DNA-methylation- and autoantibody-biomarker development strategies for minimal invasive diagnostics

Andreas Weinhäusel, Istvan Gyurjan, Johana Luna, Regina Linhart, Manuela Hofner, Gordana Wozniak-Knopp[1], Florian Rüker[1], Friedrich Längle[2], Johann Hofbauer[3], Nina Pecha[4], Christian Singer[4], Robert Zeillinger[4], Andrea Gsur[5], Christa Noehammer, Johannes Söllner[6]

[1] Life Sciences Univ, Vienna, [2] LKH Wr Neustadt, Surgery, Wiener Neustadt, [3] LKH Wr Neustadt, Urology, Wiener Neustadt, [4] Medical Univ Vienna, Senology, Vienna,[5] Medical Univ Vienna, Inst, for Cancer Research, Vienna, Austria, [6] emergentec biodevelopment GmbH, Vienna, Austria,

& Molecular Diagnostics, Health and Environment Department , AIT- Austrian Institute of Technology GmbH, Vienna,

Translational Medicine-2014
3-5 November, 2014, Las Vegas

OMICS Group
Accelerating Scientific Discovery

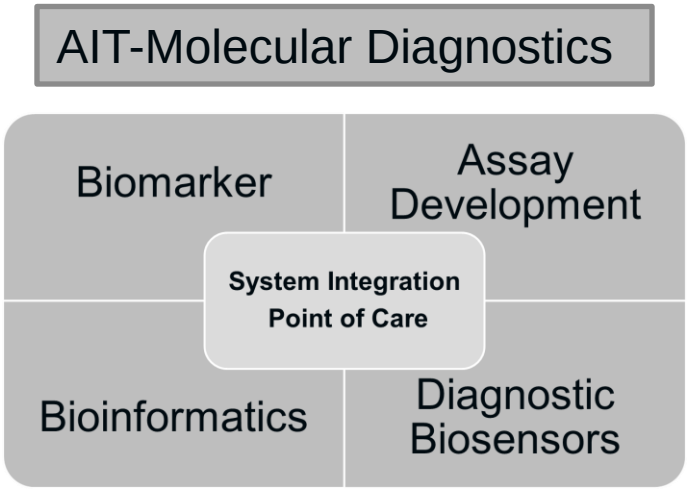
3rd International Conference on
Translational Medicine
November 03-05, 2014 Las Vegas, USA

AIT Austrian Institute of Technology



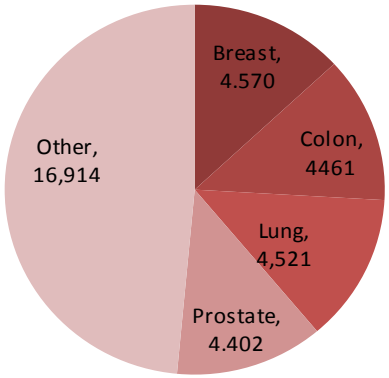
2 daughters

5 academic departments



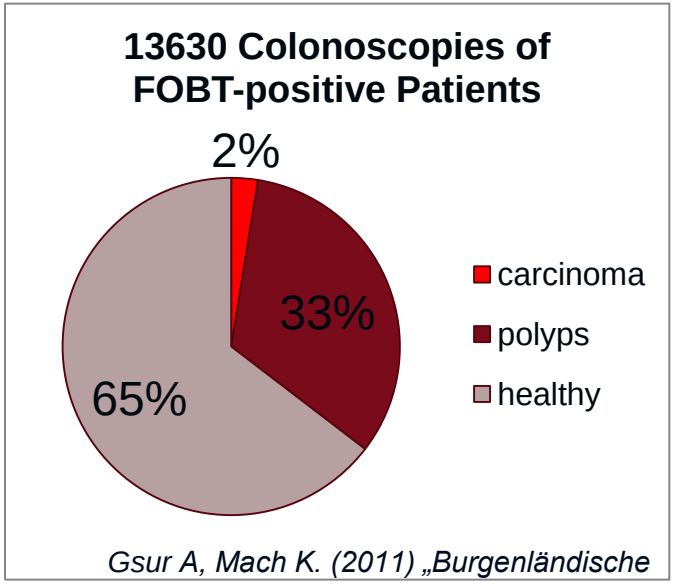
AIM: improving cancer diagnostics

- ***„the big 4“ ➔*** *breast, colon, lung, prostate cancer*
  - incidence *EU25* 1.2mio/y (2006)
  - incidence AUSTRIA 18000/y (2008)
    = 4x13% = 52%



- low therapeutic treatment success
  - **early diagnosis improves survival**
  - „personalised medicine“

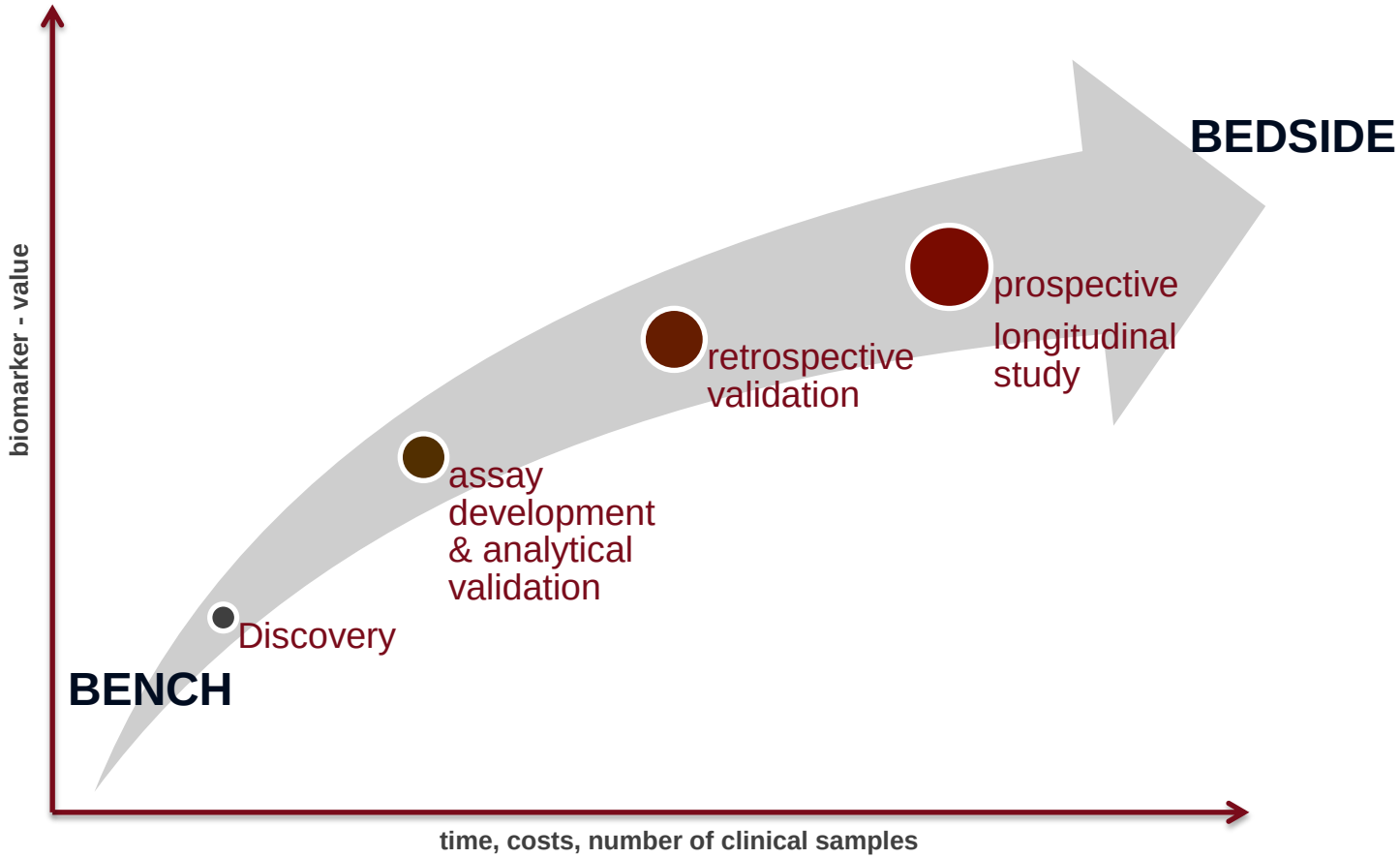
- classical diagnostic methods are „inefficient“
  - Breast: Mammography (60€), MRT (400€), US (60€): ≈80% can be detected
  - **Colon: FOBT**
  - Lung: chest X-ray (78.3% sensitivity)
  - Prostate: PSA test – AUC 0.66



Gsur A, Mach K. (2011) „Burgenländische Dickdarmkrebs –Initiative“

➔ Diagnostic Biomarkers - only a few since recent years available

The biomarker development pipeline



The **4 main phases of biomarker development**

- to move from one phase to another it needs to overcome **multiple challenges at different levels**.
- **the phases are not always distinct from each other and the pipeline is not always linear. Each phase can involve multiple studies performed at various time points throughout the biomarker's developmental lifecycle.**

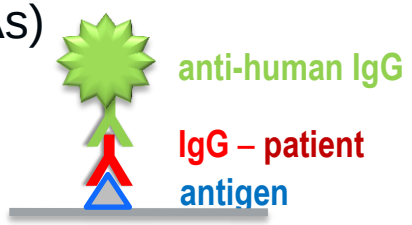
Scheme according to: Pavlou, M.P., Diamandis, E.P., and Blasutig, I.M. (2013). *The Long Journey of Cancer Biomarkers from the Bench to the Clinic.* Clinical Chemistry *59*, 147–157.

Cancer „DIAGNOSTIC“ Biomarkers in clinical use

- **RNA** *PCA3- expression in urine – prostate cancer*
- **DNA Methylation** - *testing biopsies for reconfirmation & also for serum-testing*
  - SEPT9 – Colon (*Epigenomics)*
  - SHOX2 – Lung (*Epigenomics)*
  - GSTP1 - prostate (CLIA laboratory testing- e.g http://mdxhealth.com)
- **PROTEIN** (serum assays)
  - **protein -“serum”- cancer biomarkers**
    - **9** protein cancer biomarkers that have been approved by the FDA for clinical use
    - only PSA as a DIAGNOSTIC BM (AUC 0,66)
  - **Autoantibody** based lung cancer test *EarlyCDT®-Lung test* (Oncimmune):
    - seven antigens (p53, NY-ESO-1, CAGE, GBU4-5, SOX2, HuD, and MAGE A4)
    - sensitivity of 41% with a specificity of 91%;

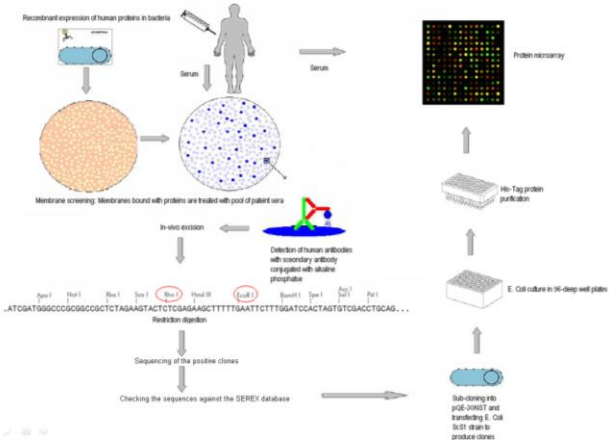
Autoantibody based serum diagnostics

- Autoantibodies in serum against tumor-associated antigens (TAAs) have been found in asymptomatic stage of cancer
- ***early / non invasive / simple == best suited for diagnosis***

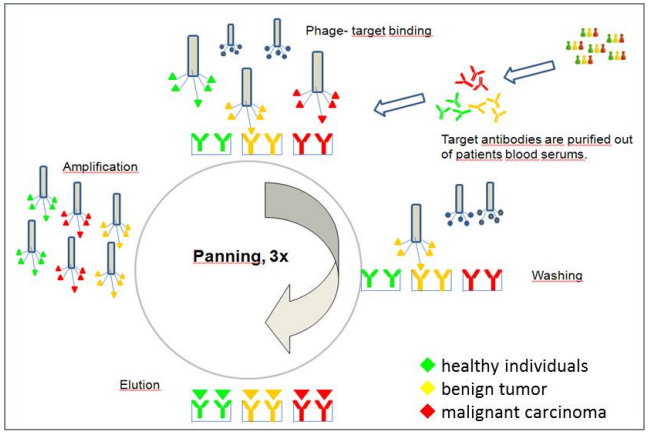


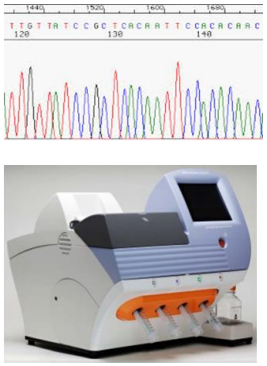
DISCOVERY

A) SEREX



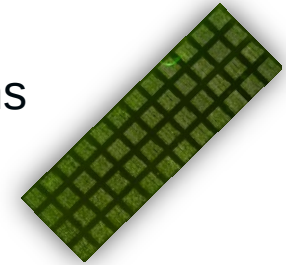
B) phage display based discovery





C) proteins form human clone libraries

- ***PROTEIN micro-arrays*** enable discovery of candidates antigens with low sample (10-30µl) volumes

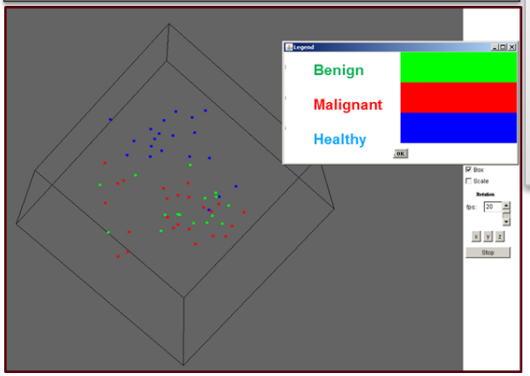


➔ **use micro-arrays to discover & refine marker selection**

Breast Cancer: TAA discovery

- **1.) DISCOVERY** *using human fetal brain*–38k macromembranes & SEREX dervied clones ➔ test on **targeted protein µ-arrays**

BrCa- 642 *Macromembrane* clones @ microarray



Syed, P (2012). In silico design and performance of peptide microarrays for breast cancer tumour-auto-antibody testing. Journal of Molecular Biochemistry; Vol 1, No 2 (2012).

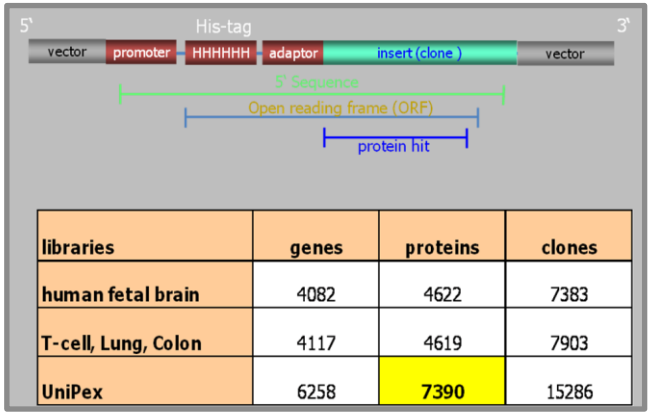
BrCa- Serex results ➔ 194 clones

107 genes
21 already known in SEREX DB
7 thereof described for BrCa/OvCa
22 clones represent potential „novel fusion genes“

Syed, P., etal (2012). Evaluation of auto-antibody serum biomarkers for breast cancer screening and in silico analysis of sero-reactive proteins. Journal of Molecular Biochemistry; Vol 1, No 2 (2012).

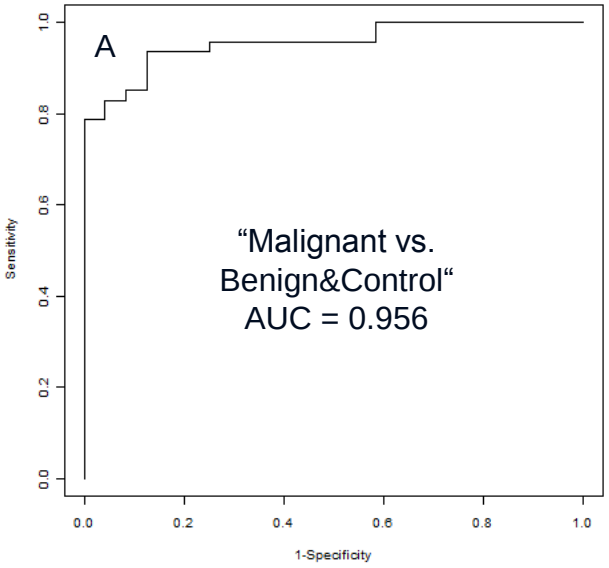
➔ AUC ≈ 0.75

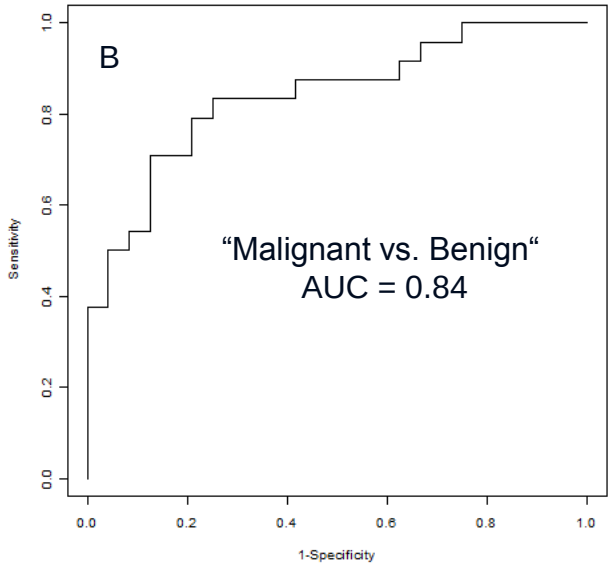
- **2.) 16k protein micro-array**
  - 15286 „UNIPEX-clones“ (his-tag) purified & duplicates printed – 32000 spots/slide
  - purified IgG & screen:
    72 malignant, 60 healthy, 48 benign
  - processing of 6x36 = 216 -*16k arrays*

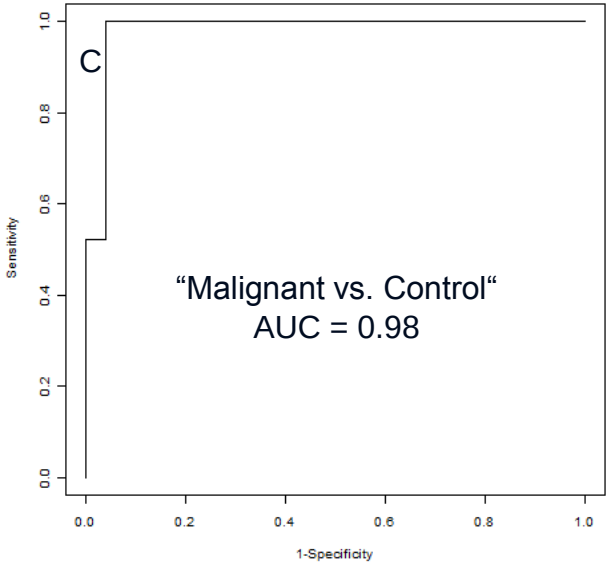


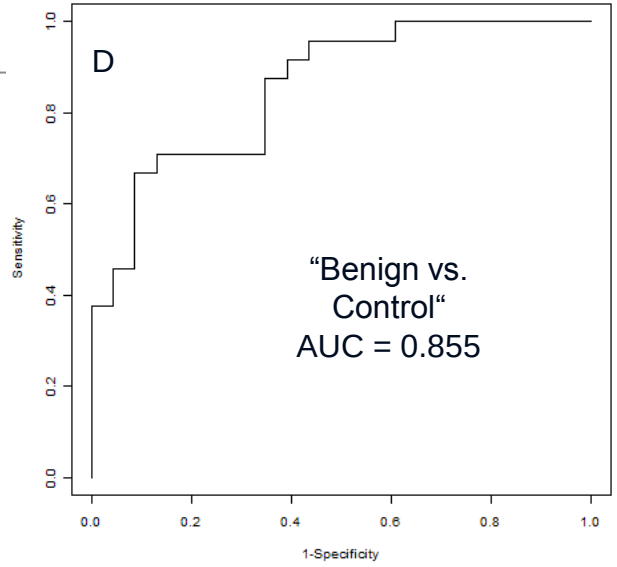
libraries	genes	proteins	clones
human fetal brain	4082	4622	7383
T-cell, Lung, Colon	4117	4619	7903
UniPex	6258	**7390**	15286

16k BrCa TAA analyses







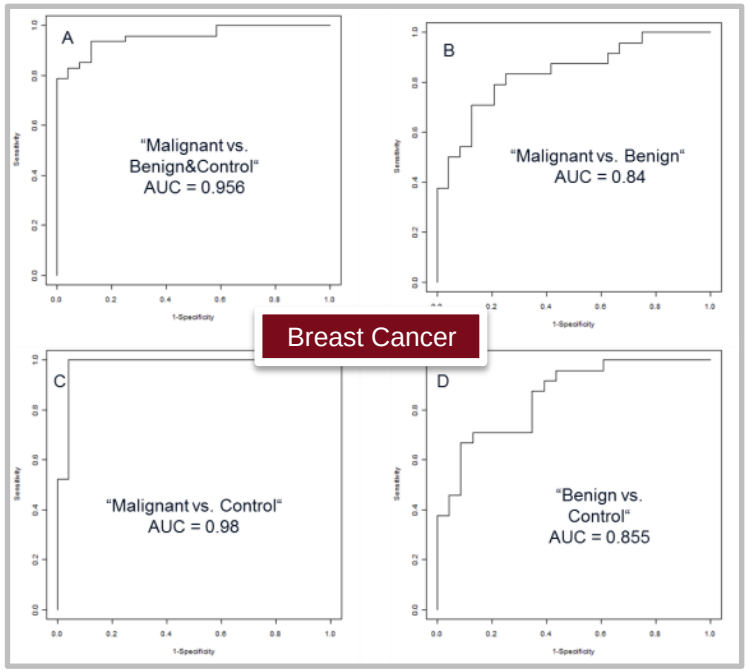


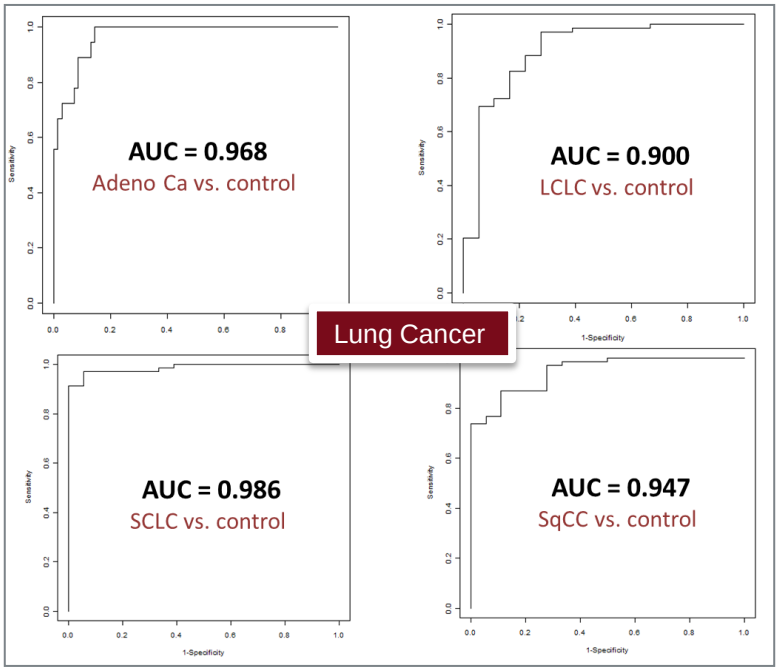
IgG, 0.3 mg/mL - from 30 µL of serum / chip

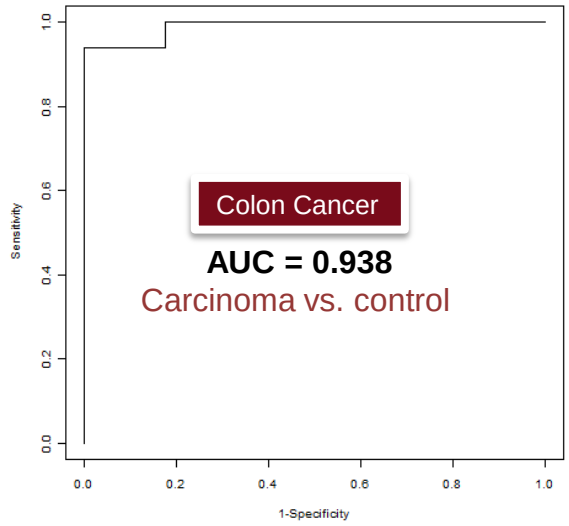
➢ **Cross-validation receiver operating characteristic (ROC) curves of antigen-classifiers derived from the 16k protein-chip screening for different sets of samples (Bayesian Compound Covariate Predictor, p<0.001)**

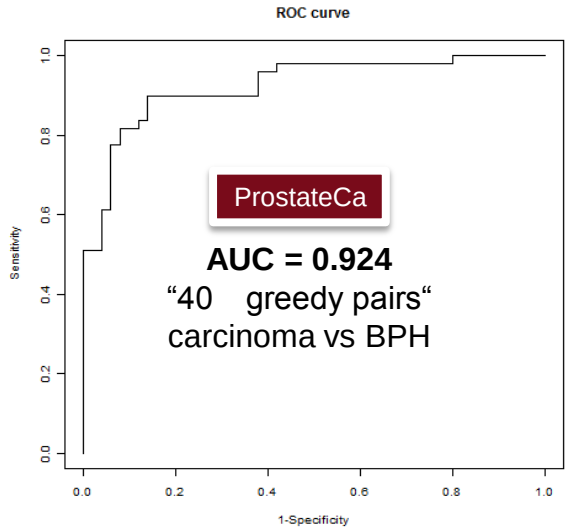
A: AUC 0.956 for malignant vs. "benign&control" samples (100 classifiers).
B: AUC 0.84 for malignant vs. benign samples (94 classifiers).
C: ROC curve for malignant vs. control samples with an AUC of 0.98 (108 classifiers).
D: AUC of 0.855 was obtained from the classification between benign vs. control samples (46 classifiers).

BIG-4 cancer entities TAA candidate markers



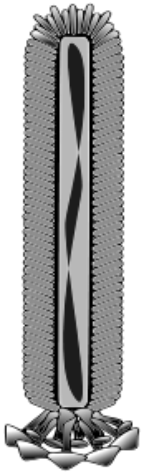


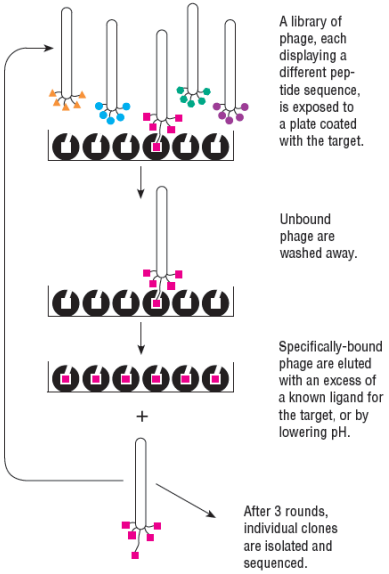


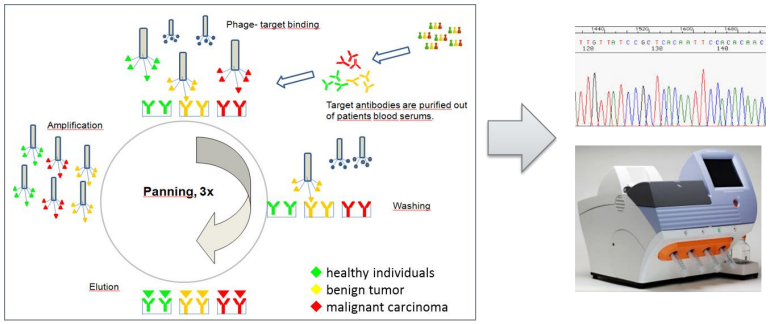


Phage display based discovery (BreastCa)

- ~50 serum-samples per group **Control, Malignant, Benign** - 100µl of each patient pooled
- random 12aa M13-phage library (from NEB) – subtractive panning



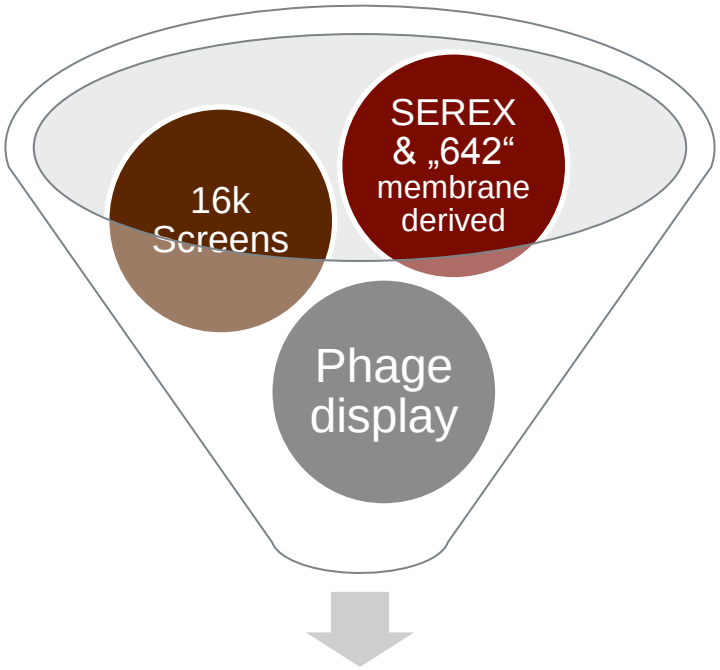




subtractive panning: non-binders from patient group "MAL" are enriched on "BEN/CTRL" column

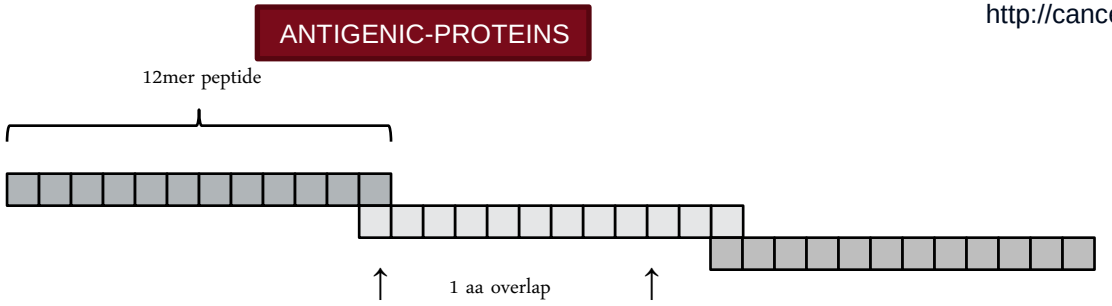
- ***ProteinA & ProteinG - IG** isolations* >> 106000 reads ➔ **3325** different peptides
- repeat - using **MelonGel for IgG** selection >> 67126 reads ➔ **8976** different peptides

179k peptide array design



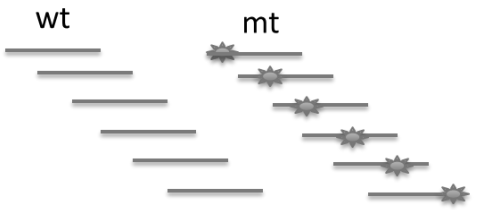
Cancer-type	Selected proteins	Corresponding peptides
Breast	656	51563
Colon	536	43831
Prostate	457	39824
Lung	898	68016
Cosmic classic	93	8685
Serex_derived	61	3446
Unique	2090	165570
Phage-display		1076
Cosmic mutations	41 proteins/496 mutations	5950
		172596 peptides

PEPTIDE DESIGN
172.596



COSMIC
Catalogue of somatic mutations in cancer

http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/

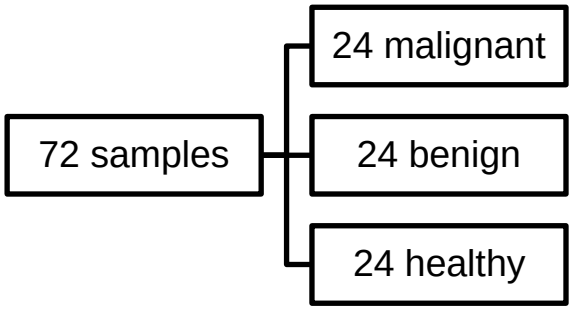


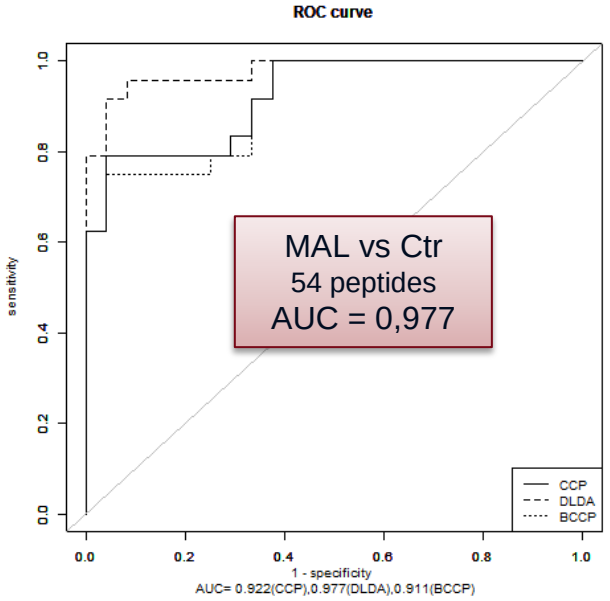
Somatic mutations

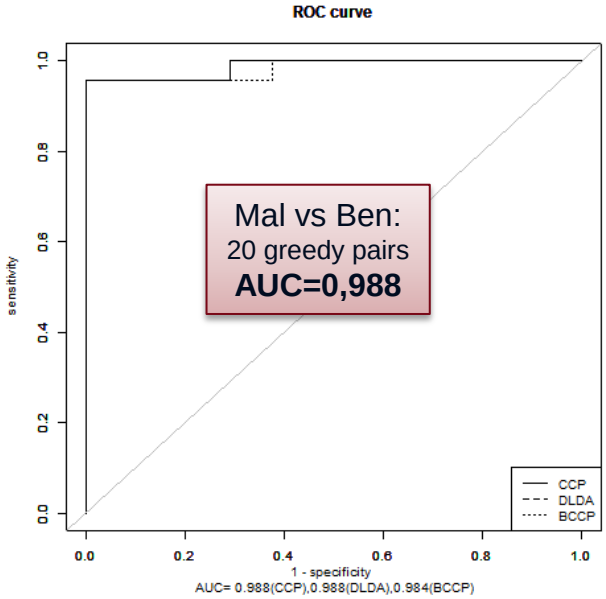
BrCa-Serum IgG -179k peptide array analyses

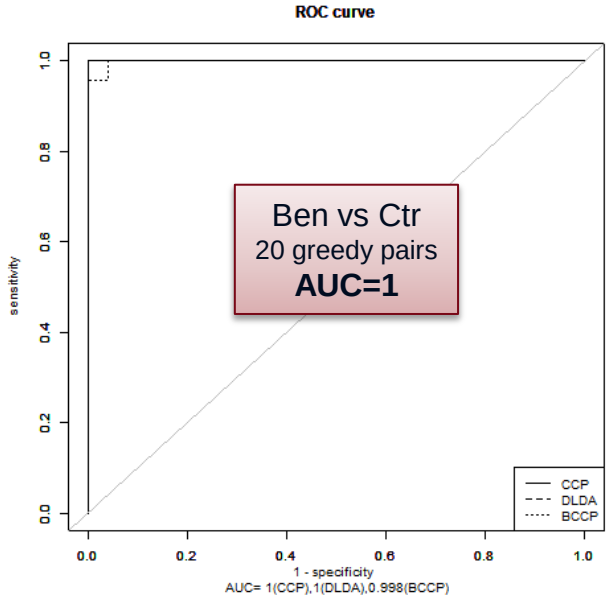
PROCESSING of 12 samples / array: 179K peptides











Mapping peptides to *biology*- KEGG pathway analyses

AUSTRIAN INSTITUTE OF TECHNOLOGY

Term	Count	PValue	Fold	Bonferroni	Benjamini
hsa05200:Pathways in cancer	151	5.7714E-019	1.90	1.0792577E-016	1.0792577E-016
hsa05213:Endometrial cancer	41	1.2172E-015	3.26	2.2837288E-013	1.1413093E-013
hsa04510:Focal adhesion	100	3.2459E-015	2.06	6.0207395E-013	2.0072832E-013
hsa04010:MAPK signaling pathway	122	6.8339E-015	1.89	1.2871926E-012	3.2174263E-013
hsa04722:Neurotrophin signaling pathway	70	2.8650E-014	2.33	5.3563820E-012	1.0712542E-012
hsa04144:Endocytosis	90	3.3834E-013	2.02	6.3280048E-011	1.0546675E-011
hsa05210:Colorectal cancer	52	7.2972E-013	2.56	1.3646317E-010	1.9494739E-011
hsa04012:ErbB signaling pathway	53	1.0432E-012	2.52	1.9507196E-010	2.4383939E-011
hsa04370:VEGF signaling pathway	48	1.1332E-012	2.64	2.1190927E-010	2.3545499E-011
hsa04520:Adherens junction	48	4.5172E-012	2.58	8.4470975E-010	8.4470986E-011
hsa05223:Non-small cell lung cancer	38	5.7094E-012	2.91	1.0676432E-009	9.7058472E-011
hsa04664:Fc epsilon RI signaling pathway	48	8.7498E-012	2.54	1.6362086E-009	1.3635071E-010
hsa04540:Gap junction	51	6.9965E-011	2.37	1.3083399E-008	0.000000001
hsa05220:Chronic myeloid leukemia	45	1.4324E-010	2.48	2.6785937E-008	1.9132813E-009
hsa05212:Pancreatic cancer	43	4.8576E-010	2.47	9.0837903E-008	6.0558605E-009
hsa04910:Insulin signaling pathway	66	7.5530E-010	2.02	1.4124070E-007	8.8275446E-009
hsa05221:Acute myeloid leukemia	37	7.8081E-010	2.64	0.000000146	8.5889219E-009
hsa05215:Prostate cancer	49	1.1224E-009	2.27	2.0989738E-007	1.1660967E-008
hsa04730:Long-term depression	41	1.6403E-009	2.45	3.0674176E-007	1.6144306E-008
hsa04810:Regulation of actin cytoskeleton	91	0.000000004	1.75	7.5211631E-007	3.7605829E-008
hsa04666:Fc gamma R-mediated phagocytosis	50	5.5360E-009	2.17	1.0352268E-006	4.9296536E-008
hsa05130:Pathogenic Escherichia coli infection	35	1.0184E-008	2.54	1.9043671E-006	8.6562218E-008
hsa04662:B cell receptor signaling pathway	42	1.1084E-008	2.31	2.0727626E-006	9.0120201E-008
hsa04660:T cell receptor signaling pathway	54	1.2619E-008	2.07	2.3596877E-006	9.8320431E-008
hsa05214:Glioma	37	1.8440E-008	2.43	3.4482019E-006	1.3792830E-007
hsa05216:Thyroid cancer	22	5.4328E-008	3.13	1.0159360E-005	3.9074654E-007
hsa00051:Fructose and mannose metabolism	24	9.4325E-008	2.92	1.7638597E-005	6.5328691E-007
hsa03040:Spliceosome	58	0.000000137	1.90	2.5626877E-005	9.1525692E-007
hsa04930:Type II diabetes mellitus	29	2.1670E-007	2.55	4.0522486E-005	1.3973544E-006
hsa04914:Progesterone-mediated oocyte maturation	43	4.9465E-007	2.07	9.2495500E-005	3.0833212E-006
hsa04912:GnRH signaling pathway	47	6.1215E-007	1.98	0.0001144649	3.6926222E-006
hsa04720:Long-term potentiation	36	9.2436E-007	2.19	0.0001728408	5.4017285E-006
hsa05222:Small cell lung cancer	41	2.1058E-006	2.02	0.0003937043	1.1932712E-005
hsa04960:Aldosterone-regulated sodium reabsorption	25	2.5188E-006	2.52	0.0004708992	1.3853143E-005
hsa04062:Chemokine signaling pathway	74	3.1038E-006	1.63	0.0005802406	0.000016583

➔ all 170k unique peptides have been mapped back to 21657 distinct proteins in UNIPROT matching **4139 genes** which served as background for pathway enrichment analysis and are obviously cancer related; :

Class distinction via genes/proteins *mapped back* from peptides

CONTRAST	distinct gene-symbols	distinct proteins
Diseased vs (Benign & Healthy)	473	1149
Benign vs (Diseased & Healthy)	81	171
Healthy vs (Diseased & Benign)	27	60
Diseased vs Benign	2217	6766
Diseased vs Healthy	2280	7394
Benign vs Healthy	480	1113

Pathway analyses - Diseased vs(Benign and Healthy)

- based on 473 distinct genes

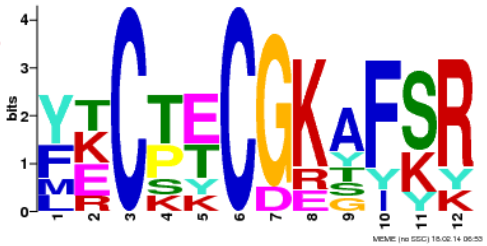
AUSTRIAN INSTITUTE OF TECHNOLOGY

Term	Count	PValue	Fold	Bonferroni	Benjamini
phosphoprotein	353	7.648821E-016	1.29	3.19411E-013	3.19411E-013
alternative splicing	270	9.711410E-008	1.25	3.99131E-005	1.99567E-005
splice variant	272	3.941179E-008	1.26	8.44953E-005	8.44953E-005
acetylation	160	8.152820E-007	1.39	0.0003350249	0.0001116874
polymorphism	336	0.00000525	1.15	0.0021554213	0.0005392914
coiled coil	102	1.971754E-005	1.47	0.0080712398	0.0016194849
sequence variant	349	3.223855E-006	1.14	0.0068881229	0.0034500127
DNA replication, recombination, and repair	9	0.0001954446	4.80	0.004096325	0.004096325
GO:0007017~microtubule-based process	28	0.00000719	2.50	0.0154884027	0.0051896872
IPR002453:Beta tubulin	8	6.018783E-006	7.96	0.0055818487	0.0055818487
GO:0033043~regulation of organelle organization	24	6.620922E-006	2.75	0.0142712558	0.0071612698
IPR013838:Beta tubulin, autoregulation binding site	8	1.813977E-005	7.16	0.0167286315	0.0083995923
GO:0022402~cell cycle process	41	5.800941E-006	2.07	0.0125149079	0.0125149079
GO:0022403~cell cycle phase	30	2.573152E-005	2.26	0.0543321216	0.0138688912
GO:0019899~enzyme binding	44	4.304963E-005	1.86	0.0251639057	0.0251639057
microtubule	21	0.0003844832	2.33	0.1461954466	0.0259982337
GO:0007010~cytoskeleton organization	37	7.943965E-005	1.94	0.1584167503	0.0339059295
GO:0032318~regulation of Ras GTPase activity	13	9.544284E-005	3.62	0.1871561772	0.0339464852
GO:0000279~M phase	24	0.0001125073	2.34	0.2167216411	0.0342934855
GO:0046578~regulation of Ras protein signal transduction	21	0.0001590614	2.47	0.2920272824	0.0422502104
GO:0007049~cell cycle	46	0.0002664809	1.70	0.4393212101	0.0435322801
GO:0000278~mitotic cell cycle	27	0.000259364	2.10	0.4305886647	0.0458451935
GO:0051056~regulation of small GTPase mediated signal transduction	24	0.000221065	2.25	0.3812073781	0.0468647912
GO:0000087~M phase of mitotic cell cycle	18	0.0002553345	2.62	0.4255843478	0.0491511651
GO:0000226~microtubule cytoskeleton organization	16	0.000209262	2.88	0.3651428932	0.0492307888

Motif enrichment search:

enriched in **Malignant** samples:

E-value 3.4e-015



Zn-finger proteins

E-value 1.7e-005



diverse proteins (e.g. MUC5), redundant sequences

E-value 2.6e-003



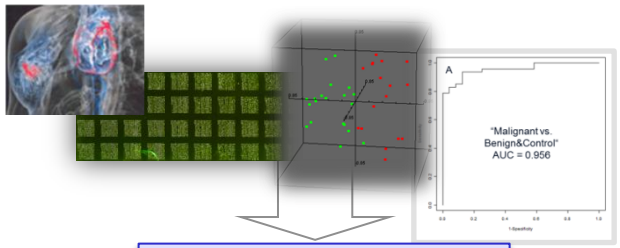
Zn-finger proteins

PROTEIN

Peptides

confirmation & validation

Breast Cancer



16k Protein-chip Classifiers,
Serex
COSMIC
Phage-Display Peptides



Fluoreszently anti-human IgG detection antibody

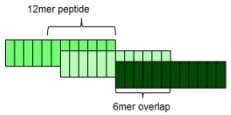
IgG – Patient vs. **Control**

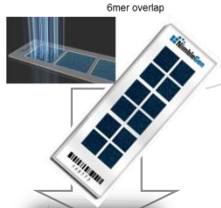
immobilised Antigen

healthy | patient

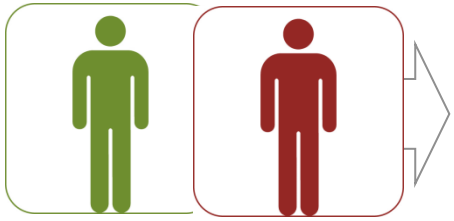


Discovery Set

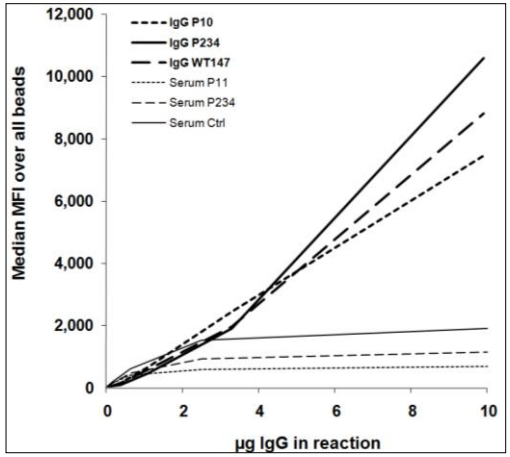




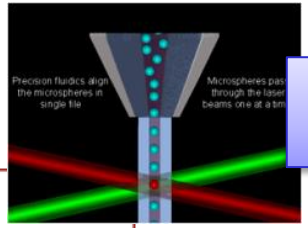
custom 179.000 peptide array



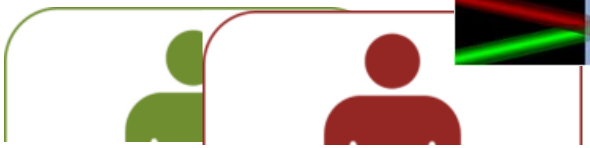
Bioinformatics



Discovery
& Training
&
Validation-SET



200⁺ plex Luminex-bead array



2x 100-plex Luminex-ASSAYS – using biotinylated peptides on avidin coated magnetic beads … almost ready

Antigenic / Immunomics -profile based Biomarkers

- microarrays improve discovery
  - 30µl serum / plasma are sufficient… *SALIVA works also*
  - IgG preparation from serum/plasma➔ wide range of linearity
  - *microarrays* & bioinformatics - well established & suited for discovery
  - *bead arrays* established for validation studies
- <u>from proteins to peptides</u>
  - direct peptide based screening - using phage display & NGS
  - tiling design and high density peptide arrays are efficient for defining chemically synthesizable peptides
  - specifying antigenic epitopes
  - peptide arrays improve classification success
  - peptide array data are biologically meaningful

….The technologies are ready.

….The technologies are ready …
& cooperations are welcome….

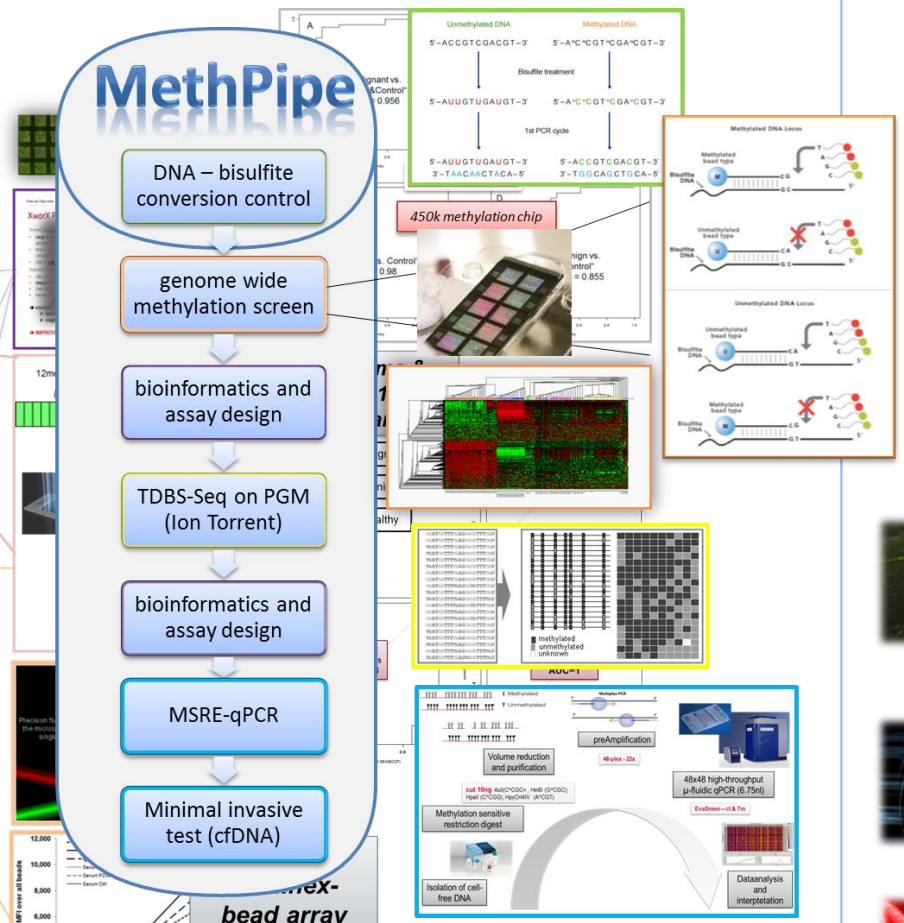
AUSTRIAN INSTITUTE OF TECHNOLOGY

PepPipe

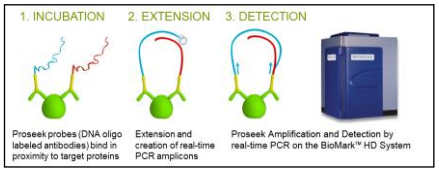
- targets of interest
- bioinformatics
- peptidome / custom peptide array
- bioinformatics
- bead array
- bioinformatics
- ELISA

MethPipe

- DNA – bisulfite conversion control
- genome wide methylation screen
- bioinformatics and assay design
- TDBS-Seq on PGM (Ion Torrent)
- bioinformatics and assay design
- MSRE-qPCR
- Minimal invasive test (cfDNA)



Olink's Proseek panel assays



analyse **92 markers** from 1µl of serum



Oncology



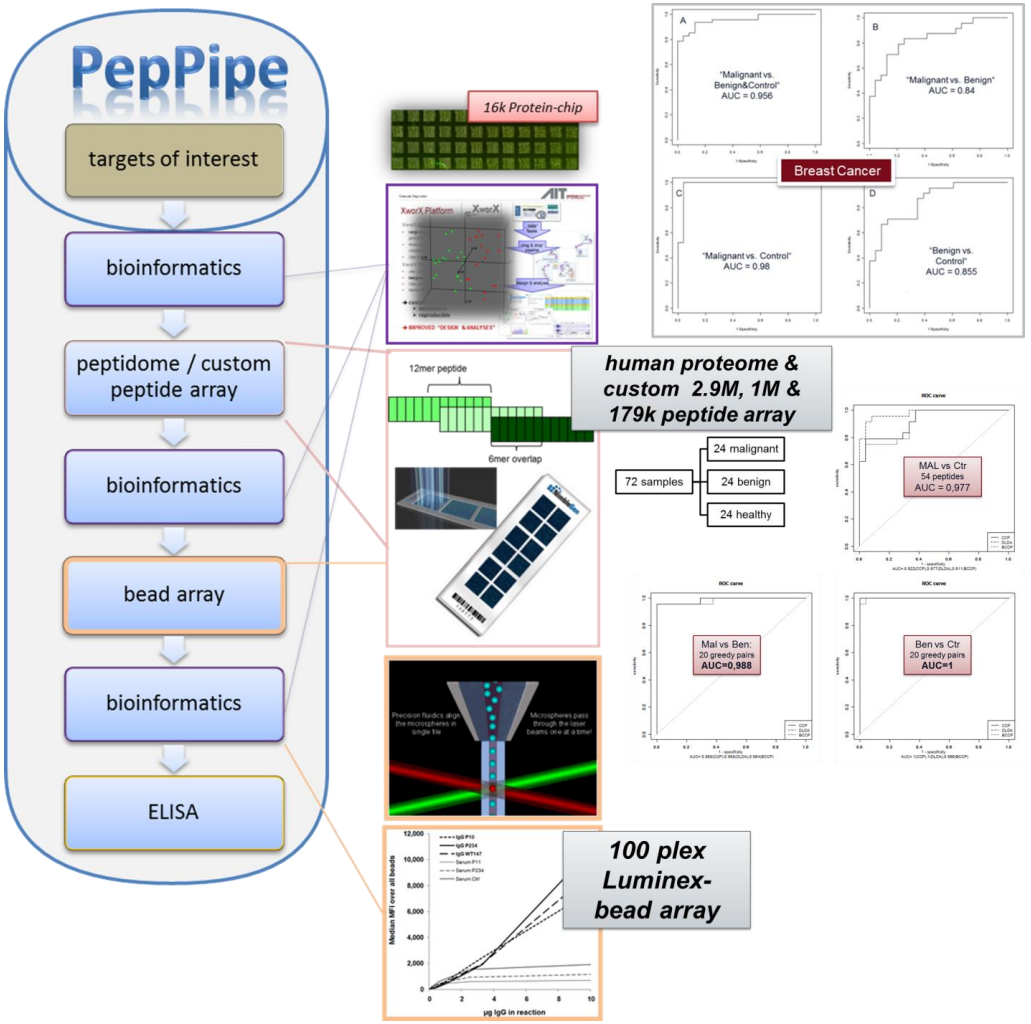
Cardio-Vascular / CVD



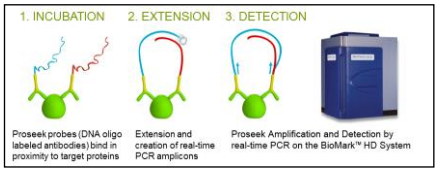
Inflammation

….The technologies are ready …
& cooperations are welcome….

AUSTRIAN INSTITUTE OF TECHNOLOGY



Olink's Proseek panel assays



analyse **92 markers** from 1µl of serum



Oncology



Cardio-Vascular / CVD



Inflammation

Acknowledgement



AIT's –"*AUTOANTIBODY*" team: Istvan Gyurjan, Johana Luna, Steffanie Brezina, Regina Soldo, Tina Malovits, Regina Linhard, Roman Kreuzhuber, Lisa Milchrahm, Olga Peinando, Sandra Rosskopf, Manuela Hofner, Christa Noehammer, Roni Kulovics, Oskar Morgenbesser, Gabriel Beikircher, Michael Stierschneider, Klemens Vierlinger

see also poster:

#522. Identification of differentially reactive antigen biomarkers for prostate cancer screening using protein microarrays. **J.A. Luna-Coronell**, M. Gamperl, R. Kulovics, P. Hofer, A. Gsur, J. Hofbauer, F. Längle, A. Weinhäusel

Cooperations:

C. Singer, R. Zeillinger, C. Rappaport, N. Picha, ……**MUW Vienna**

F. Längle, J. Hofbauer, I. Berger, B. Hameed, **LKH Wr. Neustadt**

A. Gsur. & P. Hofer, ICR – **MUW**

G. Leeb & K. Mach, **LKH Oberpullendorf**

C. Jungbauer, **Austrian Red Cross**

J. Söllner, **Emergentec**

Funding & Support:

OESTERREICHISCHE NATIONALBANK

life science krems

tecnet equity

Wiener Wissenschafts-, Forschungs- und Technologiefonds