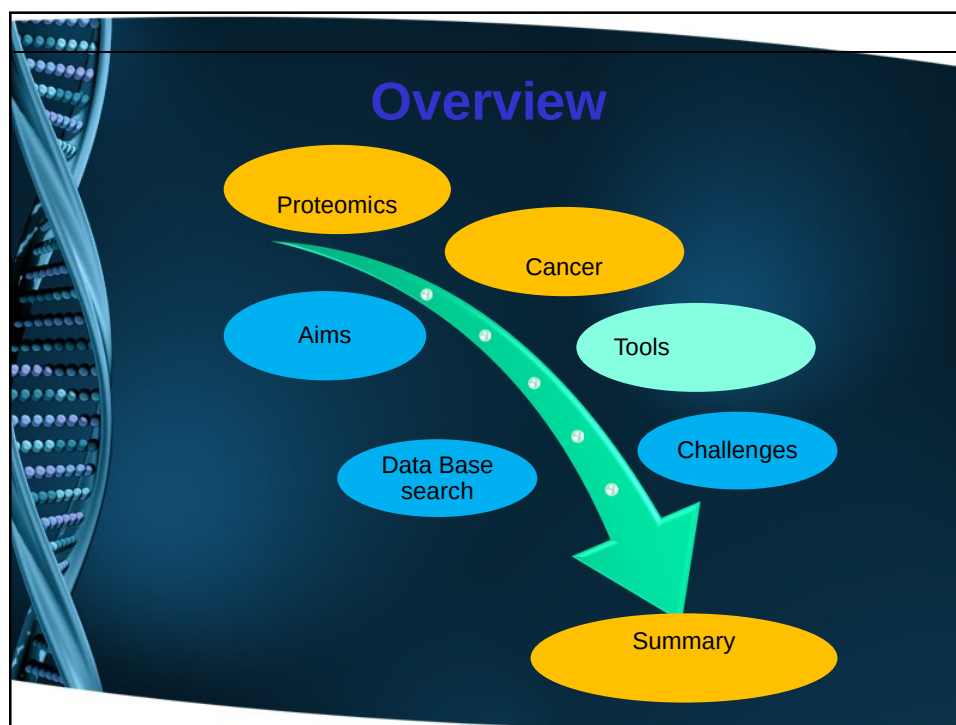
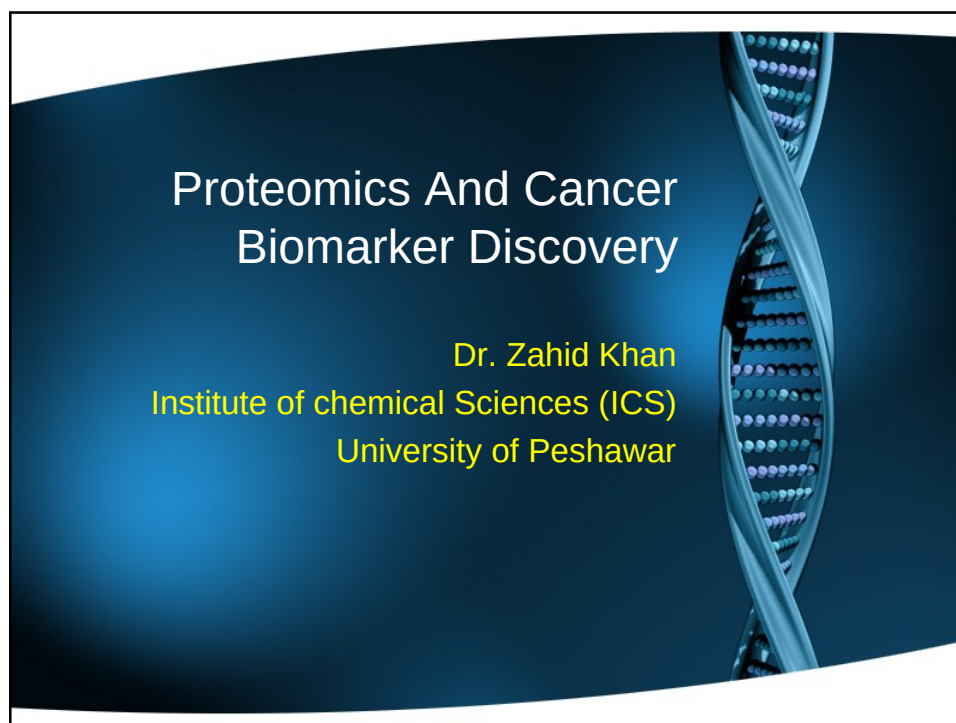
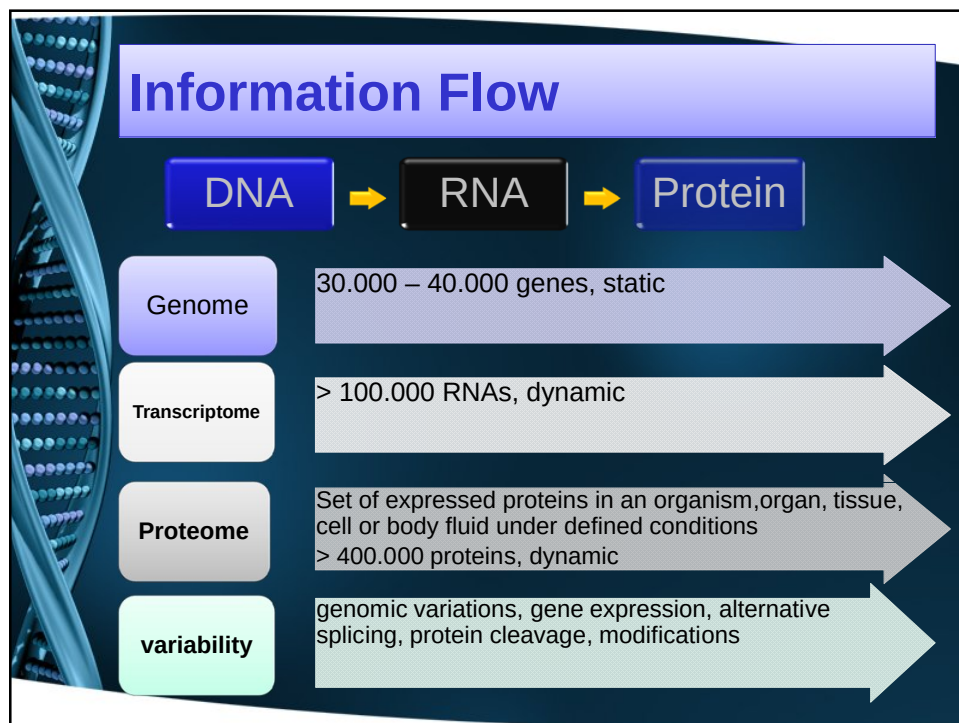
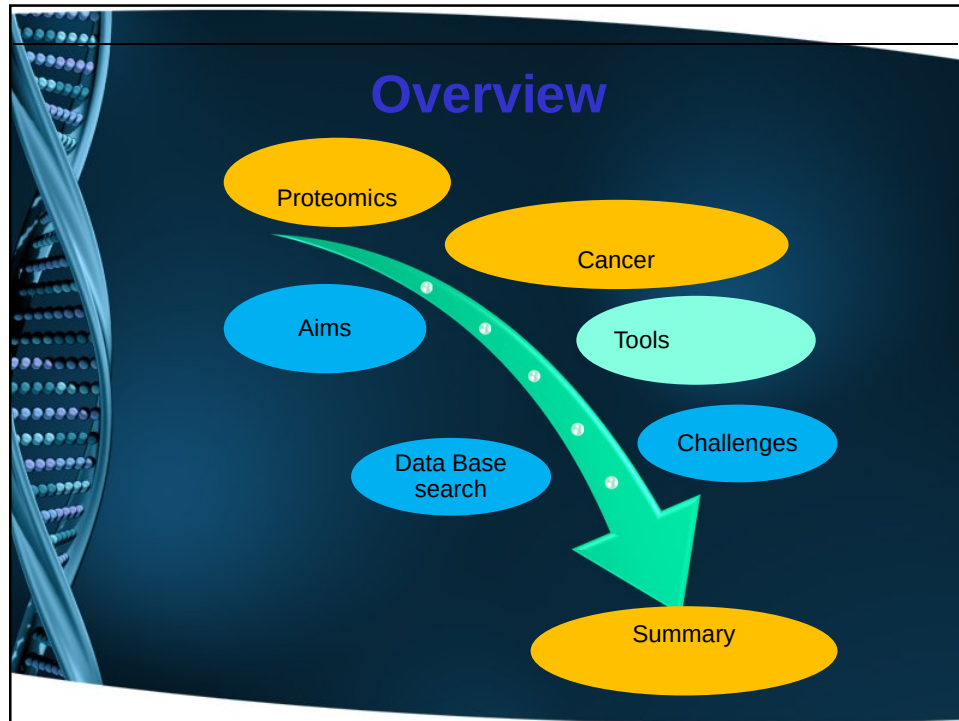
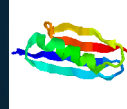




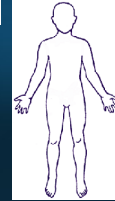


proteomics

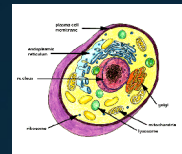
➤ *Proteins*: A chain of amino acids including hormones, enzymes and antibodies.



➤ *Proteome*: All the proteins in a cell or bodily fluid at a given point of time under certain conditions.



➤ *Proteomics*: Proteomics is the study of proteins and proteomes using high-throughput technology.



Many facts of proteomics

■ Proteomic analysis (global profiling proteomics)

- Large scale identification and characterization of proteins and all their properties.

■ Expression proteomics

- Comparison between normal and disease. Also called differential display proteomics.

□ Functional proteomics:

- posttranslational modifications
- protein-protein, protein-ligand interactions
- sequence-structure-function relationships



Why perform proteomics ?

Genomics → Transcriptomics → Proteomics

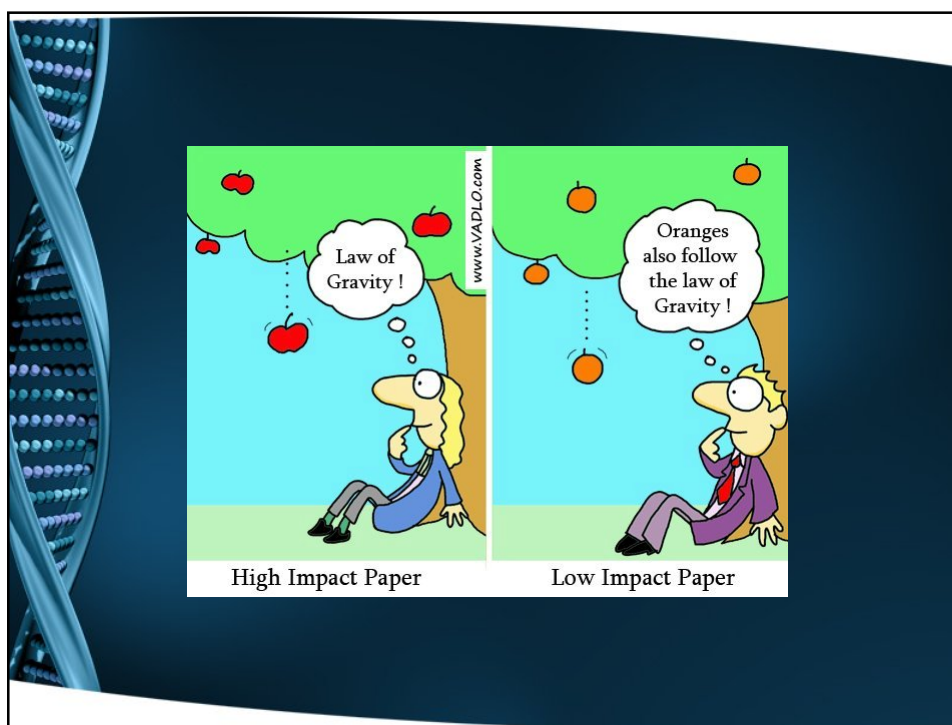
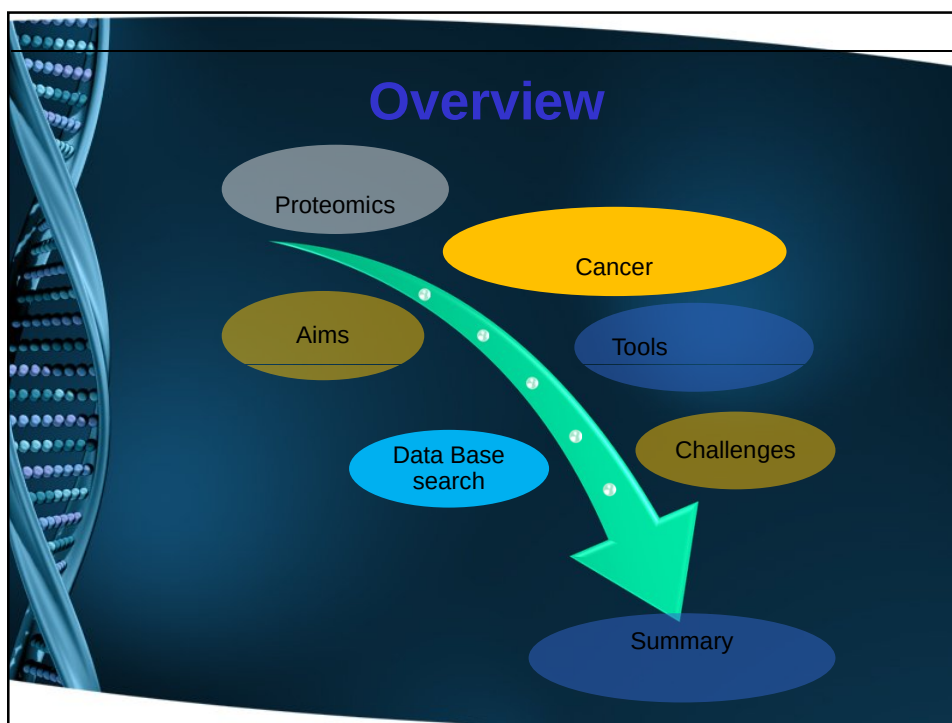
- Protein **alterations** cannot be fully deduced from DNA.
- RNA expression does not always reflect protein levels: translational control, degradation, turnover
- Body fluids are not suitable for RNA expression analysis.
- Proteins are the physiological/pathological active key players.

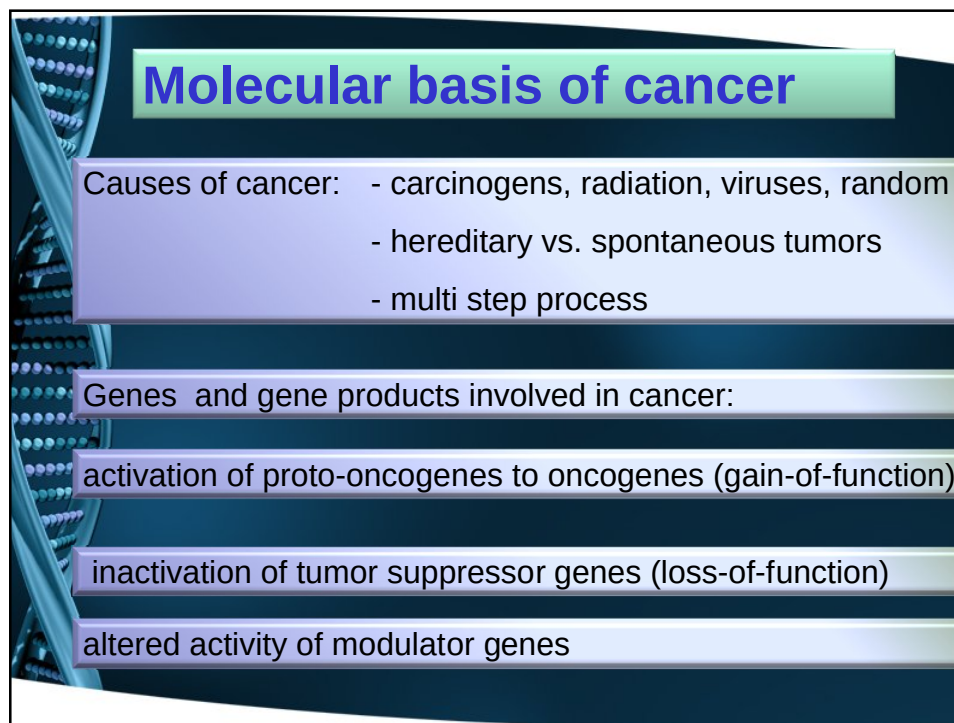
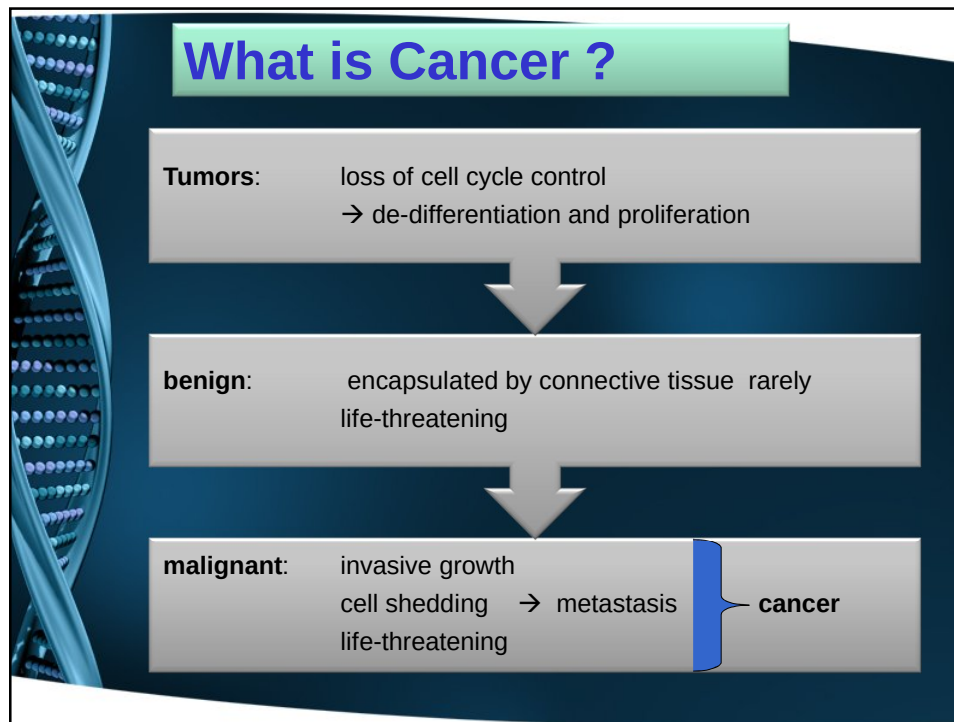


DNA tells us what possibly

RNA what probably and

Proteins what actually happens





Characteristics of cancer cells

- **general changes:**
 - loss of division limits (immortality)
 - uncontrolled proliferation
- **genetic changes:**
 - point mutations ...
 - chromosomal changes
- **structural changes:**
 - less organized cytoskeleton
 - increased membrane fluidity
- **biochemical changes:**
 - altered protein expression
 - altered protein modification

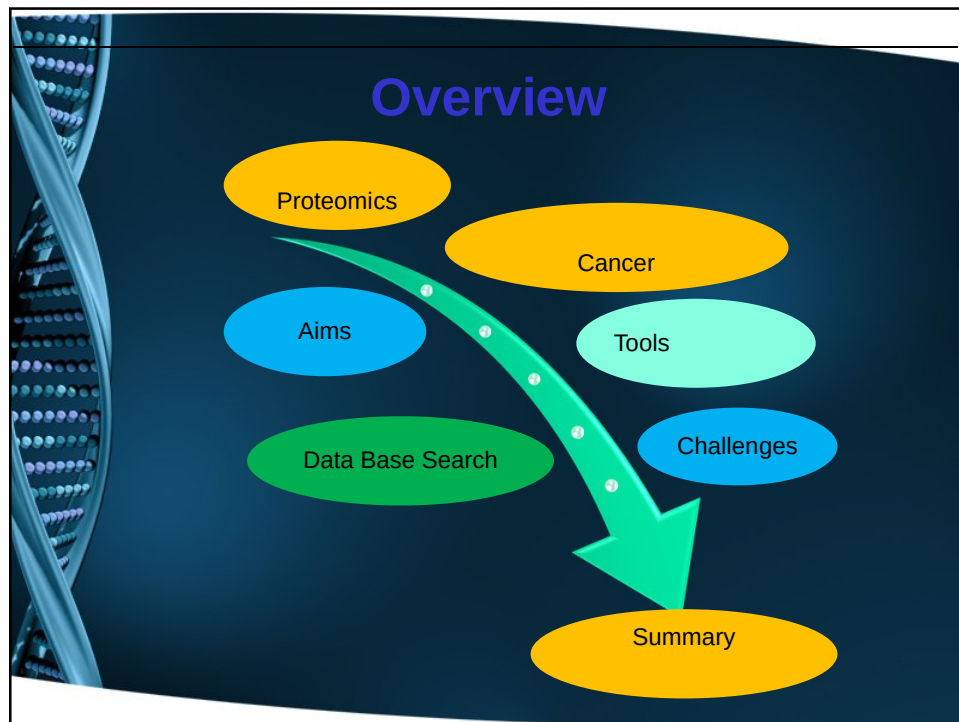
Cancer facts and treatment

- > 100 different types of human cancer
- ≥ 20 % of the mortalities in industrialized nations

detection → classification and localization → therapy

imaging
histology
biomarkers

surgery
radiation
chemotherapy





Proteomic aims in cancer research

General goal:

- better understanding of genesis and progression of cancer

Clinical goals:

- early cancer detection using biomarkers
- identification of potential therapeutic target structures
- efficient monitoring of therapy control

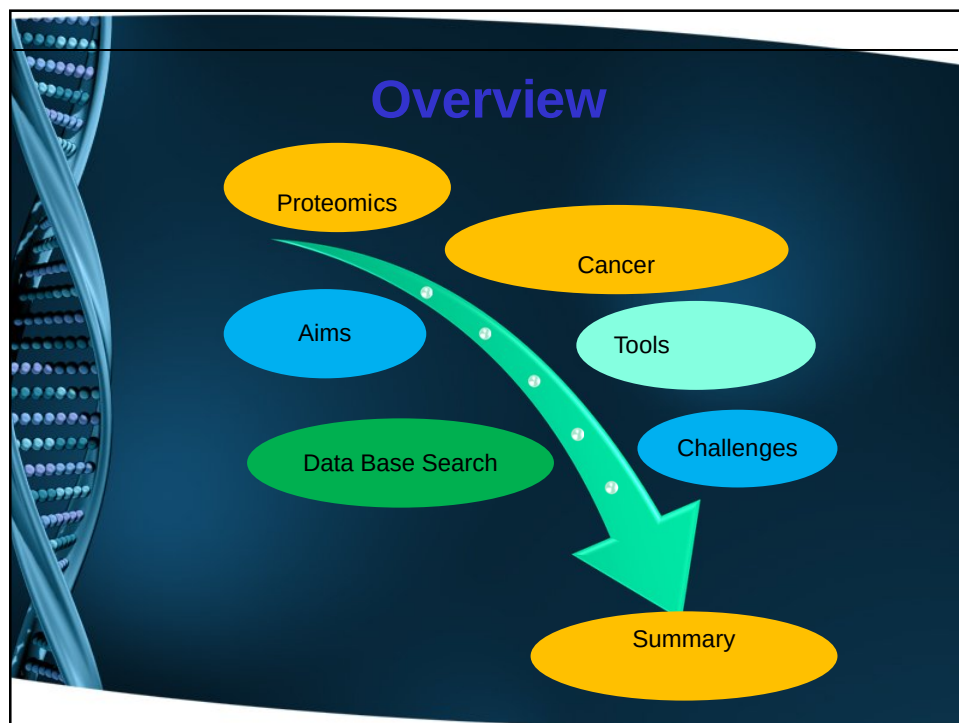


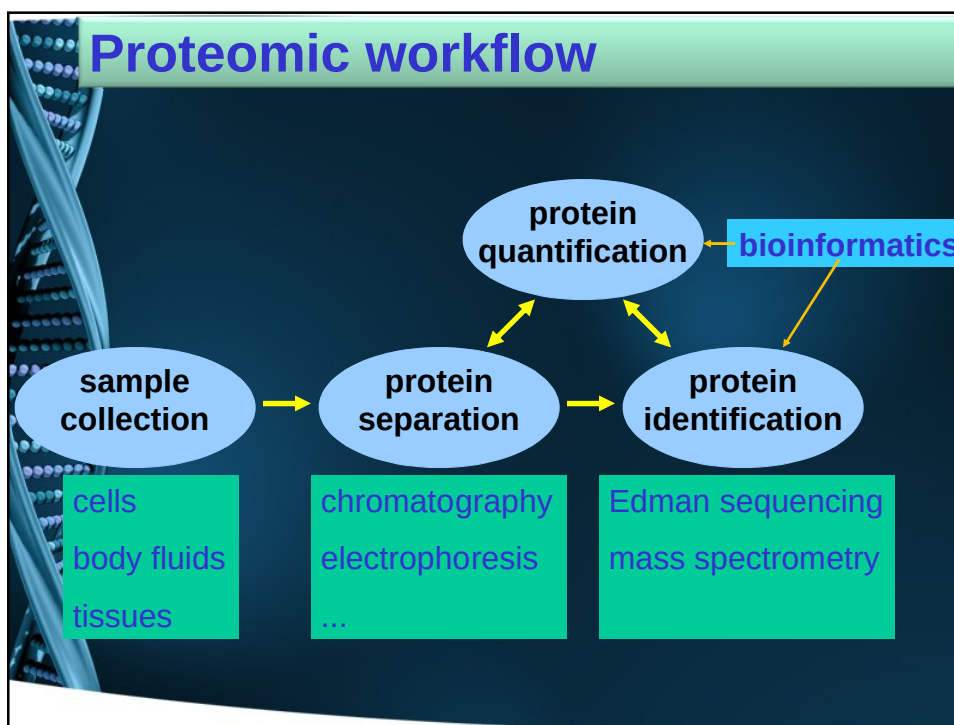
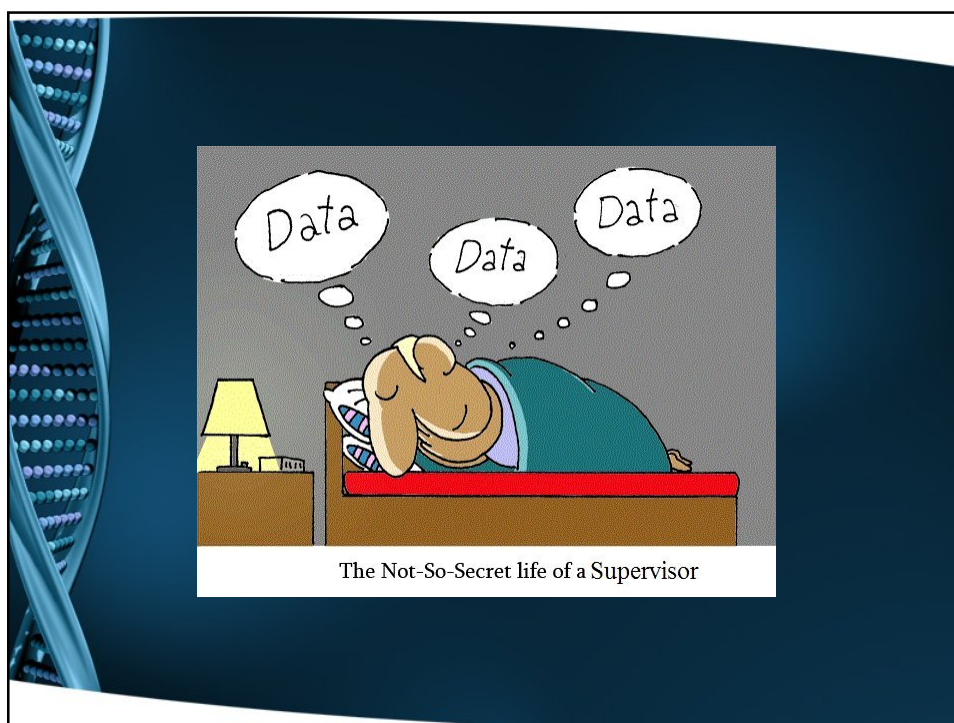
Biomarker

- Substances present in increased or decreased amounts in body fluids or tissues that indicate exposure, disease or susceptibility to disease.

Tumor markers

- **diagnostic markers:** detection of malignant disease
- **sensitivity** = detected positives / real positives
- **specificity** = detected negatives / real negatives
- **prognostic markers:** malignant potential, disease recurrence
- **predictive markers:** response to different therapies
- **positioning markers:** positional information (e.g. for surgery)





Sample collection

- **sample type**: cells, body fluids, tissues
- **time** of sample collection (daily and monthly variations)
- **prevention of contaminants** (necrotic tissue, blood ...)
- **biological variability** (patient-to-patient, tumor-to-tumor)
- **identical collection, treatment and storage**

"The devil is in the details !"



<http://www.meine.de/images/teufel.gif>

Sample collection: Tissue and LCM

- **Laser capture microdissection (LCM):**
Laser-cutting of stained tissue slides under microscopic control
→ defined procurement of cells
- **intra-patient control:** tumor vs. non-tumor tissue





Protein separation

- Protein-specific **biochemical-biophysical parameters**:
isoelectric point (IP), molecular weight (MW), affinity.
- **chromatographic methods**:
HPLC
2D-HPLC
ProteinChips
- **electrophoretic methods**:
SDS-PAGE
2-dimensional (2DE)

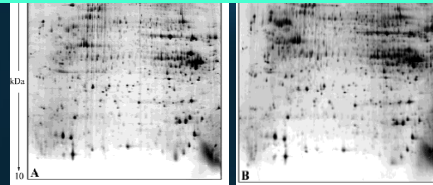
2-Dimensional gel electrophoresis

- 1. dimension: isoelectric focussing (IEF) → separation by **IP**
- 2. dimension: SDS-PAGE → separation by **MW**

- **staining** → > 1000 proteins /gel

- Example: breast cancer cell lines:

Fontana *et al.* Proteomics 2004, 4, 849



- **molecular analysis** by mass spectrometry, HPLC, Westernblot ..

DIGE

- DIGE can be done in one-or two-dimensions. Same principle.
- Requires fluorescent protein stains (up to three of these), a gel box, and a gel scanner.
- Dyes include Cy2, Cy3 and Cy5 (Amersham system).
- These have similar sizes and charges, which means that individual proteins move to the same places on 2-D gels no matter what dye they are labeled with.
- Detection down to 125 pg of a single protein.

Cyanine

Probe	Ex (nm)	Em (nm)	MW
Cy2	489	506	714
Cy3	(512);550	570;(615)	767
Cy5	(625);650	670	792

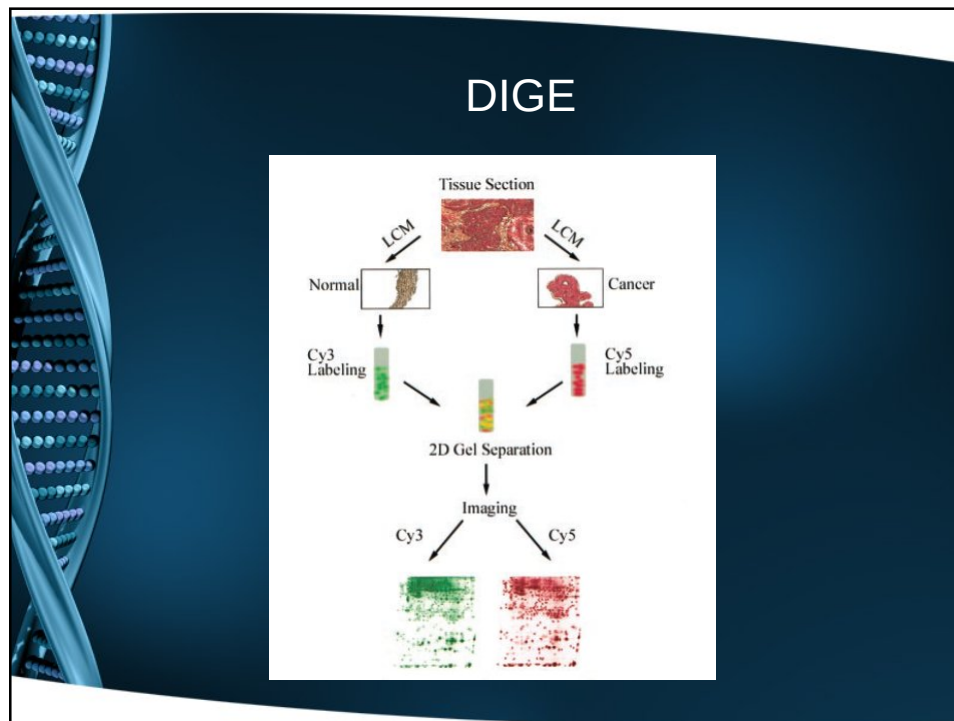
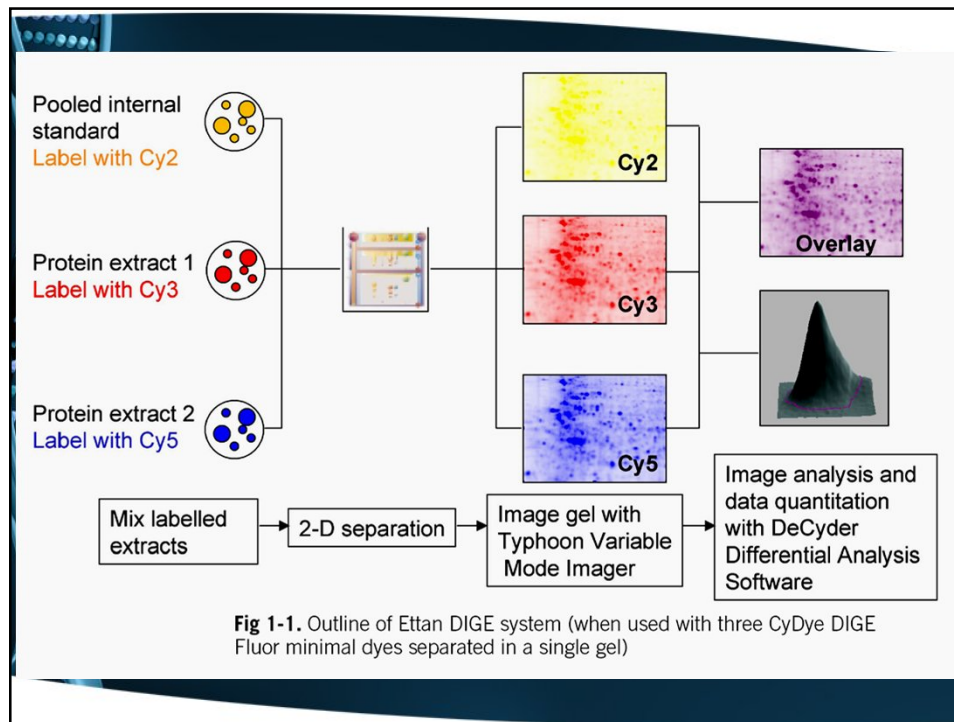
Cy3

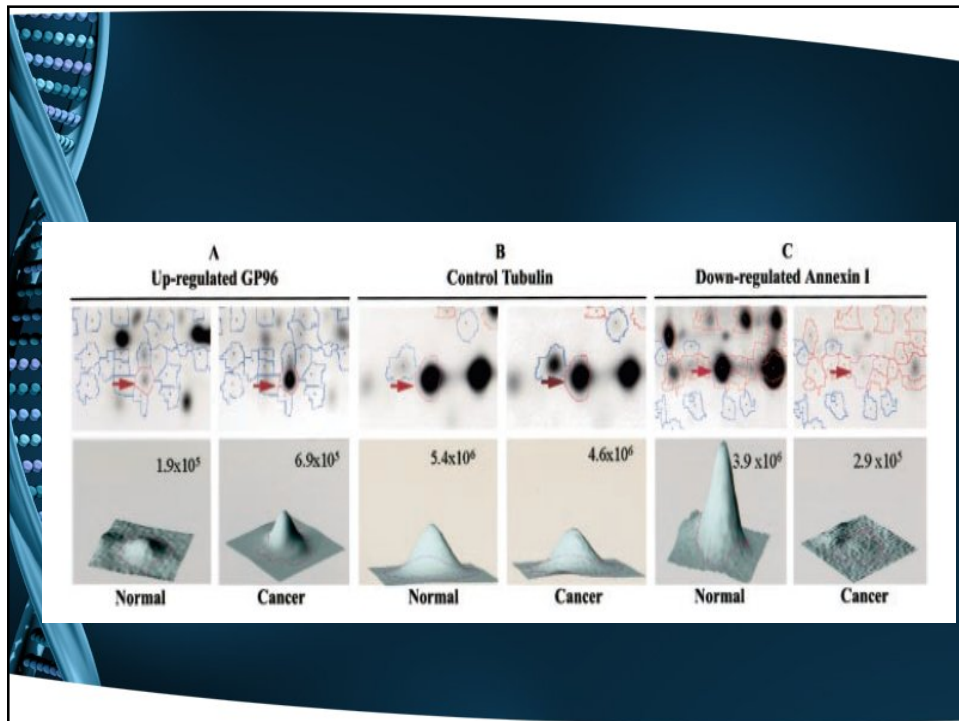
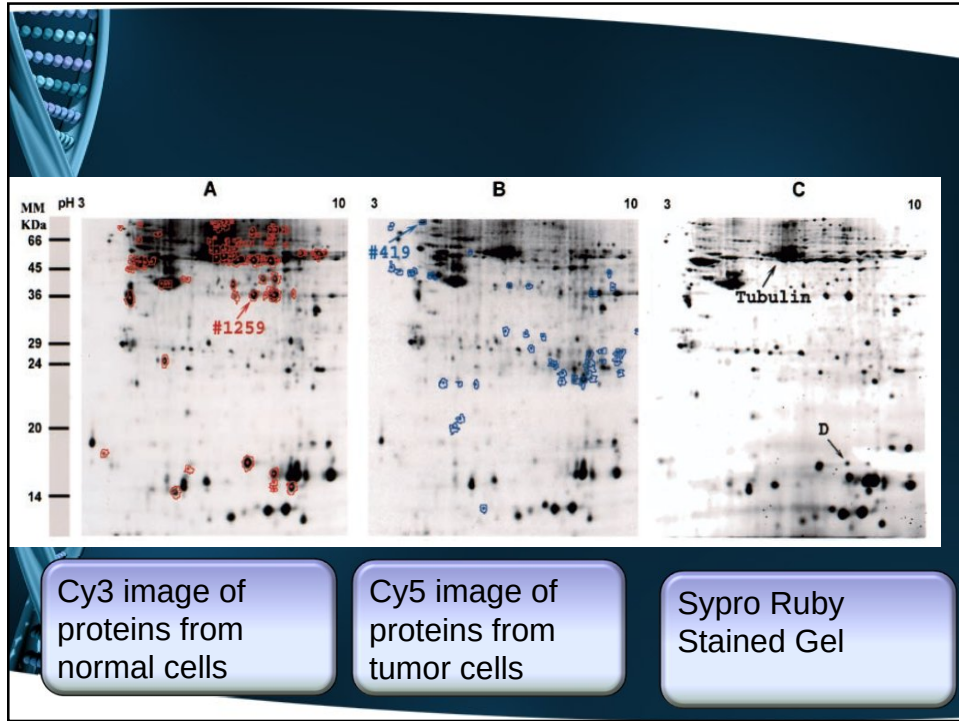
Cy5

- Linear response to protein concentration over a 10^5 concentration range.
- Standard chemistry links dyes to proteins via lysine side-chains.
- Proteins samples are produced by homogenizing cells.

NHS ester reactive group

Fig 1-2. Schematic of the minimal labelling reaction. CyDye DIGE Fluor minimal dye containing NHS ester active group covalently binds to the lysine residue of a protein via an amide linkage.





Protein identification

- Identification and characterization (modifications) by combined **bioanalytics** and **bioinformatics**

Edman sequencing: N-terminal sequence tags

mass spectrometry: peptide finger prints (digests)
(tandem MS, MS/MS, MSⁿ)

- **data bases** and **bioinformatic algorithms** (scoring)

Mass spectrometry

- Measurement of mass/charge (m/z) ratios

ionization

matrix-assisted laser
desorption/ionization
(MALDI)
electrospray ionization
(ESI)



m/z analysis

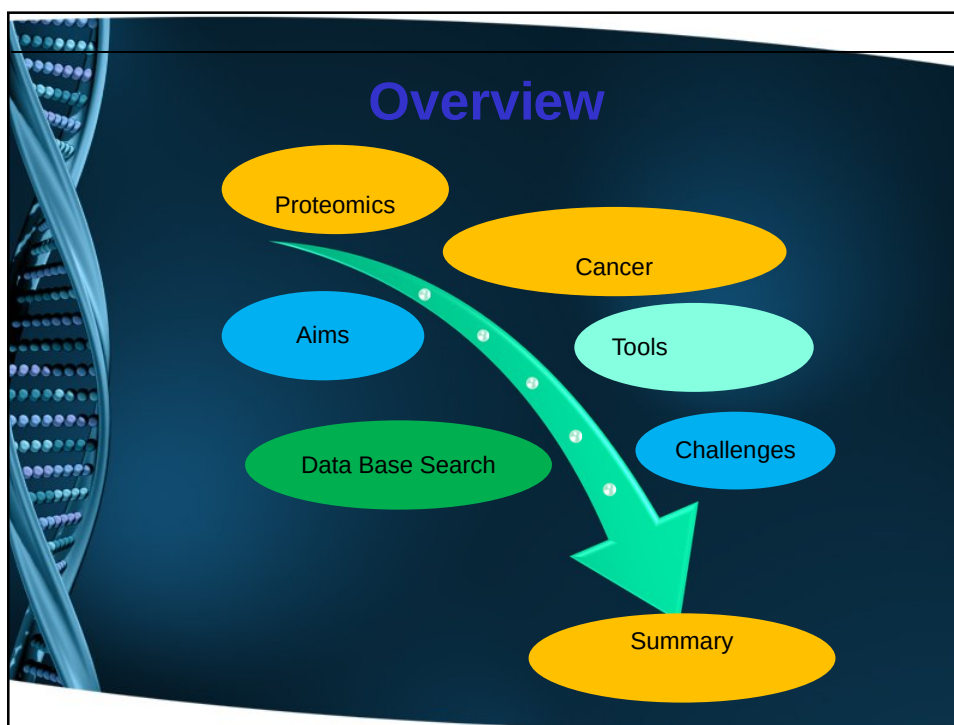
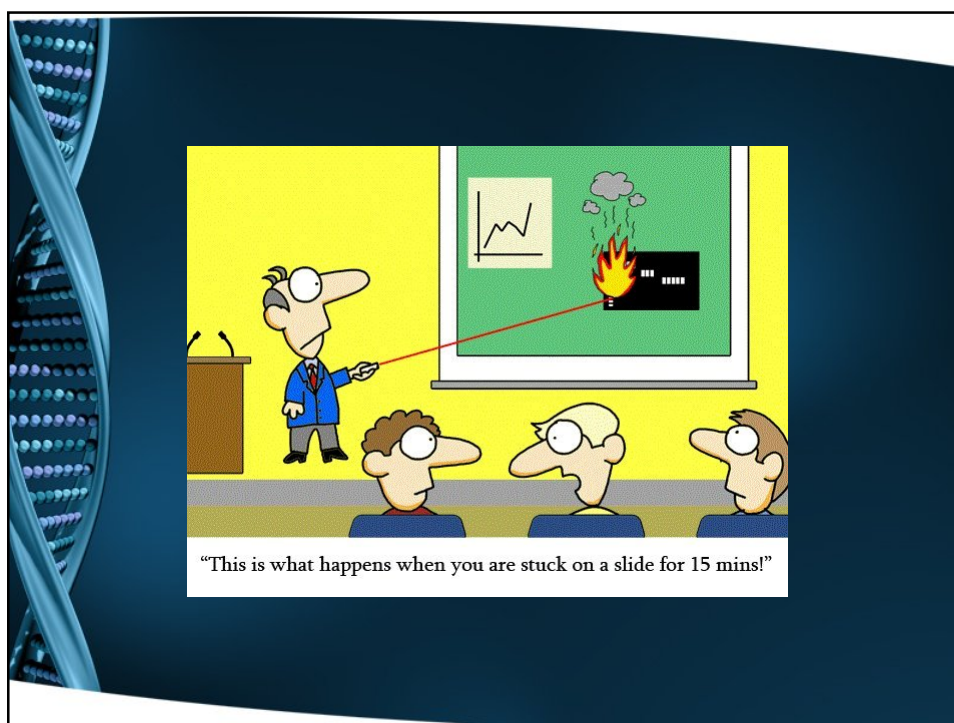
time-of-flight (TOF)
quadrupol (Q)
ion trap (IT)
etc



detection

secondary electron
multiplier
micro channel plates

- **different combinations:** accuracy, sensitivity, mass range ...





Database search

Proteomics & Bioinformatics / MBI
 PAPDFKATAASGNARIPAPDFKATAA
 SCAYVVKQASASCHAAVVKQKQSA
 SYVVLFFVSDVKGKFPYVVLFFVSDY
 KPLDLFFVPLDFTTTPPLDLFFVDFE
 EKATASACHARYVVLFDATSGARA
 YVRVKQVLVGNARAYVVKQVYVLA
 RSGNKATAASGNTRKSGNSIEVPLDFT
 LPDETATA

Peptide mass fingerprinting provides evidence for the **most probable identity** of a protein.

The genome should be verified for the organism that you are working on.

If not, then the next most ideal situation is that there is good cDNA data available.

If neither of these are the case then it is worth checking if there are any expressed sequence tags (EST) that can be used.

The quality of the Protein identification will depend upon:

- Quality of the mass spectrometry data
- The accuracy of the database
- The power of the search algorithms and software used



Tools for fingerprinting

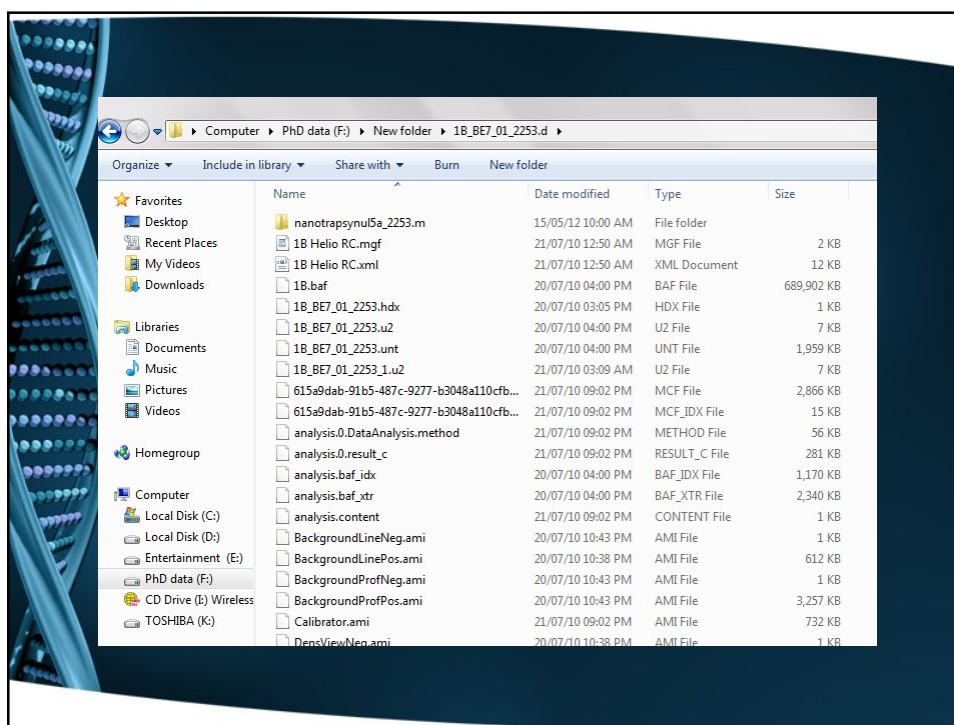
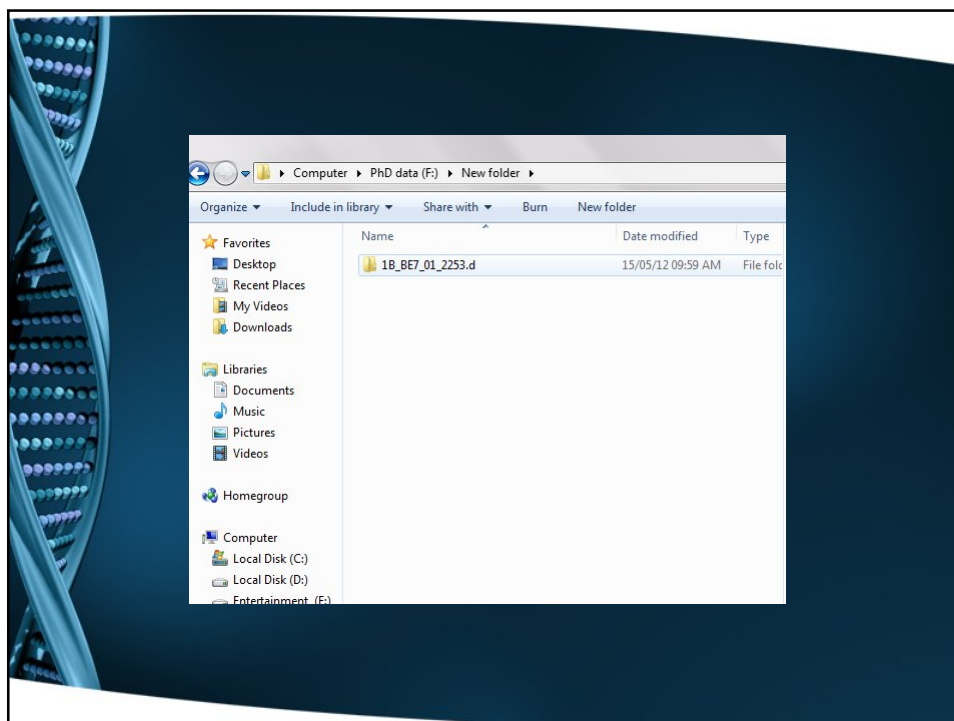



Mascot

Aldente

Ms-Fit

Profound





```

2253.d)
### Instrument: microTOF-Q
### 1B
### PeakWidth 5.00, MiniumIntensity 50, S/N 10
### Retain Residuals no, Create neutrally no,
### Proteins Parameters: AdductIon (+)+H (-)-H, LowMass 1000,
HighMass 20000, MaxCharge charge decon 50, MinPeaks 3
### Perf. isotope decon no, Max Charge Istopic decon 2, MW
agreement 5.00, Abund. cutoff 10.00, Envelope cutoff 75.00
### Global Charge Limit: no 3, Prefer FullScan Result: yes,
Export what: 2 (0=non deconv/1=deconv only/2=mixed),
NoMSMSsignals: 50, Intensity threshold: 100, NoMSsignals
(deconv): 2, NoMSsignals (non-deconv): 2, NormalizeData: no,
SingleCharged: no
###FS: #m/z: 1068.069733 #charge-128
###MaxRes:#m/z: 0.000000 #charge-128
###MS: 4004/
###MSMS: 4007/
BEGIN IONS
TITLE= Cmpd 1, +MSn(1068.069702), 35.5 min
PEPMAS=1068.0697334 18947

101.0674348 238
249.1575387 1488
277.1510831 1338
292.1219320 1146
345.2172017 675
367.2287165 665
395.2248222 732
405.2063915 5598
406.2132207 1088

```



MASCOT Peptide Mass Fingerprint

Mascot PMF

Your name: Proteomics & Bioinformatics Email: MBI@Helsinki.fi

Search title: Peptide Mass Fingerprint Example

Database: SwissProt

Taxonomy: NCBI

Enzyme: SwissProt

Fixed modifications: ABRF_PRG_2007

Protein mass: 814.430000

Mass values: 958.350000

Data file: 1000.330000

Query: 1165.390000

Decoy: 1182.440000

Missed cleavages: (K)

Protein N-term: (N-term)

Protein C-term: (C-term)

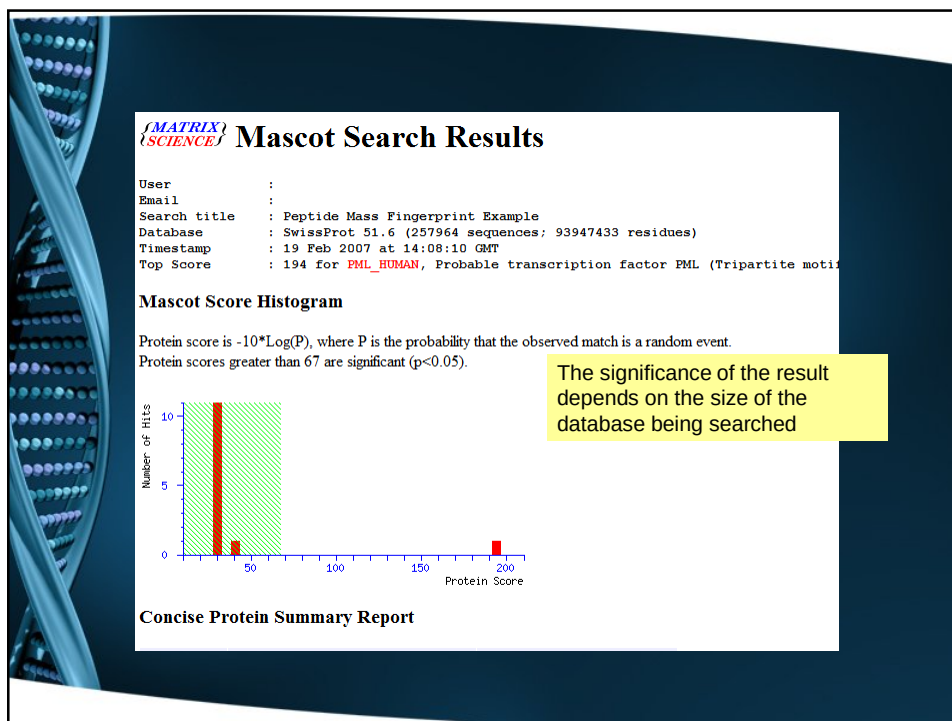
Protein C-term: (Protein C-term)

Da: Da

average: average

hits: hits

Get Form



Mascot PMF results

Entry name	Mass	Score	Expect	Queries matched	Probability to be random	Coverage similarity
PML_HUMAN	97455	Score: 112	Expect: 1.7e-06	Queries matched: 15		
Probable transcription factor PML (Tripartite motif-containing protein 19) (RING finger domain)						
ANMK_RHOF2	38245	Score: 51	Expect: 2.2	Queries matched: 7		
Anhydro-N-acetylmuramic acid kinase (EC 2.7.1.-) (AnhMurNAc kinase) - Rhodospirillum rubrum						
PSF2_DEBHA	18601	Score: 47	Expect: 4.9	Queries matched: 6		
Penicillin-binding protein 2 (PBP2) - Bacillus subtilis (strain 168)						
PURA_THEVO	48976	Score: 40	Expect: 26	Queries matched: 6		
Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase) (AdSS) (AMPSase) - Thermotoga maritima						
KSGA_BRUMI	30303	Score: 40	Expect: 26	Queries matched: 6		
Dimethyladenosine transferase (EC 2.1.1.-) (S-adenosylmethionine-6-N', N'-adenosyl transferase) - Rhizobium sp. (strain NGR234)						
GNAS3_BOVIN	27477	Score: 38	Expect: 42	Queries matched: 6		
Neurogranin secretory protein 55 (NESP55) [Contains: LSLAL tetrapeptide; GAIPRRH peptide] - Bos taurus						
Y4GB_RHISN	16103	Score: 37	Expect: 52	Queries matched: 4		
Hypothetical 16.1 kDa protein y4gB - Rhizobium sp. (strain NGR234)						
SYQ_HUMAN	87743	Score: 36	Expect: 74	Queries matched: 9		
Glutamyl-tRNA synthetase (EC 6.1.1.18) (Glutamine--tRNA ligase) (GlnRS) - Homo sapiens						
YDIU_SHISS	54287	Score: 35	Expect: 75	Queries matched: 6		
UPF0061 protein ydiU - Shigella sonnei (strain Ss046)						

Mascot protein view

Protein View

Match to: FML_HUMAN Score: 112 Expect: 1.7e-06
Probable transcription factor FML (Tripartite motif-containing protein 19)

Nominal mass (M_r): 97455; Calculated pI value: 5.88

NCBI BLAST search of [FML_HUMAN](#) against nr
Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Homo sapiens](#)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Number of mass values searched: 18

Number of mass values matched: 15

Sequence Coverage: 22%

Matched peptides shown in **Bold Red**

```

1 MEAPARSPR PQQDPARQGE PTMPPPETPS EGRQSPSPSPS PTERAPASEE
51 EFQPLRCQQC QAEAKCPKLL PCLHTLCSCG LEASGMQCPI CQAFWPLGAD
101 TPAIDNVFFE SLQRRLSVYR QIVDAQAVCT RCKESADFWC FECEQLLCAR
151 CFEAHQWFLK HEARPLAELR NQSVREFLDG TRKTNINFCS NFNHRTPLT
201 SIYCRGCSKF LCCSCALLDS SHSELKCDIS AEIQQRQEEL DAMTQALQEQ
251 DSAFGAVHAQ MHAAVGQLGR ARAETEELIR ERVQVVVAHV RAQERELLEA
301 VDARYQRDYE EMASRLGRLD AVLQIRITGS ALVQRMKCYA SDQEVLDMHG
351 FLRQALCRLR QEEPGSLQAA VRTDGFDEFK VRLQDLSSCI TQGDAAVSK
401 KASPEAASTP RDPIDVDLPE EAERVQAQVQ ALGLAEAQPM AVVQSVFGAH
451 FVFVYAFSIK GPSYGEDVSN TTTAQKRKCS QTQCPRKVIK MESEEGKEAR
501 LARSSEQPR PSTSKAVSPF HLDGPPSPRS PVIGSEVFLP NSNHVASGAG
551 EAERVVVIS SSEDSDAENS SRELDSSS ESSDLQLEGP STLRVLDENL
601 ADQQAEDRPL VFFDLKIDNE TQKISQLAAV NRESKFRVVI QPEALFSIYS
651 KAVSLEVGLQ HFLSFLSSMR RPILACYKLW GPGLNFFRA LEDINRLWEF
701 QEATSGFLAA LPLIRERVPG ASSFKLKNLA QTYLARIMSE RSAMAAYLAM
751 RDLCLLEVS PGQQLAQHVY PFSSLCQFAS LQPLVQAAYL PRAEARLLAL
801 HNVSMELLS AHRDRQGGG KKYSRVLSLQ TTILPFAQPA FNLQALGTYP
851 EGLLEGFALA RAEGVSTFLA GRGLAERASQ QS

```

Overview

Proteomics

Cancer

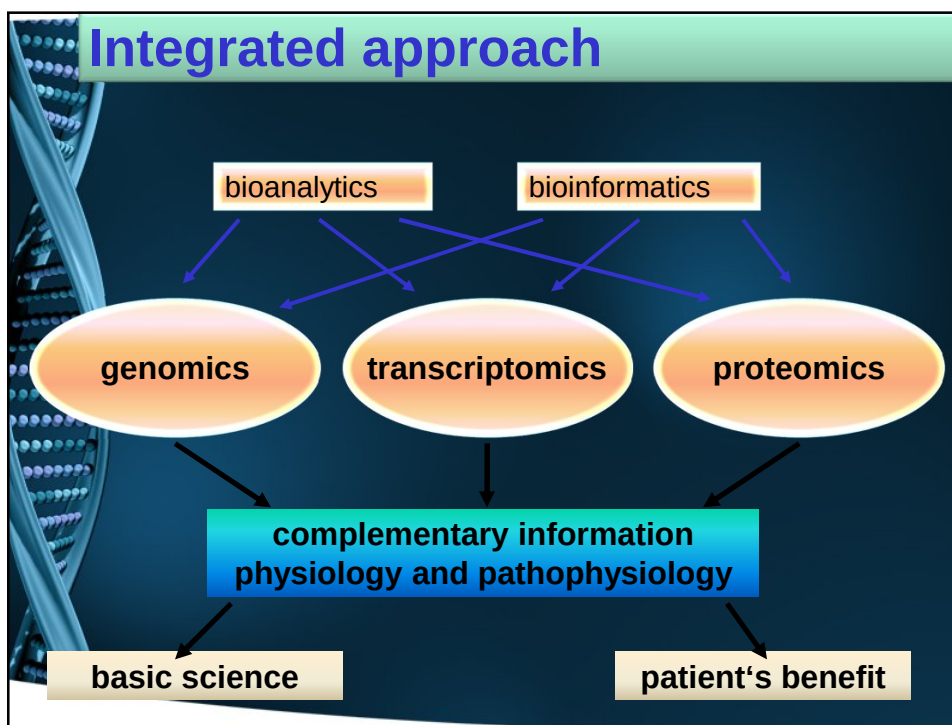
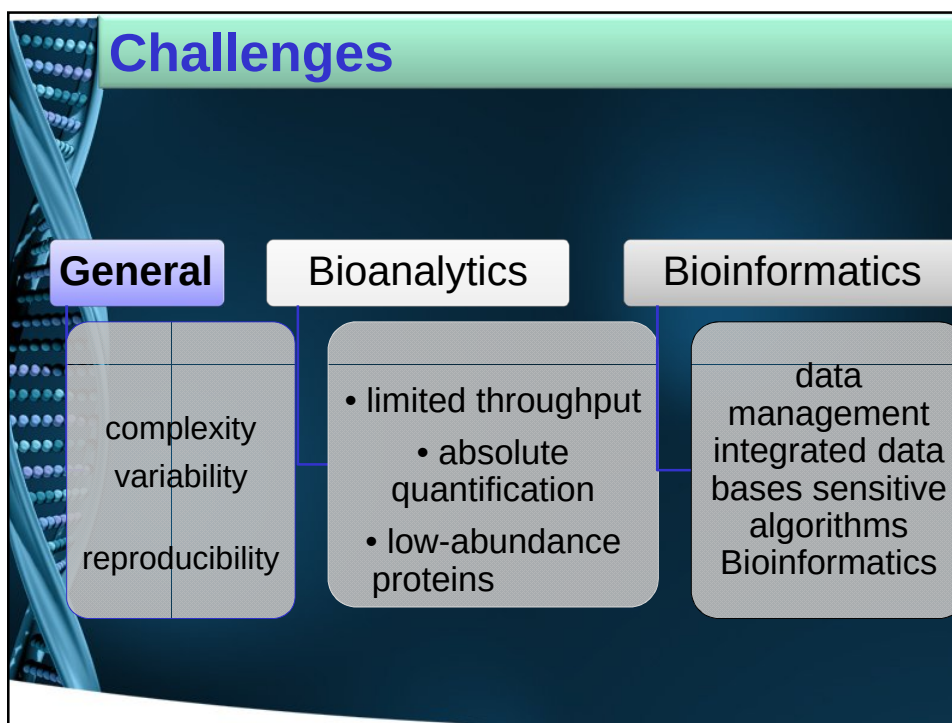
Aims

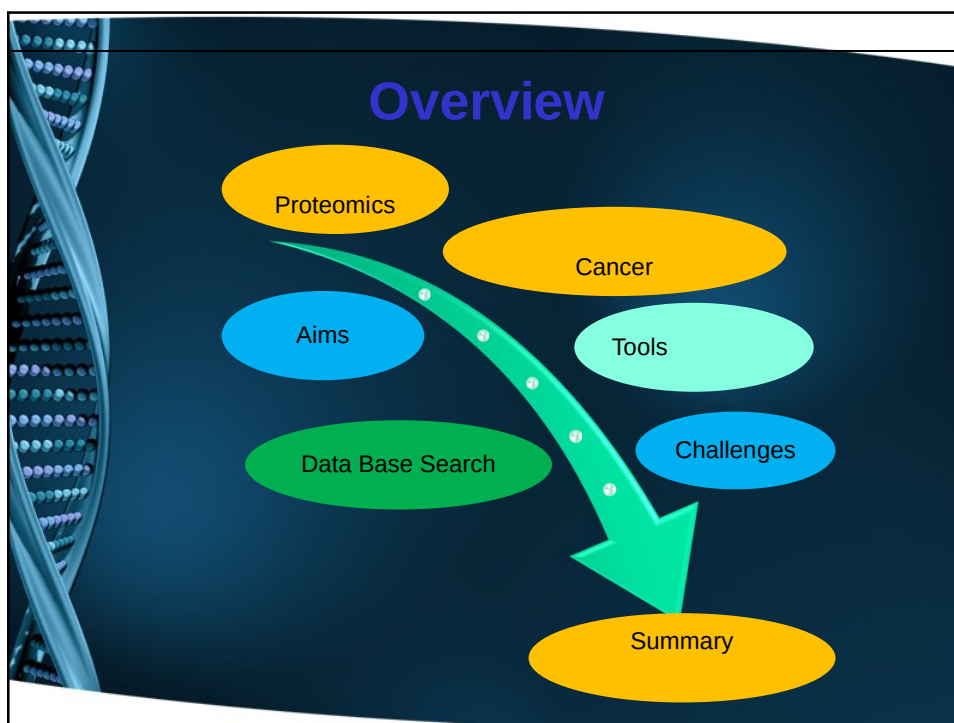
Tools

Data Base Search

Challenges

Summary





Summary

- **bioanalytics** and **bioinformatics** are important in proteomics
- **cancer** is a challenging field: scientists, physicians and patients
- **biomarker** identification and therapeutic aspects
- **workflow**: collection, separation, identification, quantification
- **limitations** (e.g. reproducibility, throughput, data management)
- **integration** of genomics, transcriptomics and proteomics



<http://www.humeny.net>

G. Zhou, H. Li, D. DeCamp, S. Chen, H. Shu, Y. Gong, M. Flaig, J. W. Gillespie, N. Hu, and P. R. Taylor, 2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers, *Molecular & Cellular Proteomics* 1 (2002) 117-123.

http://www.matrixscience.com/search_form_select.html

