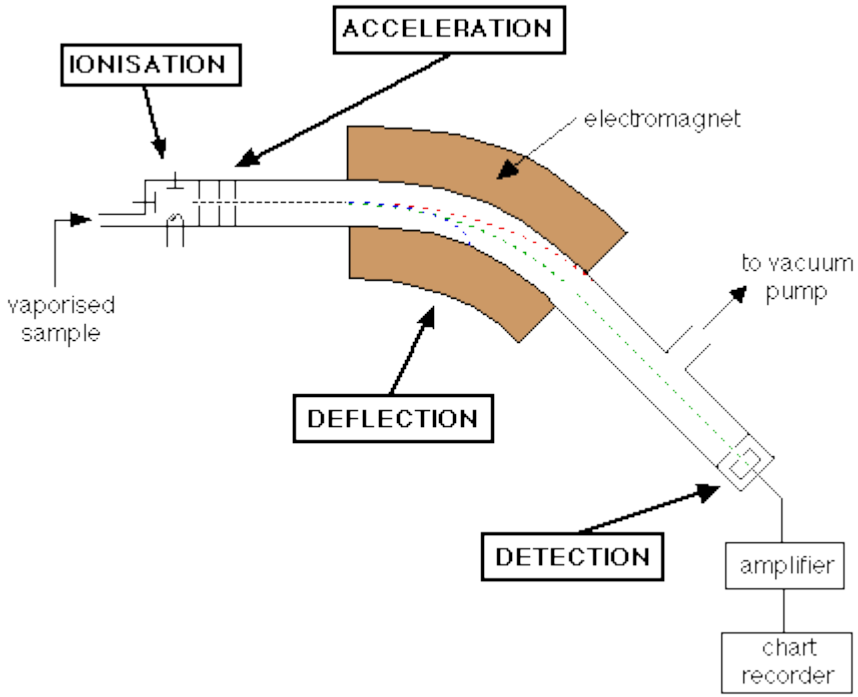
Your task: Assess how meaningful the diagram is and suggest all modifications that would be necessary prior to publication of the diagram.



Outline

1. Analytical methods
   - Protein quantification
   - Mass spectrometry
2. Data analysis
   - Sampling
   - Normal distribution
   - Probabilities
   - Testing normality

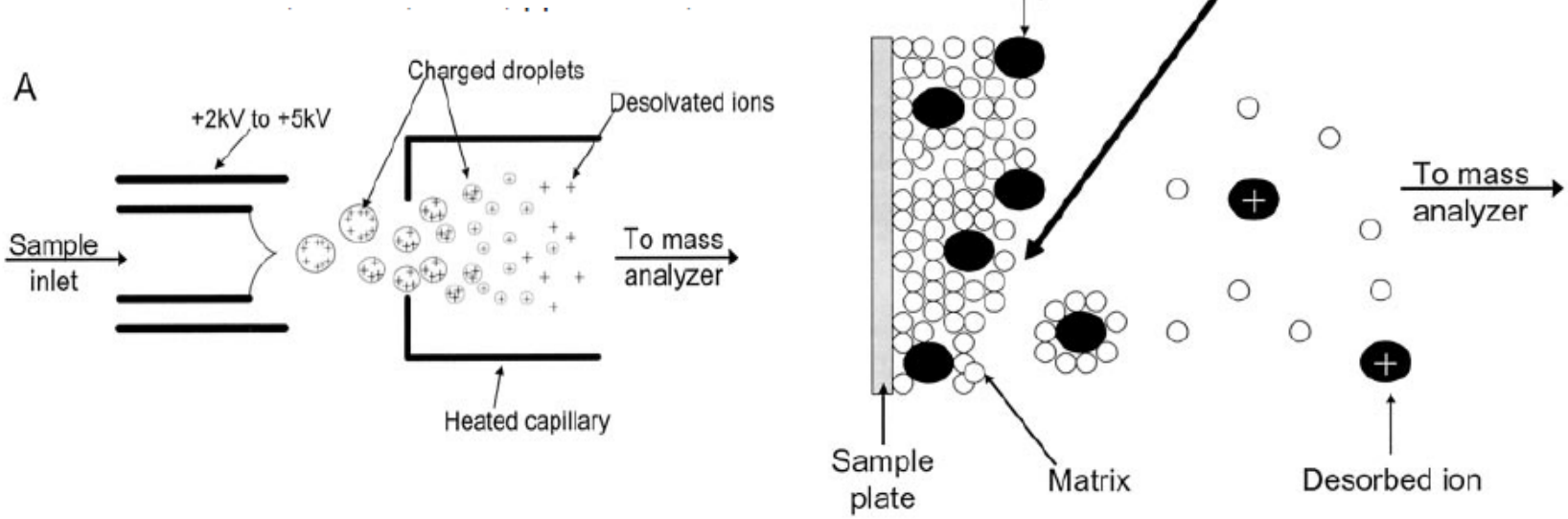
Principles of mass spectrometry



Ionization

- The purpose of ion source is to ionize, and in some cases vaporize, the sample.
- Low pressure is necessary to limit the number of ion collisions

- Methods:
  - ESI (A)
  - MALDI (B)



Acceleration

- Acceleration
  - depends on mass and charge (m/z)
  - two parallel plates; one repels, one attracts and has holes

- Deflection
  - ions pass electromagnetic field
  - ions separated by m/z-ratio

Mass analyzers

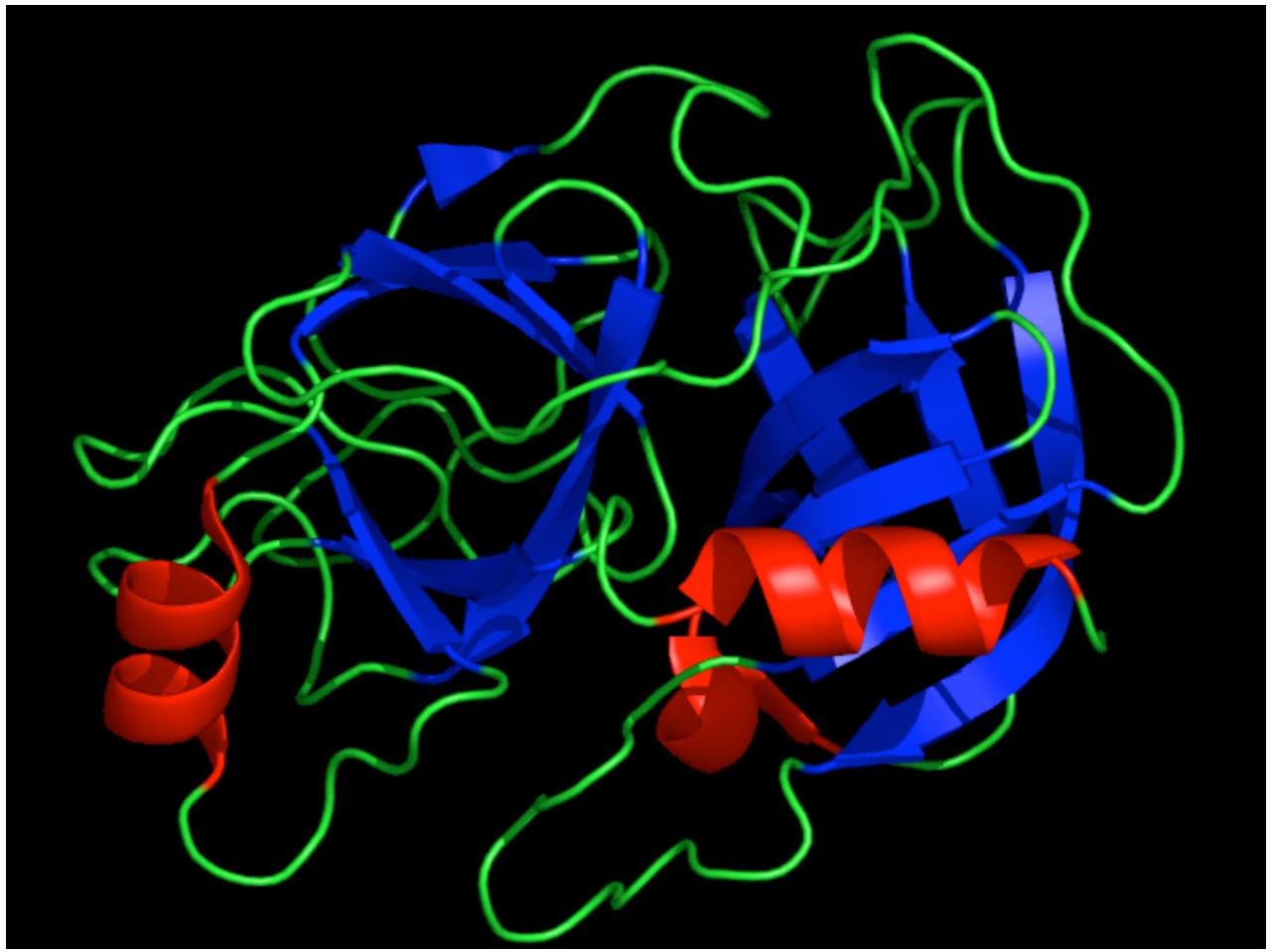
- Magnetic sector
- Quadrupole
- Ion trap
- Time of flight
- Fourier transform-ion cyclotron resonance

Characteristics of analyzers

- Upper mass limit: largest m/z that can be measured
- Transmission: fraction of ions that reach the detector
- Resolution: measure of the ability to differentiate two proximal m/z's

Detector: Secondary electron multiplier

Protein fragmentation: Trypsin



Cleaves after Arg or Lys if not followed by Pro.

PeptideCutter

PeptideCutter [references / documentation] predicts potential cleavage sites cleaved by proteases or chemicals in a given protein sequence. PeptideCutter returns the query sequence with the possible cleavage sites mapped on it and /or a table of cleavage site positions.

Enter a UniProtKB (Swiss-Prot or TrEMBL) protein identifier, ID (e.g. ALBU_HUMAN), or accession number, AC (e.g. P04406), **or** an amino acid sequence (e.g. 'SERVELAT'):

Perform the cleavage of the protein. Reset the fields.

Please, select

- all available enzymes and chemicals
- only the following selection of **enzymes and chemicals**

Arg-C proteinase	Asp-N endopeptidase	Asp-N endopeptidase + N-terminal Glu
BNPS-Skatole	Caspase1	Caspase2
Caspase3	Caspase4	Caspase5
Caspase6	Caspase7	Caspase8
Caspase9	Caspase10	
Chymotrypsin-high specificity (C-term to [FYW], not before P)	Chymotrypsin-low specificity (C-term to [FYWML], not before P)	
Clostripain (Clostridiopeptidase B)	CNBr	Enterokinase
Factor Xa	Formic acid	Glutamyl endopeptidase
GranzymeB	Hydroxylamine	Iodosobenzoic acid
LysC	LysN	NTCB (2-nitro-5-thiocyanobenzoic acid)
Neutrophil elastase		
Pepsin (pH1.3)	Pepsin (pH>2)	Proline-endopeptidase
Proteinase K	Staphylococcal peptidase I	Tobacco etch virus protease
Thermolysin	Thrombin	Trypsin

Protein Quantification using MS

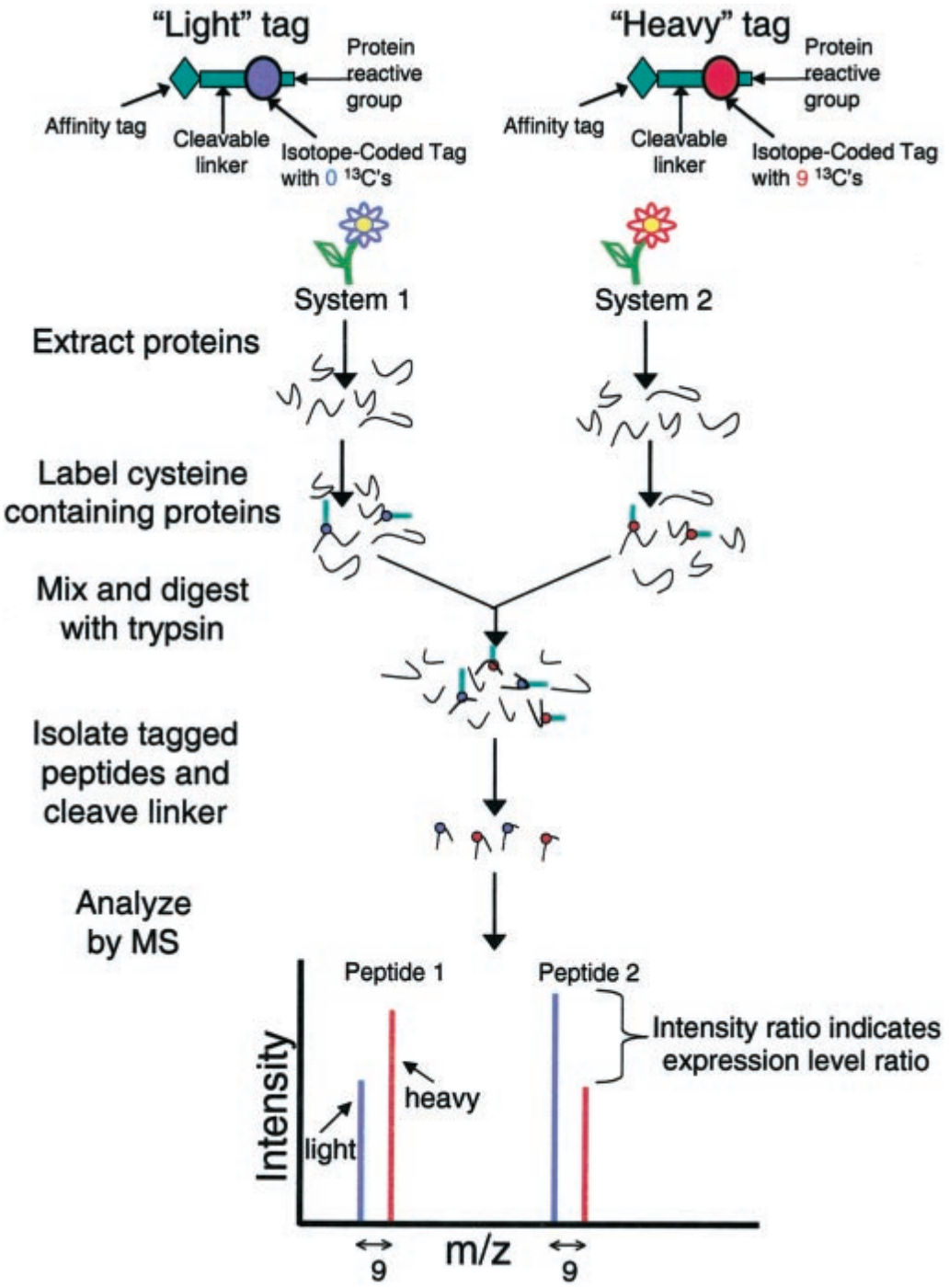
- Relative quantification of two samples

- Isotope-coded affinity tag (ICAT) approach:
  - Uses different isotope containing tags (heavy/ light) to differentiate the samples e.g. H2, C13, N15
  - Tag covalently binds to Cysteine side chains

Reference:
Erin J. Finehout and Kelvin H. Lee, An Introduction to Mass Spectrometry Applications in Biological Research
BIOCHEMISTRY AND MOLECULAR BIOLOGY EDUCATION Vol. 32, No. 2, pp. 93–100, 2004

Isotope-coded affinity tagging

1. Extract proteins
2. Label cysteine groups
3. Mix and Digest with trypsin
4. Isolate tagged peptides using affinity tag
5. Analyze using LC and MS
6. intensity ratios indicate expression level ratio



Outline

1. Analytical methods
   - Protein quantification
   - Mass spectrometry
2. Data analysis
   - Sampling
   - Normal distribution
   - Probabilities
   - Testing normality

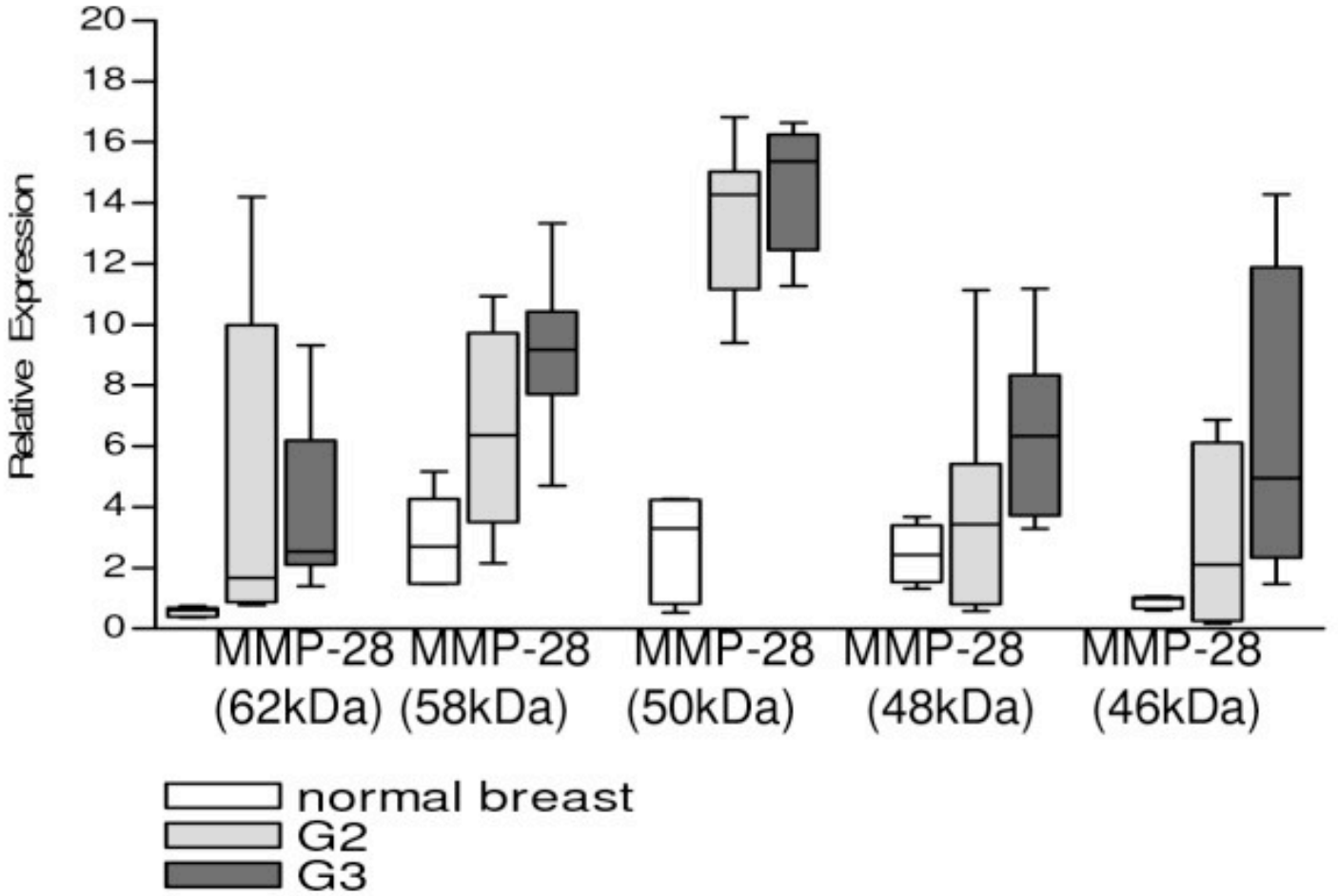
Sampling

- Samples represent population of interest ->„biological replicates“
- Measurements have errors -> „technical replicates“
- Summary measures:
  - Centre (Mean, Median)
  - Spread (Standard deviation, Quartiles)

Median and Quantiles

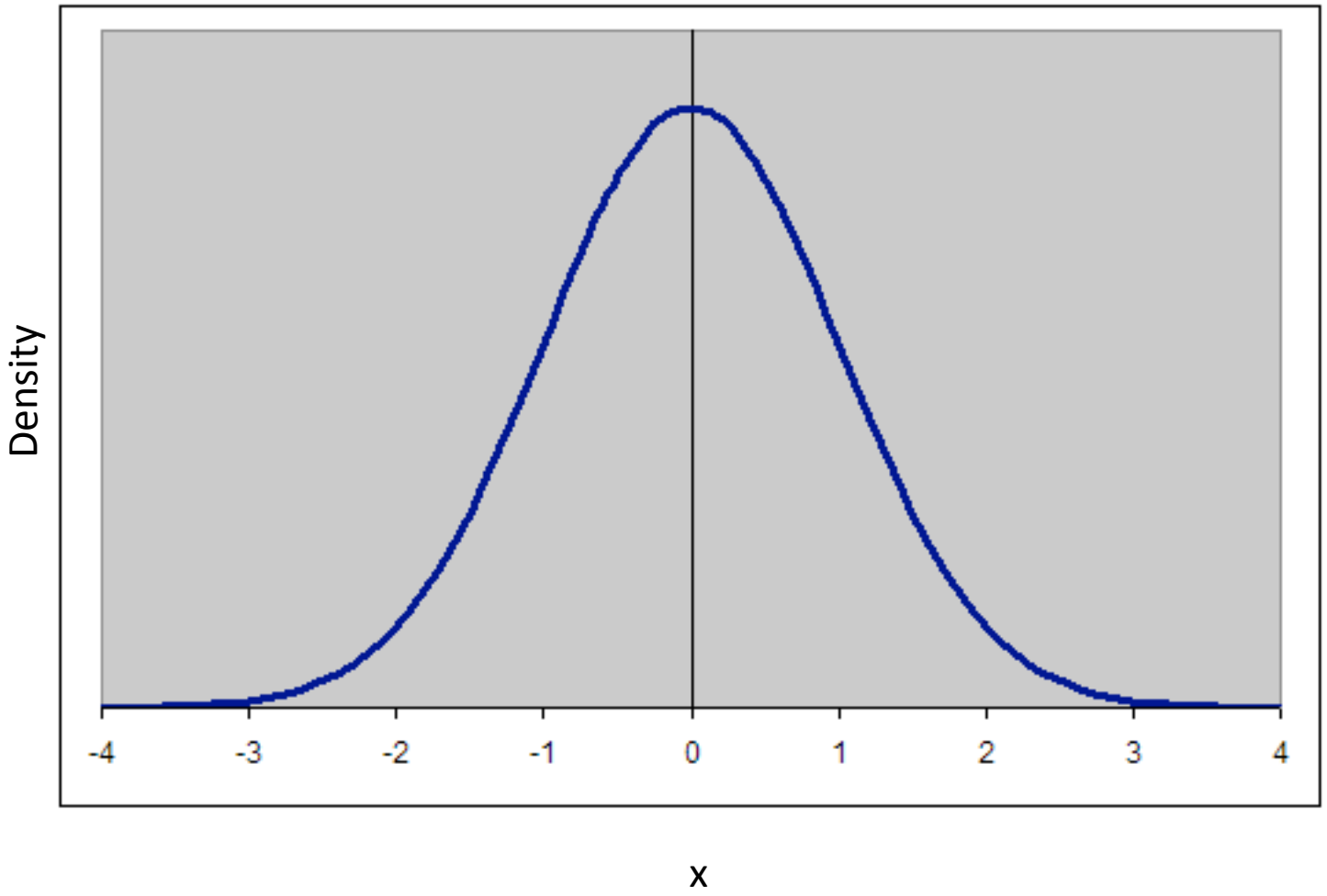
- Non parametric: data not on interval scale -> use rank statistics
- Median: „Middle number“
  - Measurement of rank (n+1)/2
  - Interpolated for even number of ranks
- Quantiles: Measure of scatter
  - Quartiles: Rank of value so that 25% of data is below/ above
  - Rank of lower quartile: (n+1)/4
  - Rank of upper quartile: (n+1)*3/4
  - Interpolated for fractional ranks, weights according to fractions

Example: Boxplots

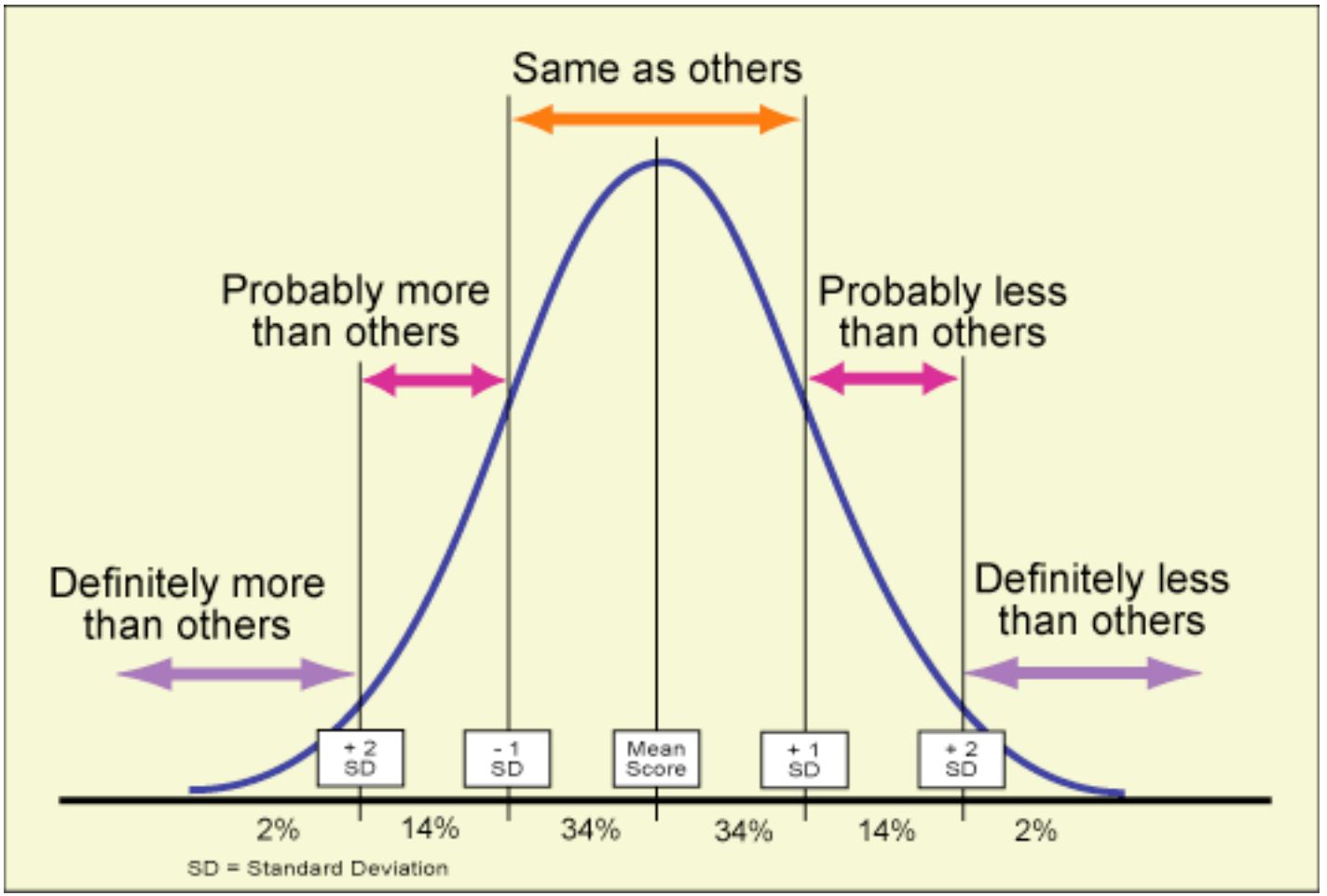


Boxplot analysis of densitometrically quantified expression of MMP protein.Protein levels were normalized to the corresponding expression of β-actin. *Köhrmann et al., BMC Cancer 2009.*

Normal distribution



Probabilities



Probability in normal distribution

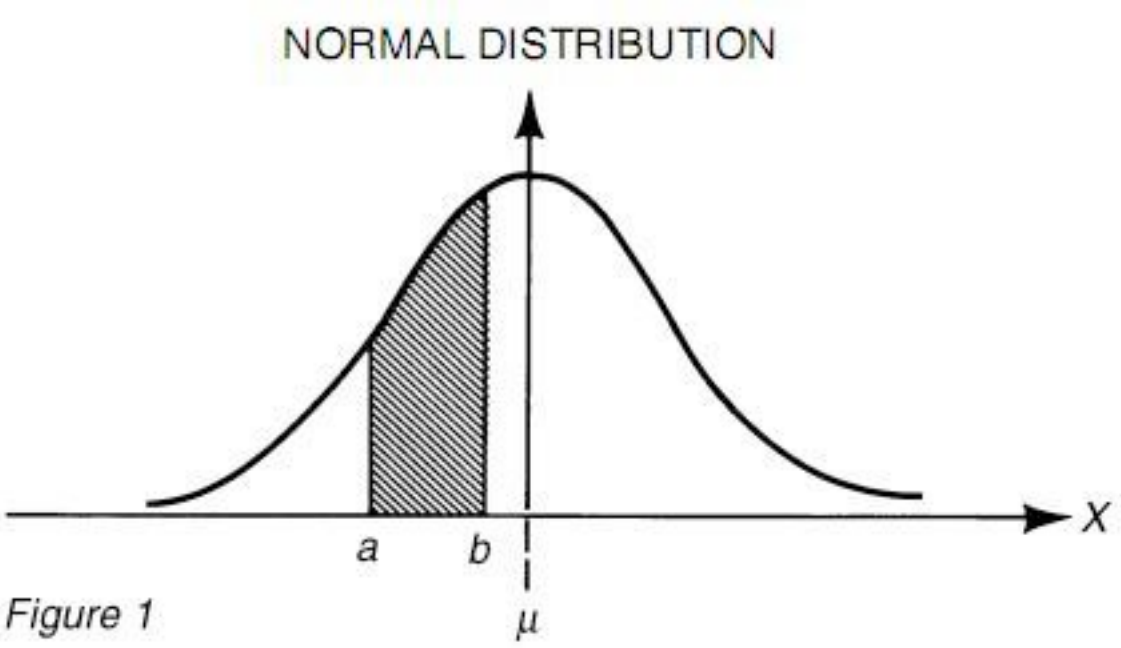


Figure 1

Probability in normal distribution

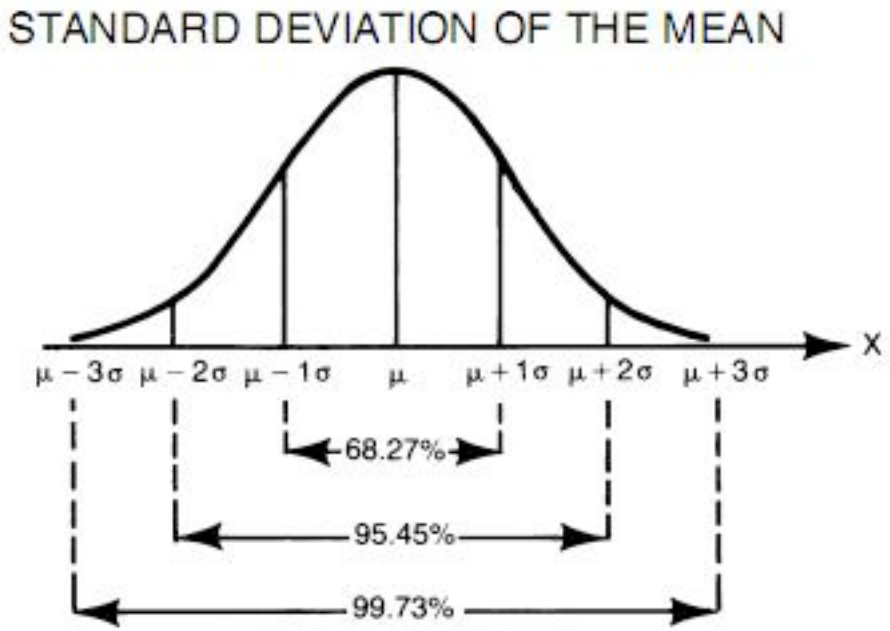
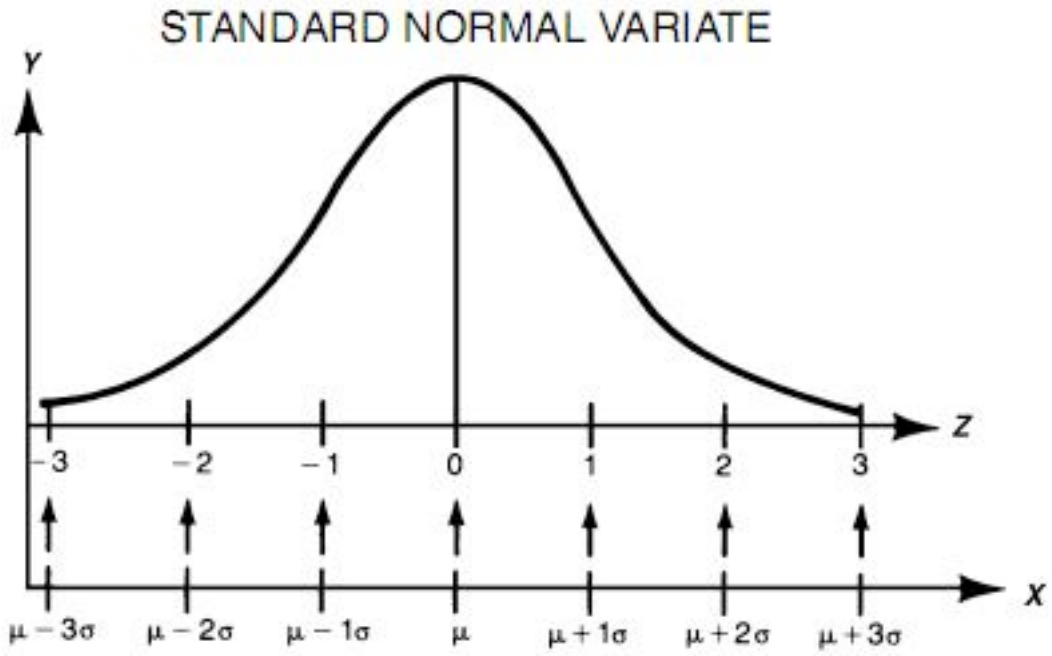


Figure 2

Percent	99.73%	99%	95.45%	95%	90%	80%	68.27%
No. of ± σ's	3.00	2.58	2.00	1.96	1.645	1.28	1.00

Normalization and Z-Score

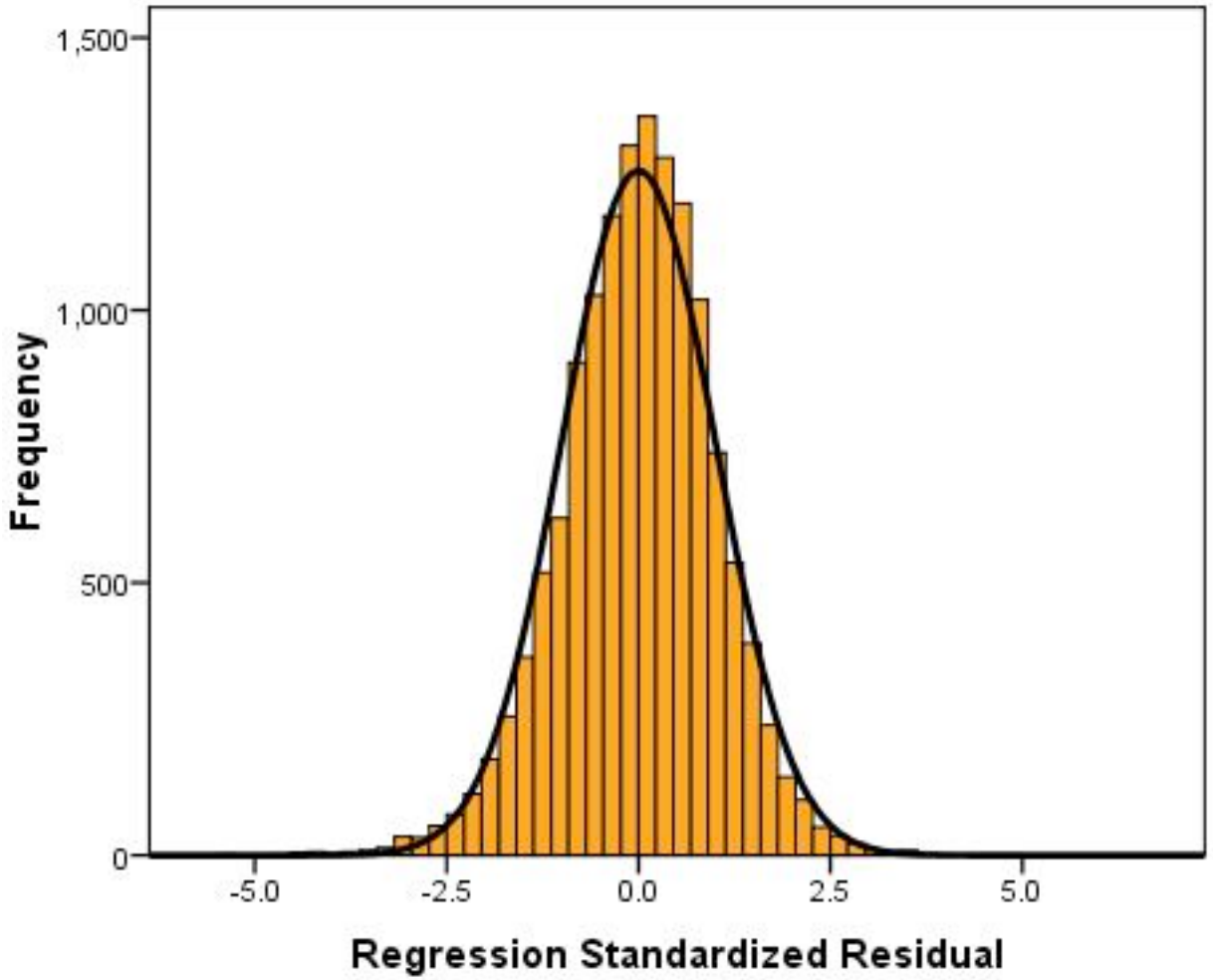
$$Z = \frac{X - \mu}{\sigma}$$



The Translation of X to Z by the Transformation $Z = (X - \mu)/\sigma$

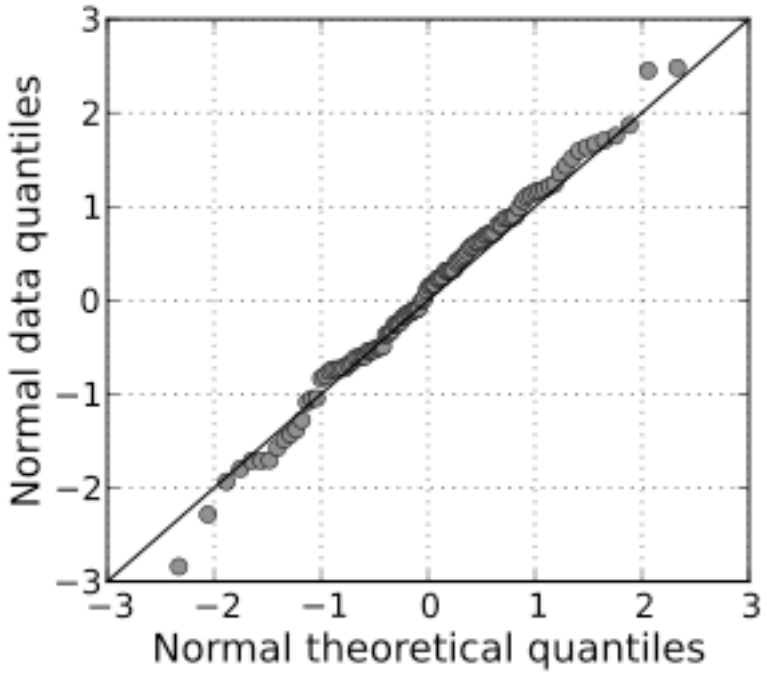
Figure 3

Testing for normality



Histogram of residuals (differences between measured values and mean) compared to theory

Testing for normality



Q-Q Plot: Diagram of data distribution in real and ideal quantiles

Testing for normality

Numerous specialized test methods, e.g. Shapiro-Wilk test:

http://sdittami.altervista.org/shapirotest/ShapiroTest.html