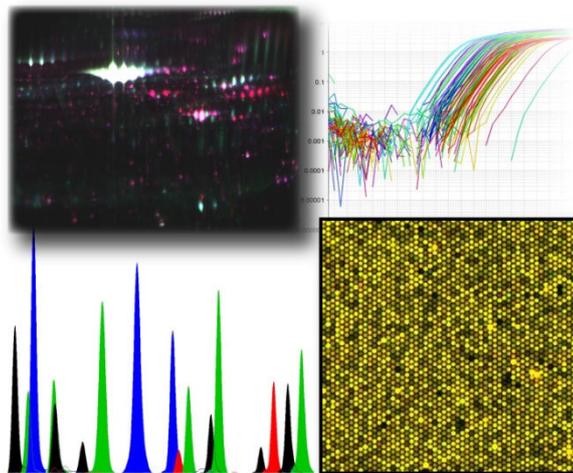


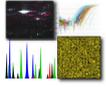
## MACPROGEN

### 3<sup>rd</sup> Genomics and Proteomics Workshop

March 1-6, 2012

Skopje, Republic of Macedonia





# MACPROGEN

3<sup>rd</sup> Genomics and Proteomics Workshop

March 1-6, 2012

Skopje, Republic of Macedonia

***Organized by:***

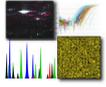
Research Center for Genetic Engineering and Biotechnology,  
Macedonian Academy of Sciences and Arts



***Sponsored by:***

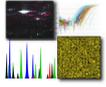
FP7 grant “National Reference Centre for Genomics and Proteomics” (acronym – MACPROGEN), funded by the Seventh Framework Programme, REGPOT-2008-1.





## CONTENTS MANUEL

- **Program**
- **Speakers**
- **Participants**
- **Talks**
- **Workshop, Training sessions**



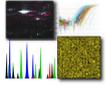
## PROGRAM

### **March 1, 2012 (Thursday)**

- 09.00 – 09.30 Registration
- 09.30 – 12.00 Practical session – Designing and performing an MLPA for DNA detection and Quantification  
Designing and performing an HRM Mutation Scanning Experiment
- 12.00 – 13.00 *Lunch break*
- 13.00 – 16.00 Practical session – Real-time gene expression analysis
- 16.00 – 17.00 Round table discussion (Troubleshooting for array MLPA analysis; Troubleshooting for HRM and gene expression analysis)
- 19.30 *Workshop dinner*

### **March 2, 2012 (Friday)**

- 09.00 – 12.00 Practical session – DNA methylation arrays
- 12.00 – 13.00 *Lunch break*
- 13.00 – 16.00 Practical session – Microarray Data Analysis
- 16.00 – 17.00 Round table discussion (Microarray data analysis and web databases resources)



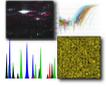
**March 5, 2012 (Monday)**

- 09.00 – 12.00 Practical session - Proteomics  
MALDI-TOF MS and MS/MS analysis of proteins/peptides
- 12.00 – 13.00 *Lunch break*
- 13.00 – 16.00 Practical session - Proteomics (*continues*)  
MASCOT database searches
- 16.00 – 17.00 Round table discussion (Possibilities and limitations of MALDI-TOF technology)

**March 6, 2012 (Tuesday)**

- 09.00 – 12.00 Practical session - Proteomics  
Bacterial identification by MALDI-TOF MS
- 12.00 – 13.00 *Lunch break*
- 13.00 – 13.10 Welcome address - Dijana Plaseska Karanfilska
- 13.10 – 13.55 *Max Chaffanet, France: “Genomic analysis in cancer”*
- 13.55 – 14.40 *Stephane Audebert, France: “Proteomic tools in cancer research”*

Closing



## **SPEAKER LIST**

### **Max Chaffanet**

Institut Paoli-Calmettes, U1068 INSERM,

Center for Cancer Research,

Molecular Oncology Department,

Marseille, France

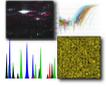
### **Stephane Audebert**

Institut Paoli-Calmettes, U1068 INSERM,

Center for Cancer Research,

Molecular Oncology Department,

Marseille, France



## **PARTICIPANTS**

**Dijana Plaseska-Karanfilska**

**Emilija Sukarova-Stefanovska**

**Katarina Davalieva**

**Katerina Popovska-Jankovik**

**Lyubomira Chakalova**

**Predrag Noveski**

**Sanja Kiprijanovska**

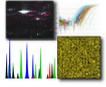
**Ivana Maleva**

**Svetlana Madzunkova**

**Mirko Trpevski**

**Zvezdana Moneva**

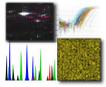
**Gorgi Bozinovski**



## TALKS

### Order:

1. **Max Chaffanet**, France: “Genomic analysis in cancer”
2. **Stephane Audebert**, France: “Proteomic tools in cancer research”



**Inserm**  
CENTRE NATIONAL DE RECHERCHE EN CANCÉROLOGIE  
Centre de Recherche en Cancérologie de Marseille  
U1068 INSERM – Molecular Oncology Department

**Oncogenomic Group**

**Integrated genomic profiling  
in breast cancers**

**Max Chaffanet**

**MACROGENE**

Skopje, March 6, 2012

**Breast Cancer**

From: *L. Liberato, Journal of Science 2007*

**Heterogeneous and complex**

**Mammary carcinogenesis**

Normal Cell

Molecular Changes

Cancer Cell

**normal**

**tumoral**

**Breast cancer**

**Characterization of  
new candidate cancer genes**

**Recurrent molecular alterations**

**Altered genes and mechanisms**

**Frequency, Clinical impact, Classification, New markers**

**How to integrate the current classification and an emerging molecular/cellular classification into a single one to understand breast cancer and deliver the best treatment?**

**Molecular portrait**

**Definition of molecular subtypes**

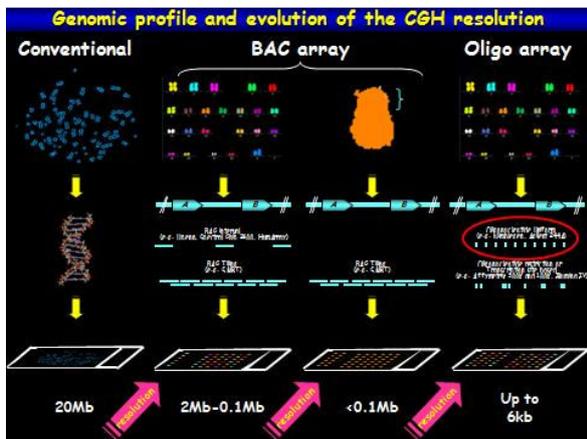
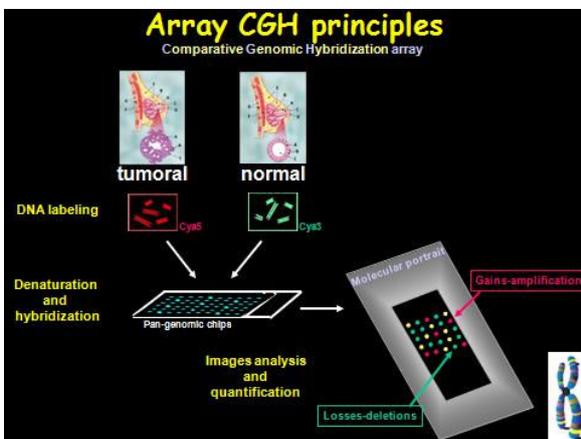
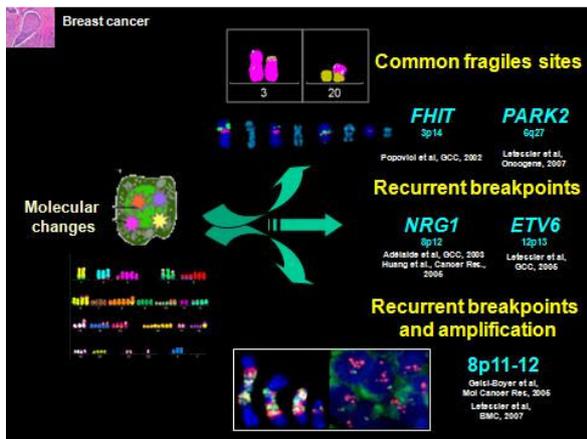
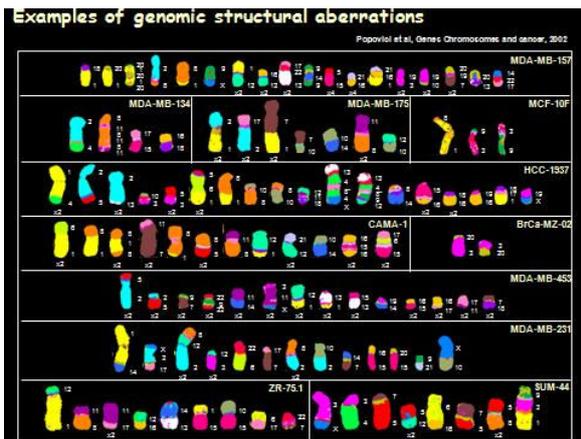
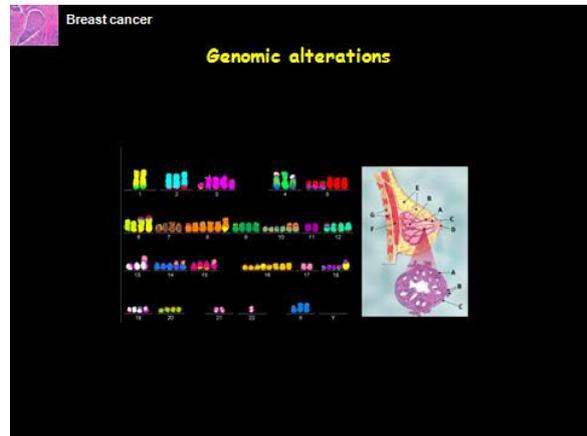
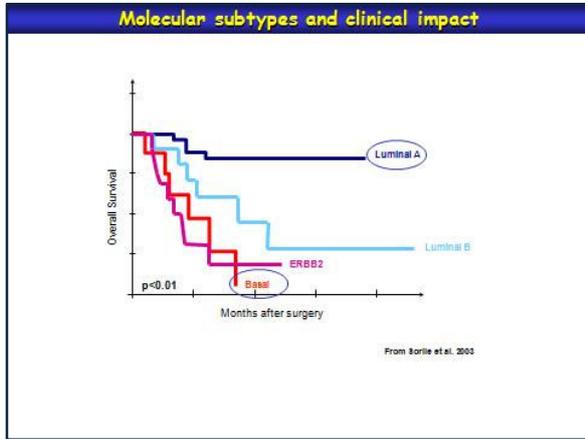
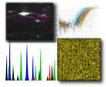
Prolif. G1  
Prolif. Ki67  
Basal  
Immune  
Stroma  
Vascular  
ERBB2  
Luminal

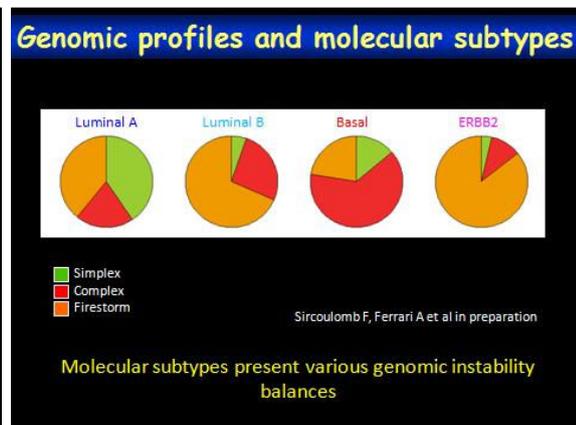
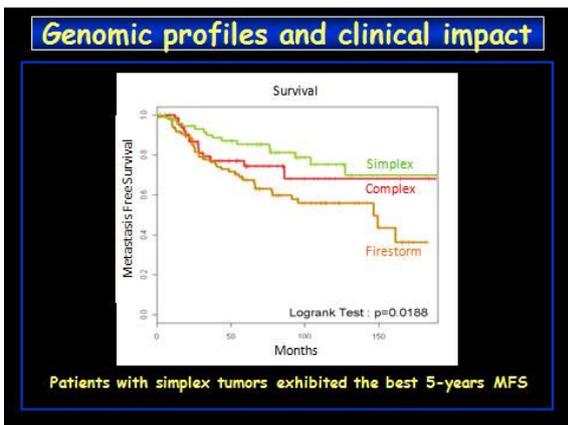
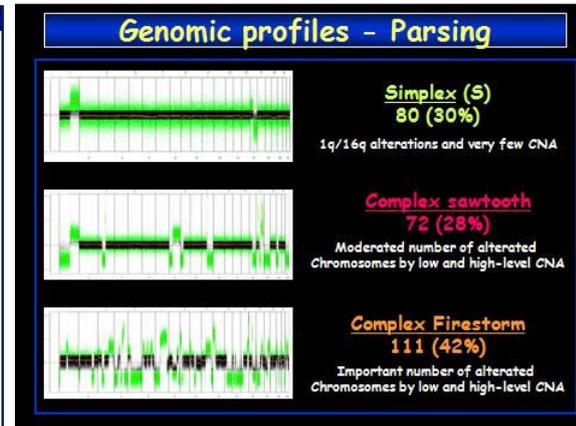
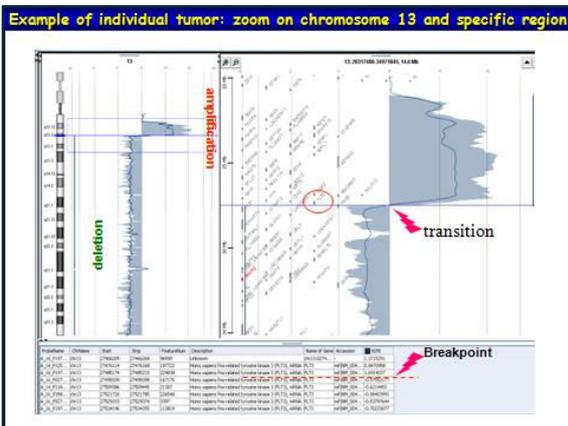
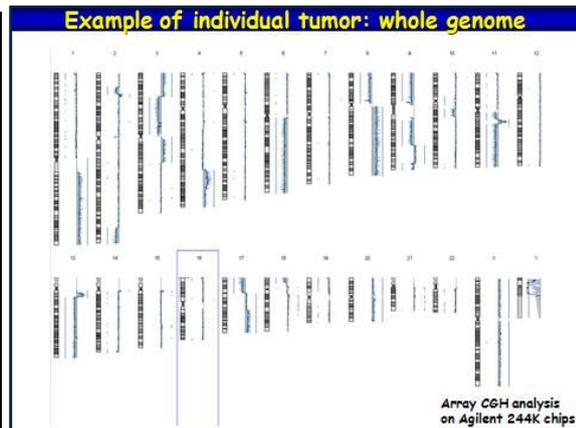
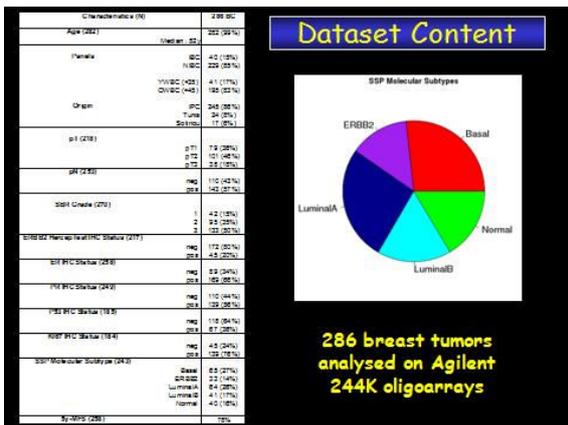
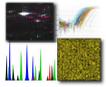
**Hierarchical clustering of 220 breast tumors Affymetrix whole-genome oligochips**

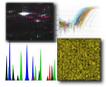
**Subtype (Sorlie et al. 2003)**

- Basal
- Luminal A
- Luminal B
- ERBB2
- Normal
- Autres

Finetti, Cervera, Bartucci et al.





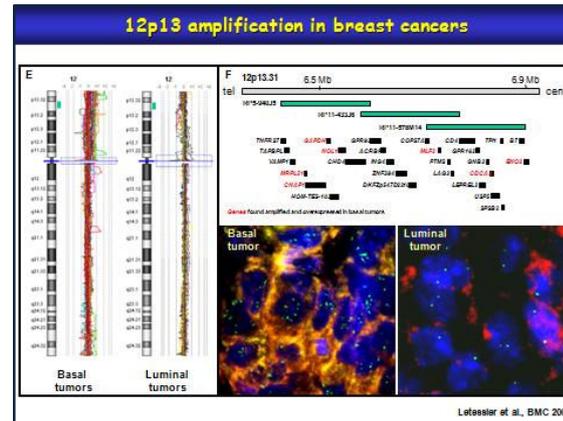
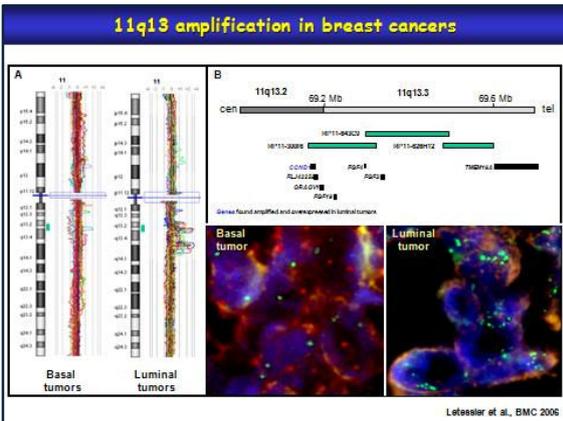
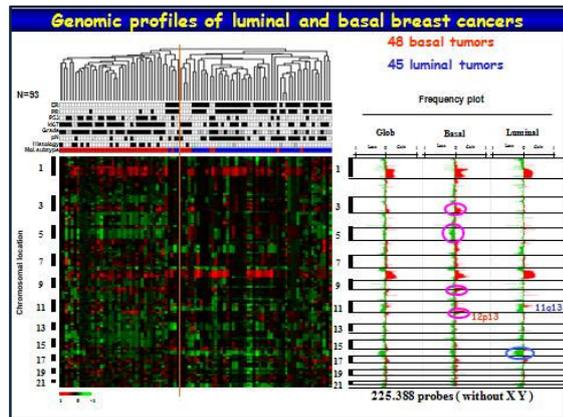


Research Article

### Integrated Profiling of Basal and Luminal Breast Cancers

José Adélaïde,<sup>1</sup> Pascal Finetti,<sup>1</sup> Imahane Bekhouche,<sup>1</sup> Laetitia Repellini,<sup>1</sup> Jeanine Geneix,<sup>1</sup> Fabrice Sircoumb,<sup>1</sup> Emmanuelle Charafe-Jauffret,<sup>1,2,3</sup> Nathalie Cervera,<sup>1</sup> Jérôme Desplans,<sup>4</sup> Daniel Parzy,<sup>5</sup> Eric Schoenmakers,<sup>6</sup> Patrice Viers,<sup>7</sup> Jocelyne Jacquemier,<sup>1,2</sup> Daniel Birnbaum,<sup>1,8</sup> François Bertucci,<sup>1,9</sup> and Max Chaffanet<sup>1</sup>

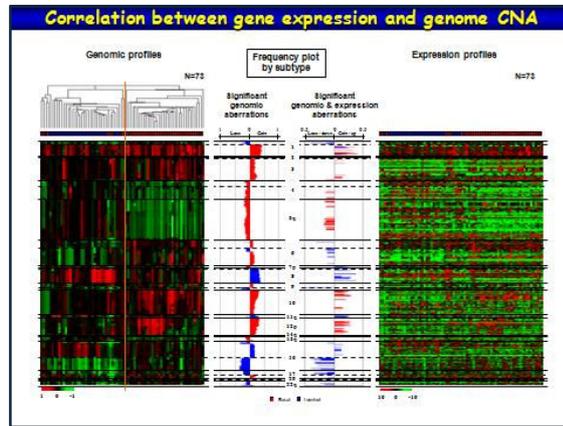
<sup>1</sup>INSERM U1024, Centre de Recherche Public-Cancer et UMR109 Institut National de la Santé et de la Recherche Médicale, Département de Biologie Oncologique, Département de Biophasiologie, and Département de Biologie Oncologique Institut Pasteur-Cancer; <sup>2</sup>Université de la Méditerranée, Institut de Médecine Translatoire du Service de Santé des Femmes, Unité de Recherche en Pharmacologie et Thérapeutique Nutrition, Marseille, France; <sup>3</sup>Department of Human Genetics, Nijmegen Centre for Molecular Life Sciences, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands; <sup>4</sup>Cancer Res 2007 December; 87 (24): 11866-76

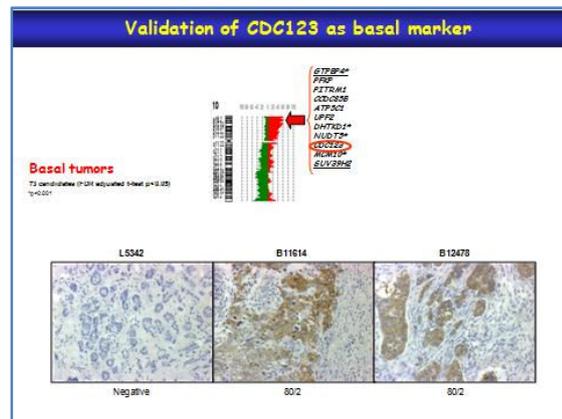
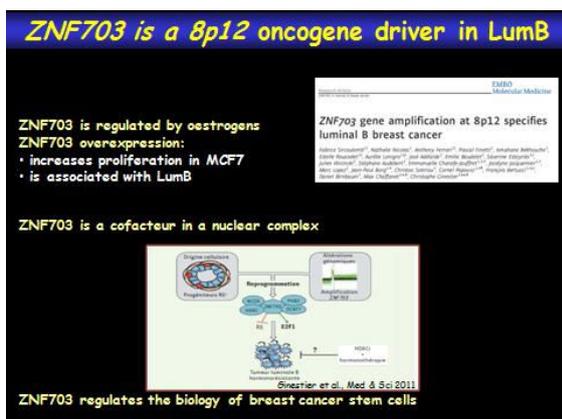
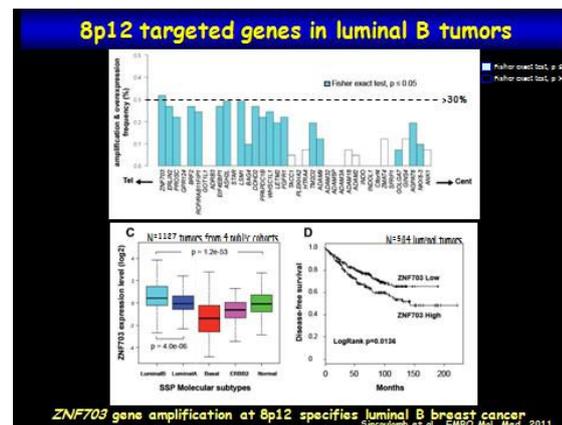
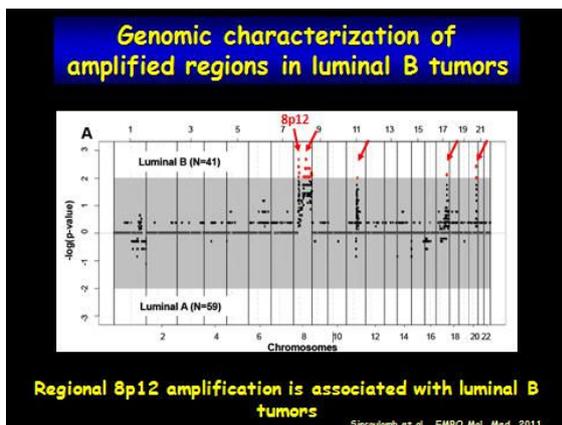
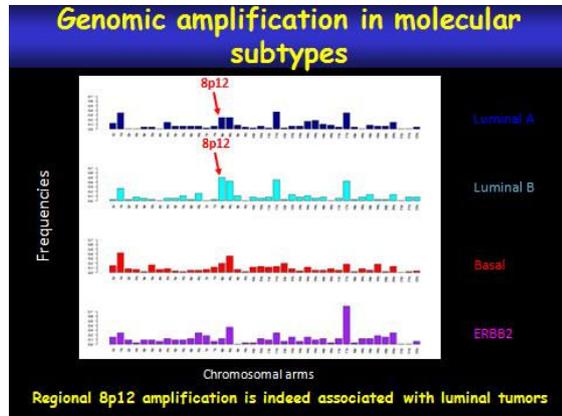
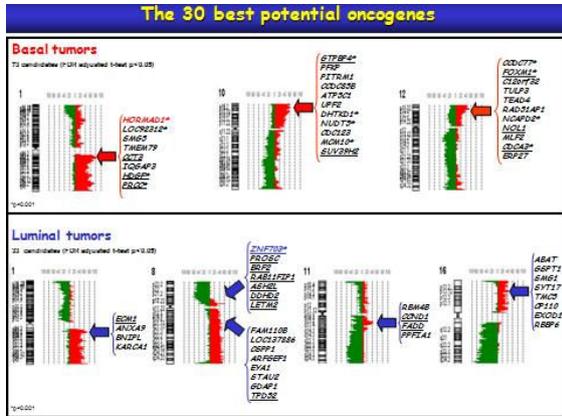
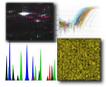


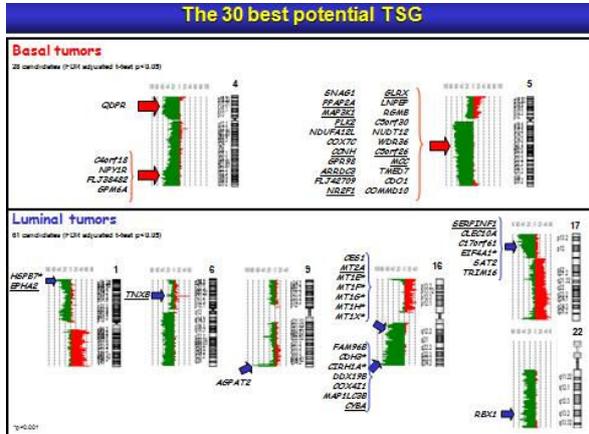
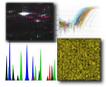
### Relation subtype/amplification signatures (FISH on TMA)

Subtype	Basal		p-value	ERBB2		p-value	Luminal		p-value
	No. of patients (%)	Non Basal		No. of patients (%)	Non ERBB2		No. of patients (%)	Non Lumina A	
<b>11q13/CCND1</b>									
Amplification	1 (5.9)	22 (49.0)	$p = 0.006$	2 (9.5)	21 (56.2)	NS	19 (57.6)	4 (12.1)	$p = 0.0003$
No Amplification	17 (94.9)	26 (59.9)		8 (75)	37 (93.8)		14 (42.4)	29 (87.9)	
<b>12p13/CDKN2</b>									
Amplification	5 (27.6)	2 (4.2)	$p = 0.02$	1 (11.1)	6 (10.5)	NS	1 (3.1)	6 (17.6)	NS
No Amplification	13 (72.4)	46 (95.8)		8 (68.9)	51 (99.5)		31 (96.9)	28 (82.4)	
<b>17q12/ERBB2</b>									
Amplification	0 (0)	6 (17.6)	NS	5 (71.4)	1 (2.2)	$p = 2.061e-06$	0 (0)	6 (20)	NS
No Amplification	19 (100)	26 (52.4)		2 (20.9)	45 (97.8)		23 (100)	24 (95)	

Letessier et al., BMC Cancer 2008







### CONCLUSION

In this pilot study, we had reported :

The integrated profiling of basal and luminal breast cancers (Adélaïde et al, Cancer Res., 2007).

We had shown that :

applying high resolution genomic analyses (244K aCGH [Agilent Technologies] + U133 Plus 2.0 human [Affymetrix]) leads to discovery of novel oncogenes and tumor suppressors, which will have a significant impact on our understanding of tumorigenesis and in the clinical management of cancer patients.

Sirotic et al. BMC Cancer 2010, 10:139  
http://www.biomedcentral.com/1471-2407/10/139

**RESEARCH ARTICLE** Open Access

### Genome profiling of *ERBB2*-amplified breast cancers

Fabrice Sirotic<sup>1\*</sup>, Imrane Belhouche<sup>1</sup>, Pascal Fivet<sup>1</sup>, José Adilola<sup>1</sup>, Azza Ben Hamida<sup>1</sup>, Julien Bonasia<sup>1</sup>, Stéphane Renaud<sup>1</sup>, Charline Innocent<sup>1</sup>, Emmanuelle Charle-Jauffret<sup>1</sup>, Colette Tappier<sup>1</sup>, Fahat Ben Ayed<sup>1</sup>, Patrice Verset<sup>1</sup>, Jocelyne Jacquemart<sup>1</sup>, François Bertucci<sup>1</sup>, Daniel Berbaum<sup>1</sup>, Max Chaffane<sup>1</sup>

\*Correspondence: sirotic@univ-st-etienne.fr

<sup>1</sup>UMR 1053, Centre de Recherche en Cancérologie, Université de Saint-Etienne, 42023 Saint-Etienne, France

Full list of author information is available at the end of the article

© 2010 Sirotic et al; licensee BMC Cancer. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sirotic et al., BMC Cancer 2010, 10:139

### *ERBB2* in breast cancers

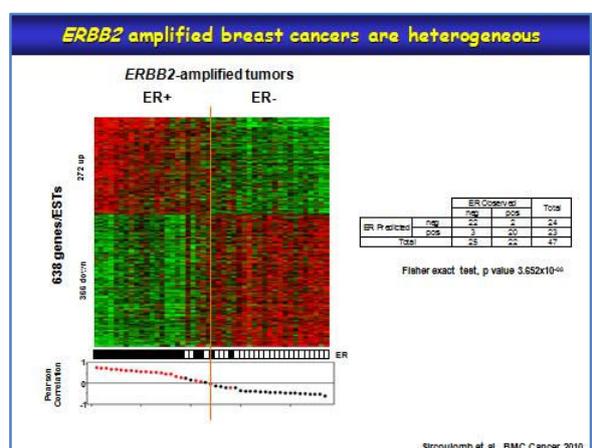
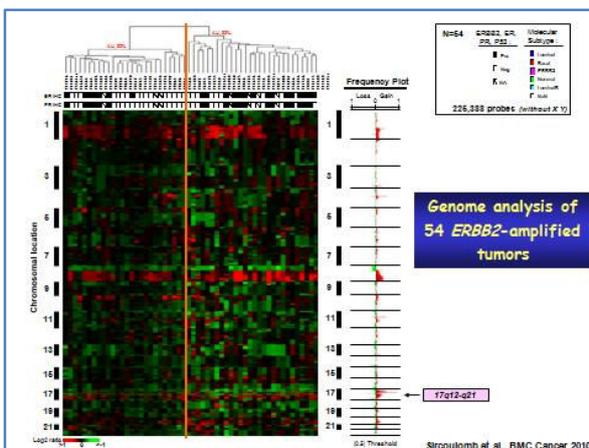
FISH on tissues

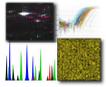
Not amplified

Amplified

Herceptin TTT

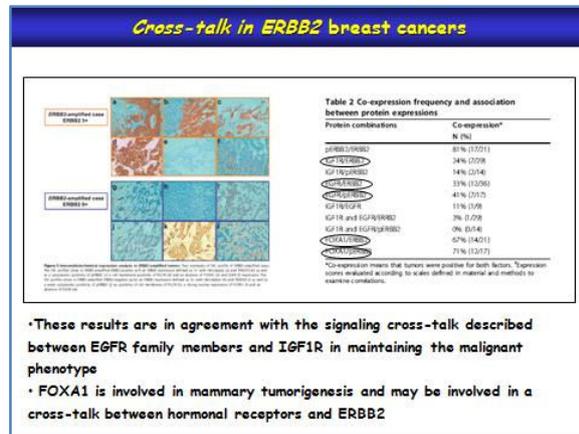
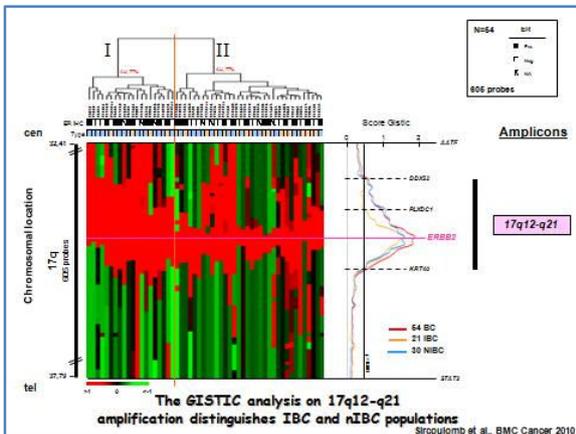
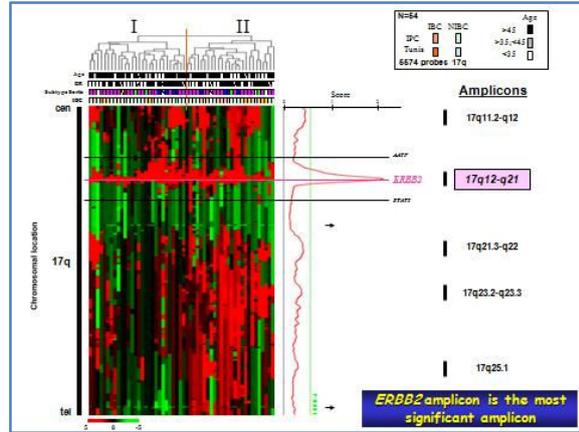
Ginestier et al., 2004





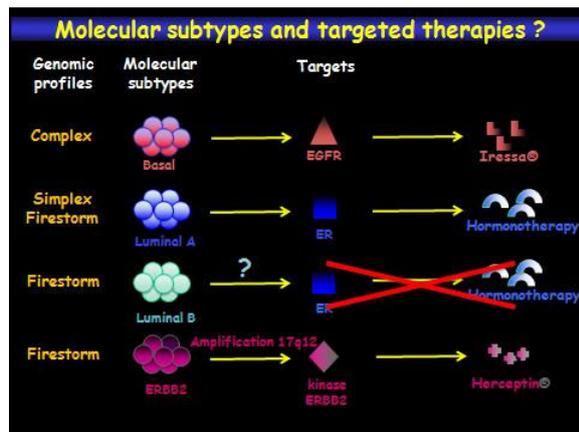
**Genes altered by CNA with a frequency significantly different in ER+ and ER- ERBB2-amplified tumors**

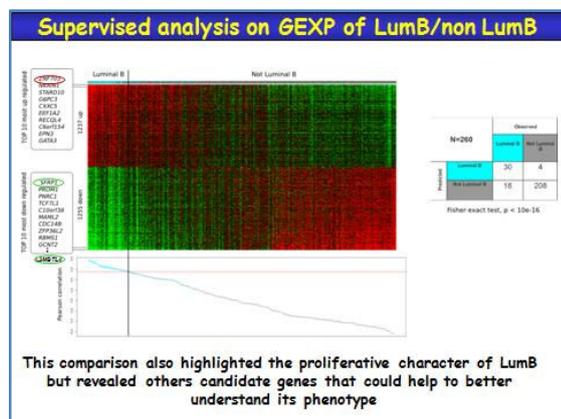
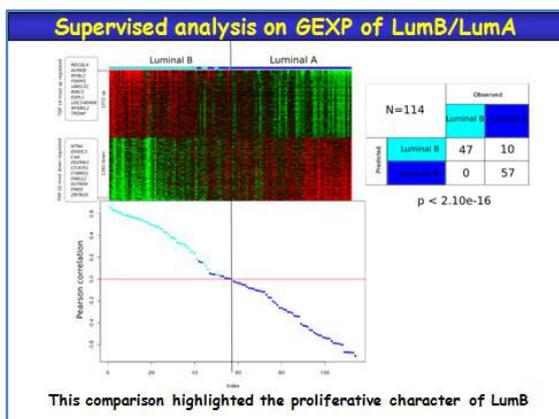
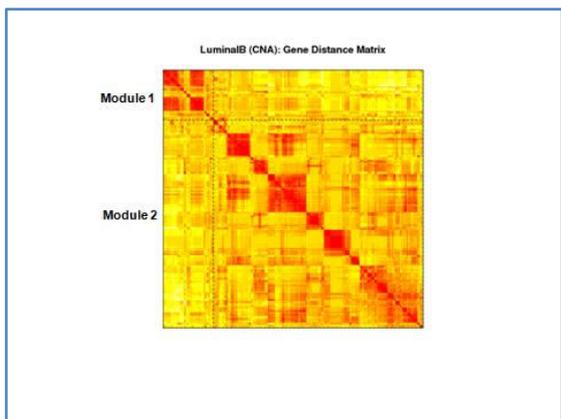
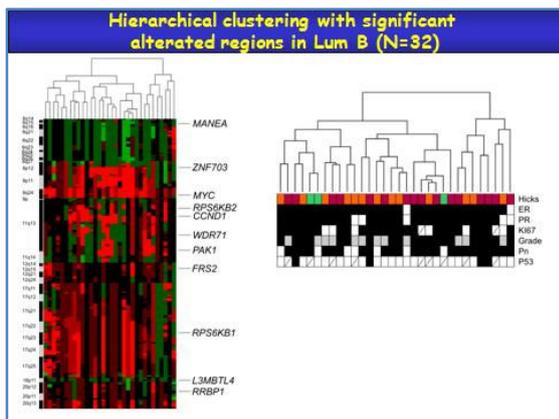
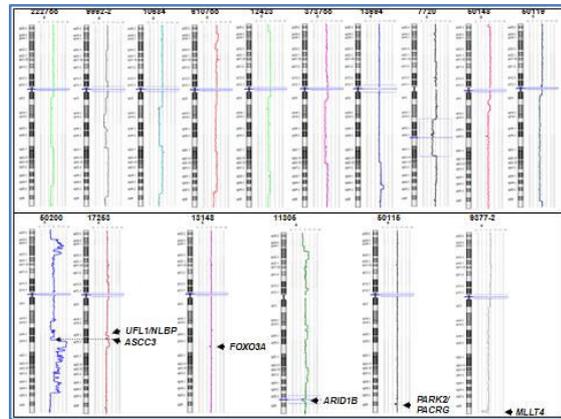
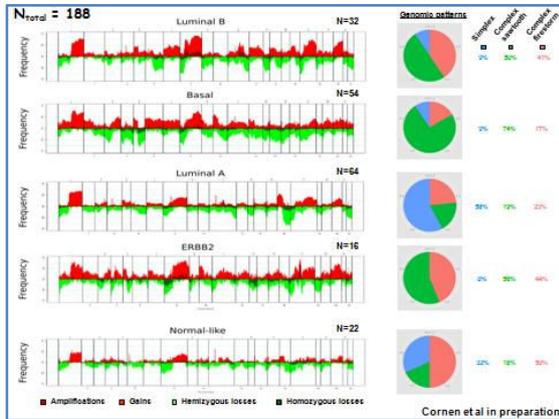
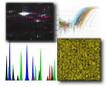
ER status of ERBB2 amplified tumors	Signaled by gene CNA	Altered gene	Location	CNA Type
ER-	gain	CDKN2A	11q13.3	
		ADAM10	9q12.12	Metastasis, IHC, CNV
		ADAM10	9q12.12	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
		CDK1	12p12.1	Metastasis, IHC, CNV
ER+	loss	CDKN2A	11q13.3	
		CDKN2A	11q13.3	



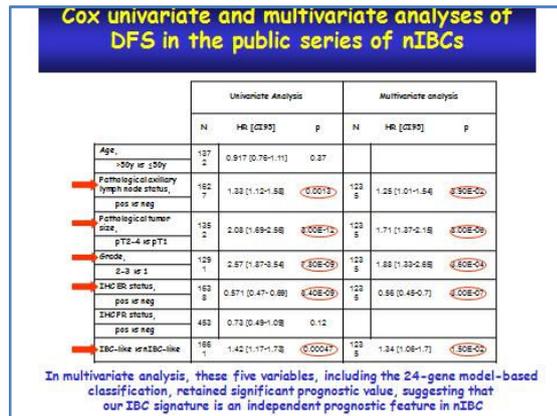
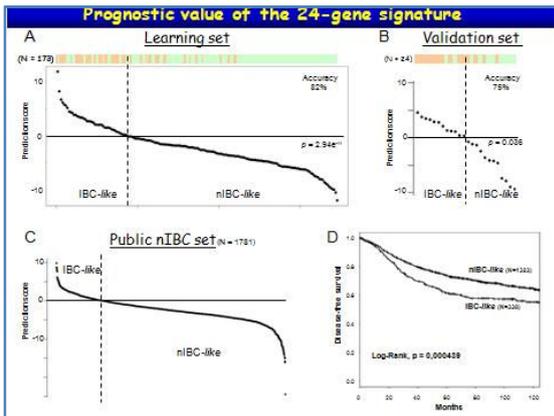
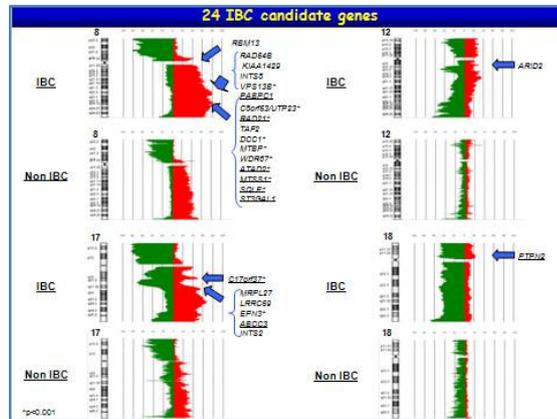
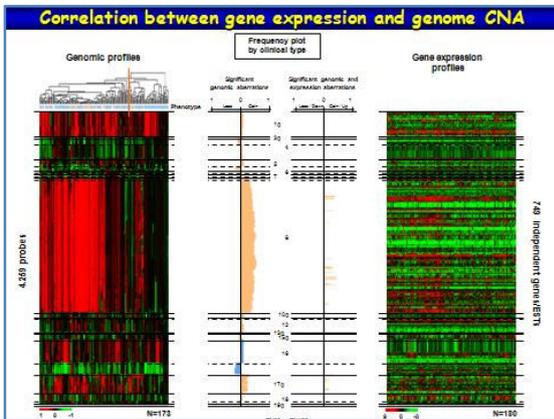
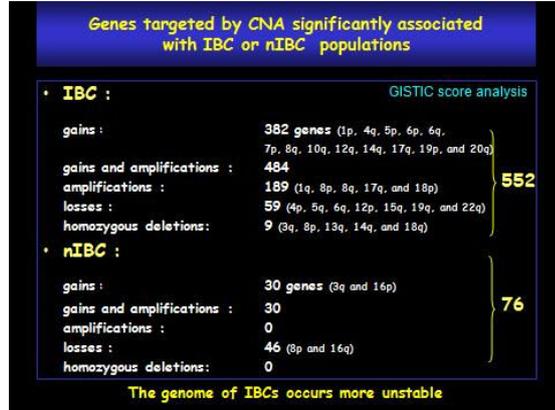
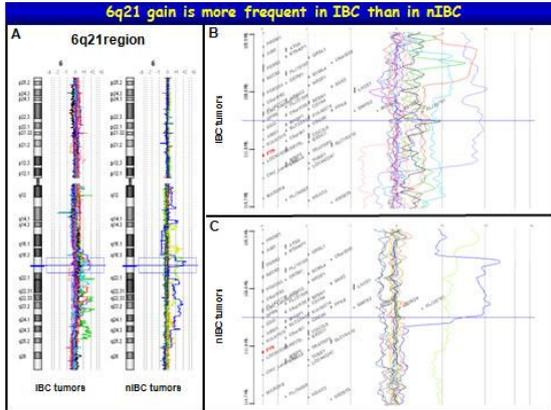
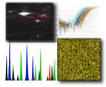
**Conclusions**

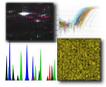
Our study shows that ERBB2 amplified BCs are heterogeneous with respect to clinical, immunohistochemical and molecular factors and identifies features that may be useful in the design of therapeutical approaches of these poor prognosis cancers.







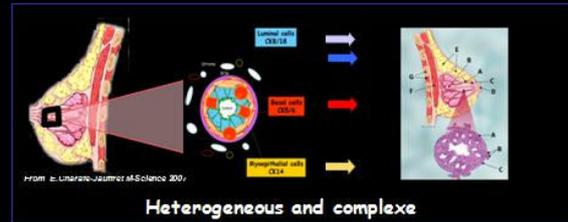




## Conclusion

- We showed the genomic complexity and heterogeneity of IBC.
- We defined, for the first time on a large scale and with high resolution, regions and genes targeted in IBCs. Many of them have not been previously reported in breast cancer, and deserve further functional validation.
- We identified 24 IBC-specific potential oncogenes that could explain, at least partially, the IBC phenotype and its aggressiveness and poor prognosis.

## Breast Cancer

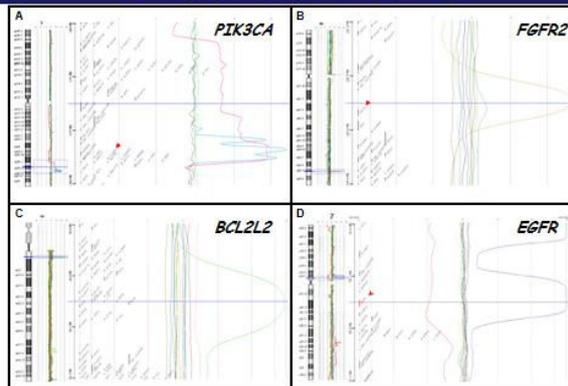


## Rare events in BCs

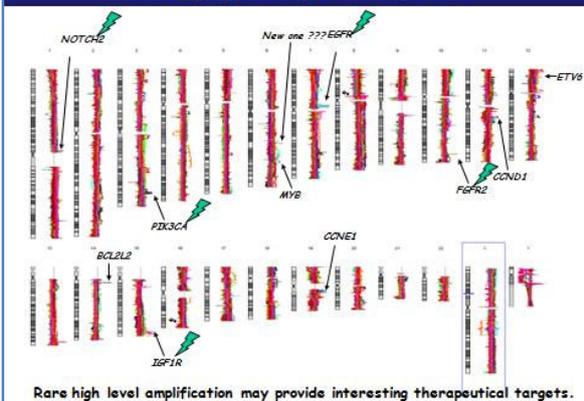
Are they useful

- 1- To guide pratican in the definition of additional therapeutic strategy ?
- 2- to define potential candidate ?

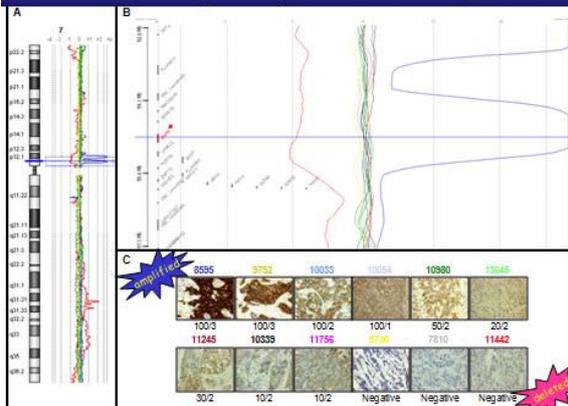
## Rare high gene amplification profiles

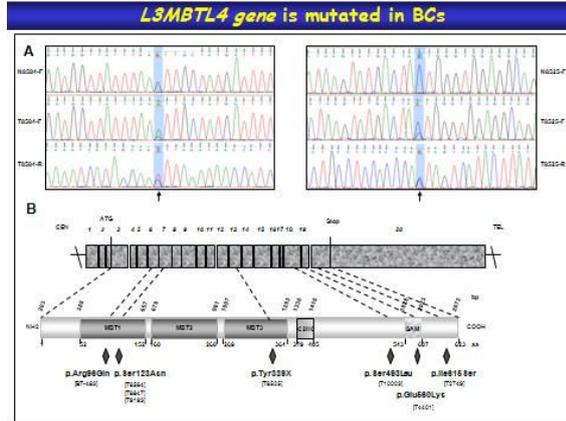
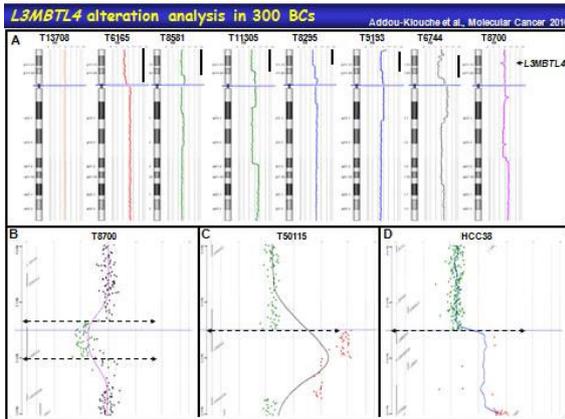
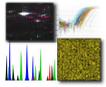


## Rare high gene amplification profiles



## Rare high EGFR gene amplification and its expression



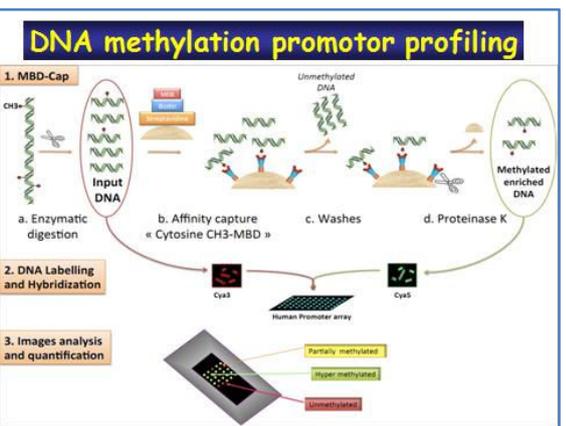
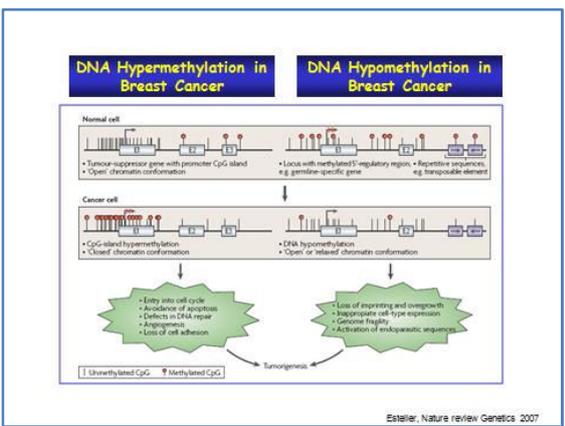


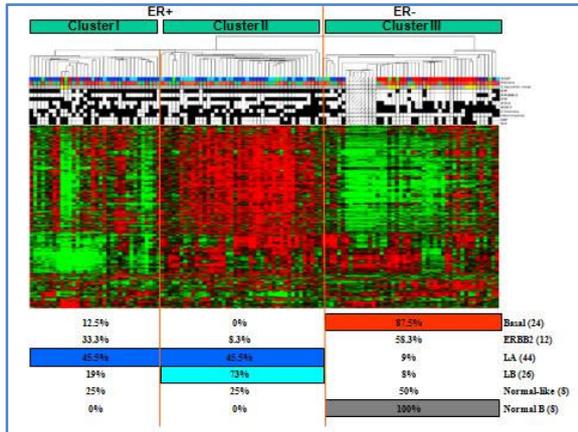
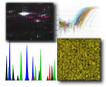
### Perspectives generated by Genomic studies

- Genome characterization
  - SCF
  - Frequent and rare events
  - List of candidate oncogenes and tumor suppressors
- Functional characterization
  - Cell models: siRNA, miRNA, shRNA, overexpression, 3D...
  - In vivo models: Transgenesis, KO
- Clinics and therapeutics
  - Role as markers, unique or within signatures
  - Role as targets (drug assays, drug development)

### Integrated profiling on BCs

Characteristics	n	%
Age	216	72.0%
Stage	199	61.9%
Pathologic type	208	64.3%
ER	208	64.3%
PR	207	63.8%
HER2	191	59.2%
Topo II	208	64.3%
HER3	208	64.3%
HER4	208	64.3%
HER5	208	64.3%
HER6	208	64.3%
HER7	208	64.3%
HER8	208	64.3%
HER9	208	64.3%
HER10	208	64.3%
HER11	208	64.3%
HER12	208	64.3%
HER13	208	64.3%
HER14	208	64.3%
HER15	208	64.3%
HER16	208	64.3%
HER17	208	64.3%
HER18	208	64.3%
HER19	208	64.3%





### Comparison of Promoter DNA methylation profiles

#### ER+/ER-

3484 promotor genes exhibited a significant DNA methylation difference between both groups, ER+ and ER- (t-test,  $q < 0.05$ ).

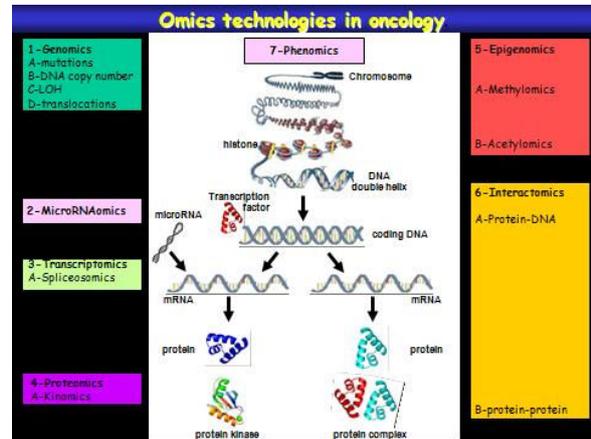
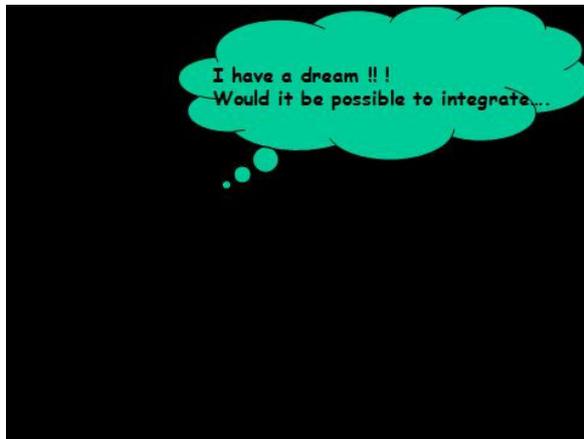
1753 exhibited a DNA methylation higher in ER+ than ER- (they included *APC, CAV1, CCND2, CDC47, CDH3, CDKN2A, CDKN2B, HEY2, RECK, RASSF1*)

1731 exhibited a DNA methylation higher in ER- than ER+ (they included *BCL2, ESRI, HSD17B4, PISD, WNK4*)

In agreement with previous studies :

Sunami et al, Breast Cancer Res 2006 ; Bediaga et al., Breast Cancer Res 2010 ; Holm et al., Breast Cancer Res 2010 ; Jovanovic et al, Molecular Oncology 2010 ; Fackler et al, Cancer Research 2011 ; Hill et al, Cancer Research 2011

The promoter array approach occurs robust, more resolute and generates similar results as Illumina method



**The breast cancer program**

Daniel Birnbaum

Max Chaffanet,  
 José Adélaïde,  
 Anne Letessier,  
 Stéphanie Cornen,  
 Marie-rose Saade,  
 Arnaud Guille,  
 Fabrice Sircoulomb,  
 Ismahane Bekhouche,  
 Stéphane Raynaud,  
 Lynda Klouche

Jocelyne Jacquemier,  
 Emmanuelle Charafe-Jauffret,  
 C. Ginestier,  
 Jeannine Genik,  
 Florence Monville,  
 Julien Wienski,

François Bertucci,  
 Nathalie Cervera,  
 Pascal Finetti

Marc Lopez, Antony Ferrari

INSTITUT PAOLI CALMETTES  
 Inserm  
 Agence Nationale de la Recherche  
 LA LIQUEUR  
 CAMPLP  
 GEFLUC

**Agilent 244K aCGH were also used in**

**Cancer du sein**

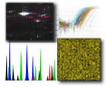
- Gelin-Boyer V, et al. Comprehensive profiling of 9p11-12 amplification in breast cancer. *Molecular Cancer Research* 2005;3(12):653-657.
- Adelaide J, et al. Integrated profiling of basal and luminal breast cancers. *Cancer Res* 2007;67(24):11265-11275.
- Adelaide J, et al. Absence of ESR1 amplification in a series of breast cancers. *Int J Cancer* 2008;113(12):2970-2972.
- Adkoy-Khorche L, et al. Loss, mutation and deregulation of L3MBTL4 in breast cancer. *Mol Cancer* 2010;9(1):213.
- Selinger R et al. BRD2 homozygous deletion, a possible alternative to BRCA1 mutation in basal breast cancer. *Genes Chromosomes Cancer* 2010;49(12):1143-51.
- Sircoulomb F, et al. Genome profiling of ERBB2-amplified breast cancers. *EMC Cancer* 2010, 10:339.
- Belkhouche J, et al. Genome profiling of inflammatory breast cancer. *PLoS ONE* 2011, 6(2):e16900.
- Sircoulomb F, et al. ZNF703 gene amplification at 9p11 specifies human B breast cancer. *EMBO Mol Med* 2011;3(3):153-66.
- Fournier G, et al. Loss of AFB1Alfa1, a marker of poor outcome in breast cancer, induces cell migration, invasiveness and tumor growth. *Oncogene* 2011, 30(36):3562-74.
- Sabatier R, et al. Down-Regulation of ECRG4, a Candidate Tumor Suppressor Gene, in Human Breast Cancer. *PLoS One* 2011;6(11):e27456.
- Corson S, et al. Mutations and deletions of ARID1A in breast tumors. *Oncogene* 2012 (online).

**Leucémies**

- Emmeo A, et al. Translocations involving chromosome 12q arm in myeloproliferative disorder: possible role of *KMG47* and *SOC37* genes. *Cancer Genet Cytogenet* 2007;178(1):80-83.
- Gelin-Boyer V, et al. Genome profiling of chronic myelomonocytic leukemia: frequent alterations of RAS and RUNX1 genes. *EMC Cancer* 2008;4(1):299.
- Morati A, et al. Alteration of the MYB locus in MYST3-linked cases. *Leukemia*, 2009;23(1):85-94.
- Bertrand F, et al. Alterations of NF1A in chronic malignant myeloid diseases. *Leukemia*, 2009;23(3):583-585.
- Gelin-Boyer V, et al. Mutations of polycomb-associated gene *ASXL1* in myelodysplastic syndromes and chronic myelomonocytic leukemia. *Br J Haematol* 2009;145(6):88-90.
- Adelaide J, et al. Gain of CBL-interacting protein, a possible alternative to CBL mutations in myeloid malignancies. *Leukemia* 2010;24(3):1259-61.
- Gelin-Boyer V, et al. *ASXL1* mutation is associated with poor prognosis and acute transformation in chronic myelomonocytic leukemia. *Br J Haematol* 2010;121(4):582-7.
- Brocqville M, et al. Rare mutations in DNMT3A in myeloproliferative neoplasms and myelodysplastic syndromes. *Blood Cancer Journal* 2011 (in press).
- Brocqville M, et al. Mutations and deletions of the *SUZ12* polycomb gene in myeloproliferative neoplasms. *Blood Cancer Journal* 2011 (in press).

**Autres cancers**

- Birnbaum DJ, et al. Genome profiling of pancreatic adenocarcinoma. *Genes Chromosomes Cancer* 2011, 50(6):456-65.



### Proteomic Tools in Cancer Research

Stéphane Audebert

March, 6th 2012



Institut Paoli-Calmettes  
Marseille, France

Centre de Recherche en Cancérologie de Marseille



Stephane.audebert@inserm.fr



### Institut Paoli-Calmettes

- Comprehensive Cancer Center
- Hospital / Cancer
- Center for Cancer Research



Centre de Recherche en Cancérologie de Marseille



Stephane.audebert@inserm.fr



### MaP: Marseille Proteomic

- Created in 2008
  - National Labelization IBiSA, 2008, 2010, 2012
  - Synergy of Marseillaises platforms to share instruments, bioinformatics resources and human expertise
- Increase proteomic strength to develop proteomic projects and to provide optimized services for proteomics



Centre de Recherche en Cancérologie de Marseille



Stephane.audebert@inserm.fr



### Our Specificities at Institute Paoli-Calmettes

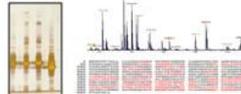
#### Basic Research

NEW PROTEIN  
COMPLEX  
DISCOVERY

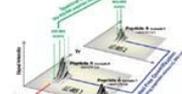
#### Clinical Research

BIOMARKERS  
DISCOVERY

#### Analytical Proteomic



#### Quantitative Proteomic



de Marseille



Stephane.audebert@inserm.fr



### MaP: Institut Paoli-Calmettes Platform

#### Staff :

- Scientific Director (JP Borg)
- Medical Oncologist (A Gonçalves)
- 5 Engineers (Mass spectrometry)
- 1 Engineers (Y2H)
- 1 bio-informatician
- 1 technician



#### Space :

- Biochemistry lab (70 m<sup>2</sup>)
- Sample Prep lab ( 20 m<sup>2</sup>)
- 2 Instrumentation labs (30 m<sup>2</sup>)
  - Office (20 m<sup>2</sup>)



Centre de Recherche en Cancérologie de Marseille



Stephane.audebert@inserm.fr



### Mass Spectrometer : MalDI-TOF-TOF

Protein identification by PMF and MSMS

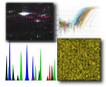


Centre de Recherche en Cancérologie de Marseille



Stephane.audebert@inserm.fr





### MALDI-TOF is a Complement to 2D-DiGE approach

1) Cy2, Cy3, Cy5 label

2) 2-DE Run

3) Statistical Analysis

4) Identification

5) Validation

Centre de Recherche en Cancérologie de Marseille

MAP  
Marseille Proteomique

Stéphane.audebert@inserm.fr

Inserm

UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### NanoLC LTQ-velos-Orbitrap-ETD

#### Dedicated to Biomarkers Discovery PTM analysis

Centre de Recherche en Cancérologie de Marseille

MAP  
Marseille Proteomique

Stéphane.audebert@inserm.fr

Inserm

UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### Automates for samples preparation

Centre de Recherche en Cancérologie de Marseille

MAP  
Marseille Proteomique

Stéphane.audebert@inserm.fr

Inserm

UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### What is proteomic ?

Analysis, description of a proteome at “t” time

Centre de Recherche en Cancérologie de Marseille

MAP  
Marseille Proteomique

Stéphane.audebert@inserm.fr

Inserm

UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### For a long time...Proteomic was...

Affinity (CoIP, Pulldown)      Differential Expression      Chromatography...

What is this Protein ?

Centre de Recherche en Cancérologie de Marseille

MAP  
Marseille Proteomique

Stéphane.audebert@inserm.fr

Inserm

UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### Now.....What are proteins in my sample ?

#### Various Biological Questions

Proteome + or - complexe

Description

- identification, PTM,
- Alternative Splices

**Stem cells**

**Qualitative Proteomic**

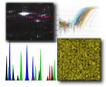
Centre de Recherche en Cancérologie de Marseille

MAP  
Marseille Proteomique

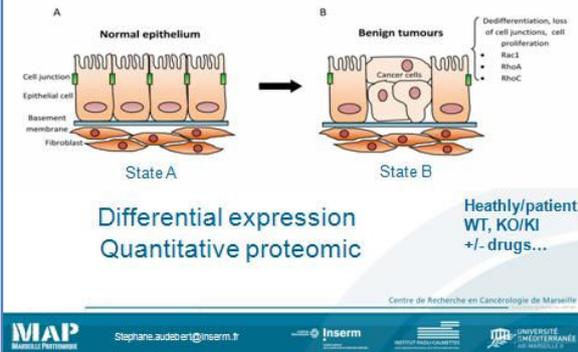
Stéphane.audebert@inserm.fr

Inserm

UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

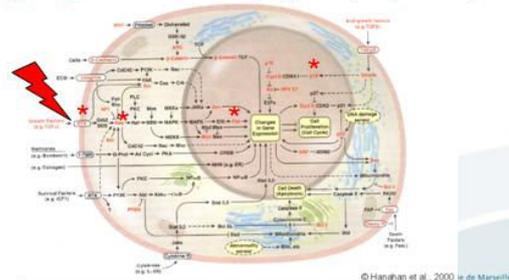


### What are Proteins change in my sample ?

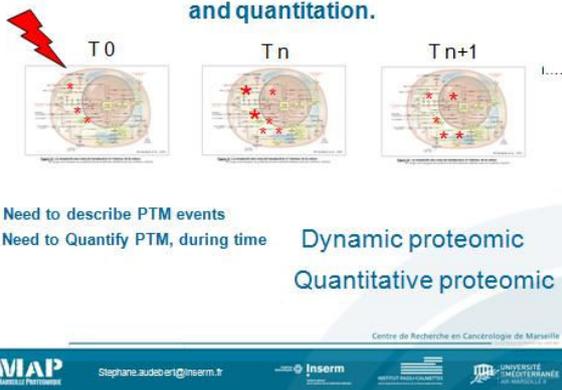


### Signalling Networks - What changes in cells upon activation ?

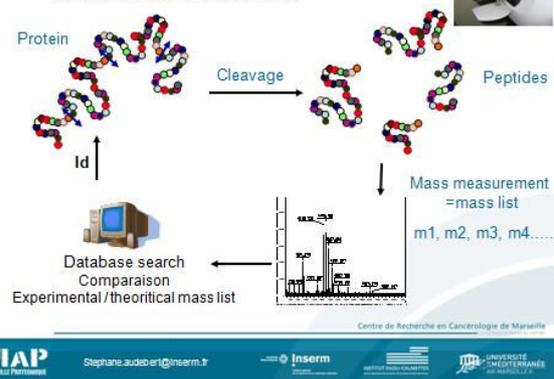
Changes in proteins PTM: Phosphorylations, Ubiquitination,



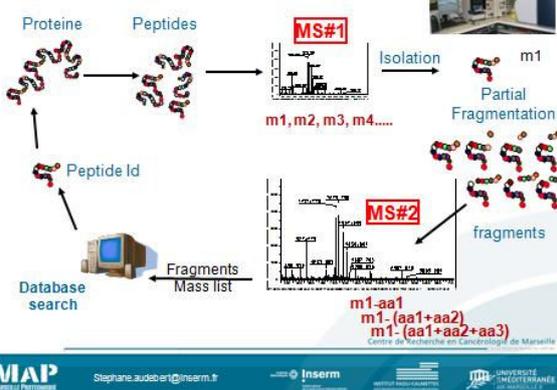
### Kinetics of PTM upon cells activation and quantitation.



### Principle of Protein Identification by PMF (Peptide Mass Fingerprint)

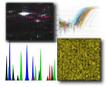


### MS/MS identification (Peptide Sequencing)



### LC-MS/MS...thousands of MS/MS/analysis





### LTQ-velos-Orbitrap-ETD

#### Two analysers

- Resolution 100,000
- Mass accuracy < 5ppm
- High Scan rate
- Dynamic range 10<sup>5</sup>

Centre de Recherche en Cancérologie de Marseille

MAP  
Stéphane.audebert@inserm.fr

## BIOMARKERS DISCOVERY

Centre de Recherche en Cancérologie de Marseille

MAP  
Stéphane.audebert@inserm.fr

### Need of Identification of new biomarkers

Centre de Recherche en Cancérologie de Marseille

MAP  
Stéphane.audebert@inserm.fr

### Need to discover new biomarkers

For diagnostic  
And more ... To predict response to therapeutic treatment

Example :

Breast cancer :  
Her2 Overexpression in 20% breast cancer  
Treatment using Herceptin (monoclonal antibody against Her2)  
Some patients do not respond.

Why ? Can we predict response to herceptin treatment

Centre de Recherche en Cancérologie de Marseille

MAP  
Stéphane.audebert@inserm.fr

### Need to discover new biomarkers

Need of full annotated sample  
[www.biobank-paca.com](http://www.biobank-paca.com)

Centre de Recherche en Cancérologie de Marseille

MAP  
Stéphane.audebert@inserm.fr

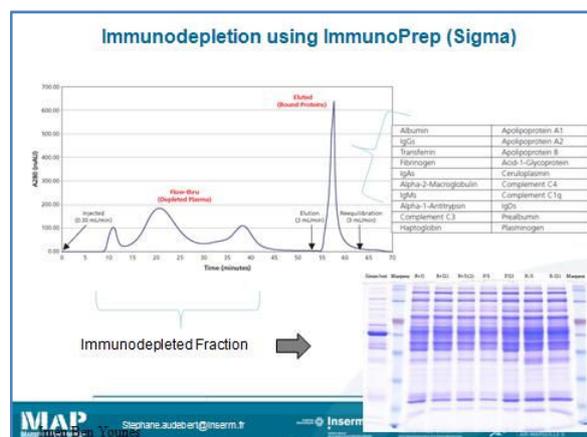
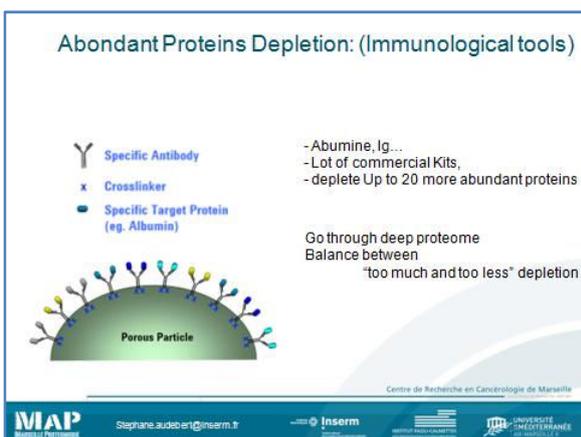
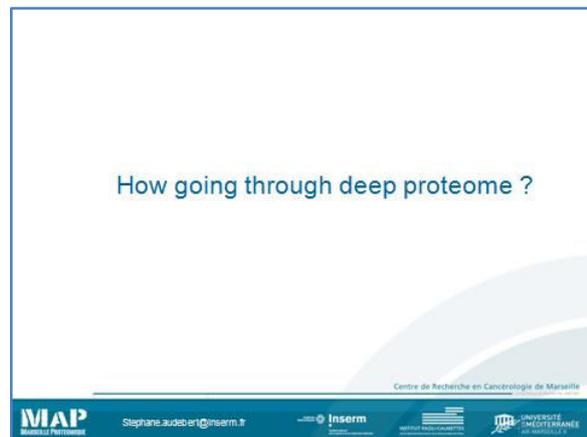
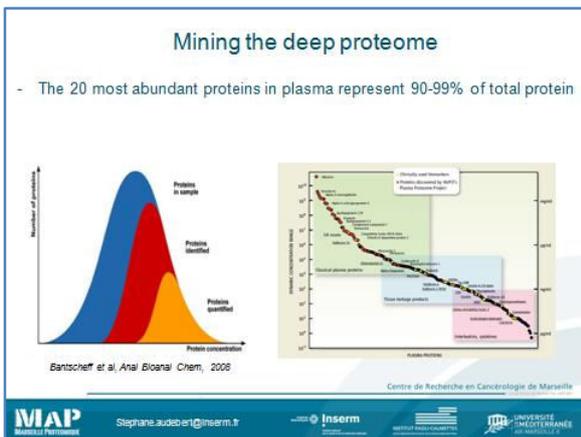
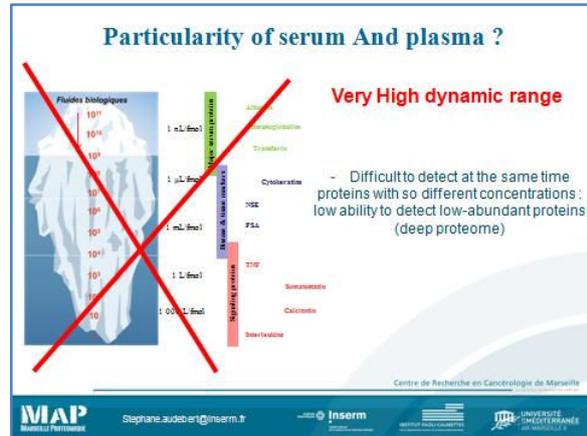
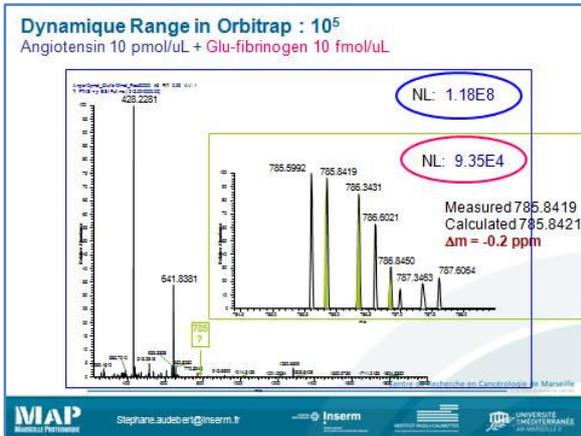
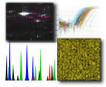
### Particularity of serum And plasma ?

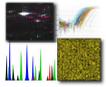
**Very High dynamic range**

- Difficult to detect at the same time proteins with so different concentrations :  
low ability to detect low-abundant proteins (deep proteome)

Centre de Recherche en Cancérologie de Marseille

MAP  
Stéphane.audebert@inserm.fr





### Equalization, Proteomer (Biorad)

- Hexapeptides synthesized according to the Merrifield approach (Science, 1965)
- Combinatorial chemistry (Split, Couple, Recombine)
- 64.10<sup>6</sup> different hexapeptides
- One unique ligand per beads

Large Dynamic range → Peptide Ligand Library (10<sup>6</sup> ligands) → Reduced dynamic range

Blue arrow: Dilute abundant and concentrate trace proteins  
Red arrow: Keep only the low abundance proteins

MAP Marseille Proteomics | Stéphane.audebert@inserm.fr | Inserm | UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### Equalization

Decrease dynamic range

Centre de Recherche en Cancérologie de Marseille

MAP Marseille Proteomics | Stéphane.audebert@inserm.fr | Inserm | UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### Fractionation

- MW, SDS-PAGE  
- 20-30 cut bands
- Isoelectric point fractionation  
- OffGel (Agilent) (12 or 24 fractions)
- Chromatographic resins,  
- Ions Exchanges (x fractions)

Complementary Approaches Depletion/Equalisation and Fractionation

Centre de Recherche en Cancérologie de Marseille

MAP Marseille Proteomics | Stéphane.audebert@inserm.fr | Inserm | UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### Biomarkers Discovery

Centre de Recherche en Cancérologie de Marseille

MAP Marseille Proteomics | Stéphane.audebert@inserm.fr | Inserm | UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### How to compare

Healthy/patient  
Response/noResponse

Centre de Recherche en Cancérologie de Marseille

MAP Marseille Proteomics | Stéphane.audebert@inserm.fr | Inserm | UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### Need of quantitative methods

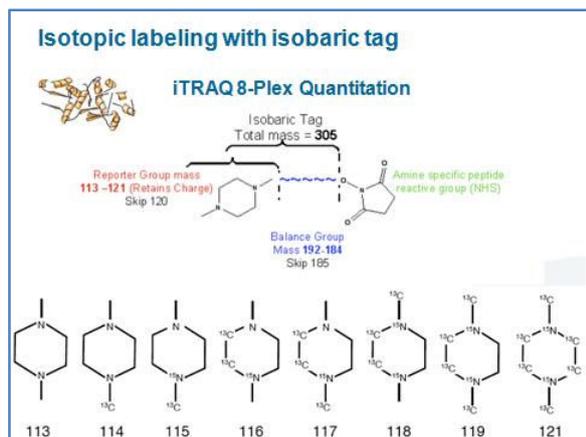
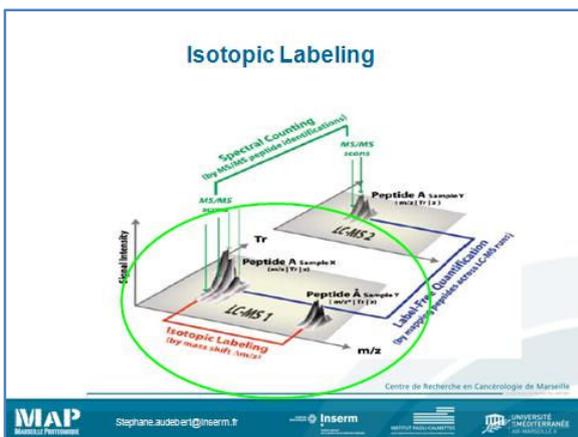
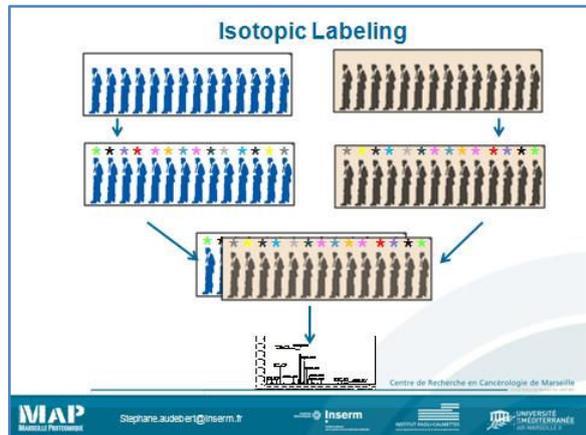
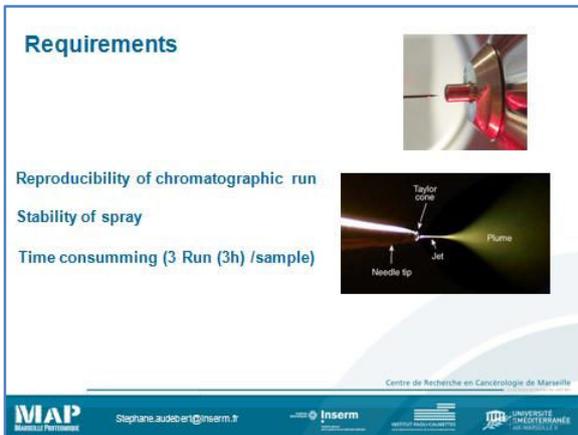
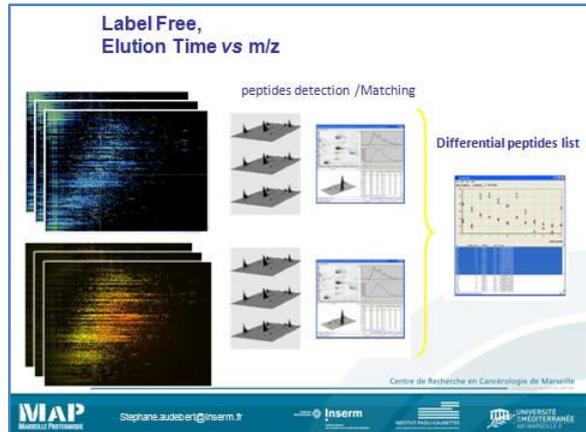
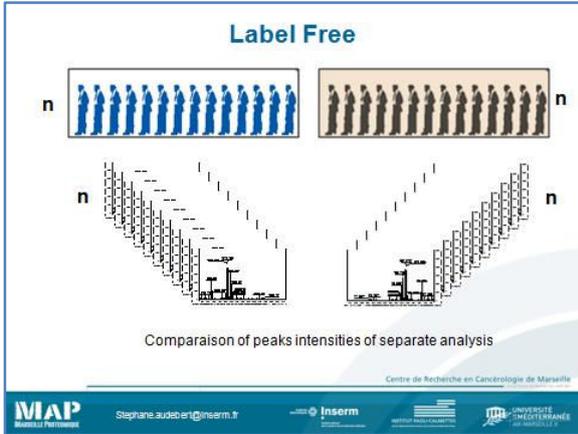
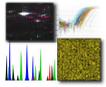
Spectral Counting (by relative peptide identification)  
MS/MS ratios  
Isotopic Labeling (by mass shift, 15N/13C)

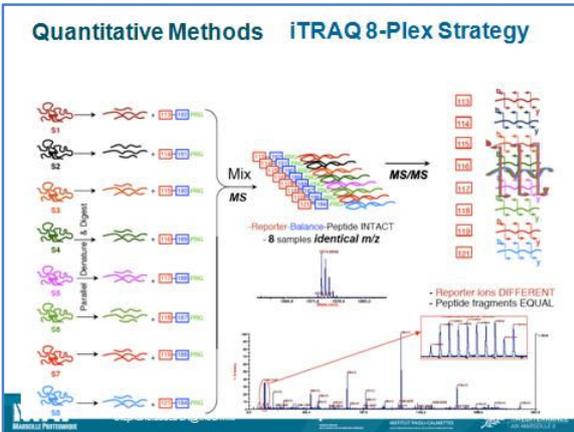
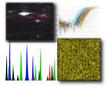
Peptide A sample X (m/z 123)  
Peptide A sample Y (m/z 123)  
Peptide A sample X (m/z 123)  
Peptide A sample Y (m/z 123)

LC-MS 1  
LC-MS 2

Centre de Recherche en Cancérologie de Marseille

MAP Marseille Proteomics | Stéphane.audebert@inserm.fr | Inserm | UNIVERSITÉ MEDITERRANÉE DE MARSEILLE





#### Advantage :

- Up to eight sample simultaneously analysed  
in one LC MSMS run

#### Limitations :

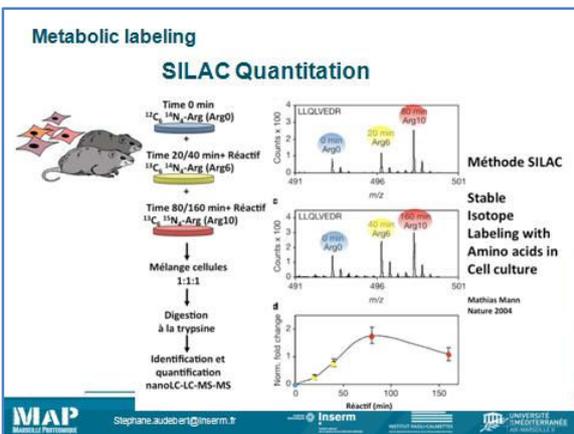
- no more than eight sample simultaneously analysed

One way to manage limitation of  
number of sample is to use Serum  
Pool (responder/non responder)

Centre de Recherche en Cancérologie de Marseille



Stephane.audebert@inserm.fr



Label free or isotope labeling Quantification are good strategies  
for Biomarkers discovery

Cancerfield

Others pathologies

...usefull tools in basic research for many projects

Centre de Recherche en Cancérologie de Marseille



Stephane.audebert@inserm.fr



### Targeted proteomic (SRM, Single Reaction Monitoring)

We know what we wish to quantify

- Specific biomarkers in serum...
- Specific biomarkers in biopsies....

**Absolute quantification**

- Need of standart peptides

Centre de Recherche en Cancérologie de Marseille



Stephane.audebert@inserm.fr



### Biomarker validation



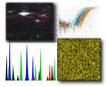
Add of Proteotypic peptide in each sample  
Heavy peptide  
Used as internal standart for absolute quantification

Centre de Recherche en Cancérologie de Marseille



Stephane.audebert@inserm.fr





### SRM Assays

Instruments : TripleQuadrupole

**Selection in MS**      **Selection in MS-MS**

Peptide Selection      Fragmentation      Fragment Selection

MS can be programmed to focus on only targets of interests  
Two levels of mass selection in MS and MS-MS  
Background ions are filtered out

Centre de Recherche en Cancérologie de Marseille

MAP Marseille Provence      Stéphanie.audebert@inserm.fr      Inserm      UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### MRM : Multiple reaction Monitoring

Advantage :  
Up to 100 proteins can be quantified in a same analysis.  
- Fast and cheaper than Elisa tests  
- Good tool for biomarkers validation

Disadvantages :  
- Price of peptides

After, biomarkers discovery

→ biomarkers Validation

Centre de Recherche en Cancérologie de Marseille

MAP Marseille Provence      Stéphanie.audebert@inserm.fr      Inserm      UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### My friends from Proteomic lab

Centre de Recherche en Cancérologie de Marseille

MAP Marseille Provence      Stéphanie.audebert@inserm.fr      Inserm      UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### Many Thanks

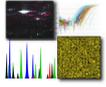
Centre de Recherche en Cancérologie de Marseille

MAP Marseille Provence      Stéphanie.audebert@inserm.fr      Inserm      UNIVERSITÉ MEDITERRANÉE DE MARSEILLE

### Many Thanks

Centre de Recherche en Cancérologie de Marseille

MAP Marseille Provence      Stéphanie.audebert@inserm.fr      Inserm      UNIVERSITÉ MEDITERRANÉE DE MARSEILLE



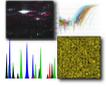
## WORKSHOP TRAINING SESSIONS

### Proteomics

- I. MALDI-TOF MS and MS/MS analysis of proteins/peptides**
- II. MASCOT database searches**
- III. Bacterial identification by MALDI-TOF MS**

Protein identification by mass spectrometry (MS) can be performed using sequence-specific peptide fragmentation or peptide mass fingerprinting (PMF), also known as peptide mass mapping. The standard approach to identify proteins includes separation of proteins by gel electrophoresis or liquid chromatography. Subsequently, the proteins are cleaved with sequence-specific endoproteases, most notably trypsin. Following digestion, the generated peptides are investigated by determination of molecular masses or generation of peptide fragments. For protein identification, the experimentally obtained masses are compared with the theoretical peptide masses of proteins stored in databases by means of mass search programs. Matrix-assisted laser/desorption ionization mass spectrometry (MALDI-MS) is the most commonly used technique to perform PMF. MALDI-MS is fast, robust, easy to perform, sensitive (low fmol range), accurate (low ppm range), tolerant to a certain level of various contaminants, and can be automated. The predominant detection of singly charged peptide molecules by MALDI-MS facilitates the evaluation of PMFs significantly. Although PMF is an effective tool for the identification of relatively pure proteins, it often fails to identify protein mixtures. Separation of complex protein samples by high-resolution two-dimensional gel electrophoresis (2-DE) is well adapted to protein identification with PMF. On the other hand, the application of PMF in combination with one-dimensional gel electrophoresis or liquid chromatography must be adjusted to the separation capacity. A frequently used strategy to identify proteins by MS is to first generate a PMF because of the simplicity of the method. Sequence-specific peptide fragmentation is necessary if no protein was identified unambiguously. Peptide fragments can be generated directly by MALDI-MS using post-source decay (PSD)-MALDI-MS or collision induced dissociation (CID)-MALDI-MS/MS. The relative new MALDI-ion traps, MALDI-Q-TOF and MALDI-TOF/TOF instruments are particularly suitable for this purpose.

Two-dimensional gel electrophoresis is the preferred method for protein separation prior to PMF. Gel separated proteins have to be enzymatically or chemically cleaved prior to mass analysis. In-gel digestion is used predominantly, but on-membrane digestions can be performed as well [5,6]. However, the blotting efficiency is dependent on the blotting conditions and different for high and low molecular proteins. Preparative 2-DE gels prior to protein identification by mass spectrometry are typically stained either by silver (without glutaraldehyde) or Coomassie Blue G-250. Silver staining even without glutaraldehyde reduces the sequence coverage considerably. Fluorescent labeling of proteins can be achieved with ruthenium II tris (bathophenanthroline disulfonate) or



SyproRuby, but needs special equipment. Different laboratories favor distinct staining techniques. As a rule of thumb, spots detectable by Coomassie Blue G-250 can be identified by PMF. Trypsin is the favored enzyme for PMF. Trypsin is used to cleave proteins by hydrolyzing peptide bonds at the carboxyl side of arginyl and lysyl residues. Trypsin is relatively cheap, highly effective, and generates peptides with an average size of about 8–10 amino acids, ideally suited for analysis by MS.

## Destaining gels (Silver and Coomassie stained)

### **Materials needed:**

Pipetors with range: 0 - 20  $\mu$ l; 0-100  $\mu$ l; 0-1000  $\mu$ l.

Eppendorf tubes 1.5 ml and 0.5 ml, razors.

### **Reagents:**

ddH<sub>2</sub>O, Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>, K<sub>3</sub>Fe(CN)<sub>6</sub>, methanol .

(All with high purity > 99.5%)

### **Solutions:**

100 mM Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>

1.24 g Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> in 50 ml ddH<sub>2</sub>O

30 mM K<sub>3</sub>Fe(CN)<sub>6</sub>

0.49 g K<sub>3</sub>Fe(CN)<sub>6</sub> in 50 ml ddH<sub>2</sub>O

Destaining buffer

500  $\mu$ l 100 mM Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> + 500  $\mu$ l 30 mM K<sub>3</sub>Fe(CN)<sub>6</sub>

(50% 100 mM Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>, 50% 30 mM K<sub>3</sub>Fe(CN)<sub>6</sub>)

(Prepare fresh destaining solution)

### **Procedure:**

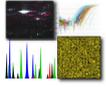
*During In Gel Digestion the purest chemicals available should be used. Gloves should be worn to avoid contamination by human epidermal proteins (keratin). To beware of talcum powder or dust contamination the gloves should be carefully rinsed with water before use. The razor should be washed with water and methanol before use.*

1. Wash the gel slab with UHQ (2 changes, 10 min each)
2. Use a clean scalpel to excise the spot of interest from the gel. Cut as close to the protein band as possible to reduce the amount of background gel. Excise a gel piece of roughly the same size from a gel region which does not carry any protein to use it as a control (blank).
3. Cut the excised piece into roughly 1 mm<sup>3</sup> cubes, and transfer them to a clean 0.5 mL microfuge tube.

If you have Coomassie stained gels continue with the procedure for In-gel digestion.

If you have Silver stained gels go through following steps:



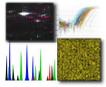


Digestion buffer: 95% 50 mM NH<sub>4</sub>HCO<sub>3</sub> / 5% Acetonitrile  
5 mL 100 mM NH<sub>4</sub>HCO<sub>3</sub> + 500 µL Acetonitrile  
bring to final volume of 10 ml

(Always prepare fresh solutions, except for the 0.1% TFA)

**Procedure:**

1. Wash the gel particles with UHQ once for 15 min. Solvent volumes used in the washing steps should roughly equal to the gel volume (20-40 µl).
2. Wash the gel particles with UHQ/ACN 1:1 (v/v), twice, 15 min/change.
3. Remove all liquid and add one gel volume of ACN to cover the gel particles for 5 min or until gel pieces have shrunk (they become milky and stick together).
4. Remove the ACN and rehydrate the gel pieces in the same volume of 100 mM NH<sub>4</sub>HCO<sub>3</sub> for 5 min.
5. Add an equal volume of ACN (to get 100 mM NH<sub>4</sub>HCO<sub>3</sub> / ACN, 1:1) and incubate for 15 min.
6. Remove all liquid and dry gel particles in a vacuum centrifuge (~15 min).
7. Swell the gel particles in a gel volume of reducing solution and incubate for 45 min at 56 °C to reduce the protein.
8. Chill tubes to room temperature, spin down condensate. Remove excess liquid, and replace it quickly with roughly the same volume of freshly prepared alkylation solution. Incubate at room temperature for 30 min in the dark.
9. Remove alkylation solution and wash the gel particles with one gel volume of 100 mM NH<sub>4</sub>HCO<sub>3</sub> for 5 min.
10. Add an equal volume of ACN (to get 100 mM NH<sub>4</sub>HCO<sub>3</sub> / ACN, 1:1) and incubate for 15 min. (All the Coomassie stain should be removed at this time. The gel particles should appear completely transparent. If a large amount of protein is analysed (more than 10 pmol) remaining Coomassie staining may still be visible. In this case, an additional 100 mM NH<sub>4</sub>HCO<sub>3</sub> / ACN washing cycle should be performed).
11. Remove all liquid and dry gel particles in a vacuum centrifuge (~15 min).
12. Rehydrate gel particles by adding one gel volume of trypsin solution. Add more if all the initially added volume is soaked up. Incubate for 45 min.
13. Remove remaining enzyme supernatant and replace it with one gel volume of the digestion buffer (without enzyme).
14. Digest overnight at 37 °C and/or for 10min at approx. 170 W in the domestic microwave oven.
15. After overnight digestion pool the supernatant into a new tube.
16. Add one gel volume of ACN/ddH<sub>2</sub>O = 1:1 to the gel particles. Incubate for 15 min and pool the supernatant and transfer into the same tube as previous supernatant.
17. Add one gel volume of ACN/1% HCOOH = 1:1 to the gel particles. Incubate for 15 min and pool the supernatant and transfer into the same tube as previous supernatant.
18. Repeat the step once more.
19. Pool all the extracts.
20. Dry the sample in a vacuum centrifuge (~1-11/2h).
21. Redissolve peptides in 10 µL of 0.1 % TFA, sonicate briefly.



## Zip Tip purification

### **Materials needed:**

Pipetor with range: 0 - 10  $\mu$ l  
Zip Tips C18

### **Reagents:**

ddH<sub>2</sub>O, ACN, TFA, CHCA  
(All with high purity > 99.5%)

### **Solutions:**

0.1% TFA	(10 $\mu$ l TFA in 10 ml ddH <sub>2</sub> O)
0.1% TFA/ACN = 1:1	(500 $\mu$ l 0.1% TFA + 500 $\mu$ l ACN)
6 mg/ml CHCA in 0.1% TFA/ACN = 1:1	(3 mg CHCA + 250 $\mu$ l 0.1% TFA + 250 $\mu$ l ACN)

(Always prepare fresh solutions, except for the 0.1% TFA)

### **Procedure:**

1. Wet the Zip Tip with 8  $\mu$ l of 0.1% TFA/ACN = 1:1, by taking in and out the solvent in the same eppendorf tube, 3 times minimum.
2. Equilibrate the tips with 8  $\mu$ l of 0.1% TFA (eppendorf 1), by taking the solvent from the tube with 0.1% TFA and throwing away the liquid into the separate waste tube, 3 times minimum.
3. Bind the redissolved peptides by aspirating the solution with the Zip Tip, in and out, in the same tube, 3 times minimum.
4. Wash the Zip Tip with 8  $\mu$ l of 0.1% TFA (eppendorf 2), by taking the solvent from the tube with 0.1% TFA and throwing away the liquid into the separate waste tube, 3 times minimum.
5. Elute the peptides by aspirating matrix solution approximately 2-3  $\mu$ l and eluting it directly on the target plate. Suck in and out the solvent from the plate with the Zip Tip minimum 3 times before letting it dry.

## Sample Preparation for MALDI MS

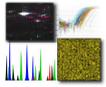
### **Sample Concentration**

Best results are obtained by loading high femtomole to low picomole quantities of sample onto the sample target.

Typically,

Peptides:	100 fmol - 5 pmol (on-target)
Proteins:	1-10 pmol

Approximately 0.5  $\mu$ l sample is loaded onto the sample target with 0.5  $\mu$ l of the appropriate matrix (>1000 - 10000 molar excess) and allowed to dry prior to analysis.



The technique is very sensitive to sub-picomole levels of sample and only a small fraction of the sample is actually consumed during analysis. However, sample handling becomes a problem at sub-picomole levels, so where possible, aim for a concentration of between 5 - 50 pmol/ $\mu$ L.

In addition, since very large proteins (>60 kDa) give weaker signals, it is better to aim for the upper part of the concentration range with these types of samples.

## Peptide and protein matrices

The most common matrices for protein and peptide analysis on MALDI MS are CHCA and SA. SA is used for protein analysis in Linear mode, while CHCA is used for peptide analysis in Reflectron mode.

1. Apply 0.5  $\mu$ L sample solution (typically 0.1-10 pmol, although smaller amounts can be detected, depending on the presence of buffers) to the MALDI target.
2. Add 0.5  $\mu$ L matrix solution (5 mg/mL in 50/50 acetonitrile/water) to the sample, and mix gently by pipetting up and down 3-4 times.
3. Allow the mixture to dry.

Greater tolerance to buffers can be achieved by increasing the SA concentration to saturating levels - approximately 50 mg/ml.

Alternative method for CHCA (6), analysis in positive and negative ion mode

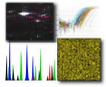
1. Apply 1  $\mu$ L of CHCA solution (10 mg/mL in 98/2 acetone/0.1% TFA) to the MALDI target and allow to dry.
2. Apply 1  $\mu$ L of sample (0.1 - 10 pmol in 0.1% TFA) on top of the dried matrix and allow to dry.

CHCA with citric acid (7) - 1 part CHCA (10 mg/mL in 50/50 ethanol/water) and 1 part citric acid (5 mM; 0.96 mg/mL in water) - for the analysis of phosphorylated, sulphated and glycosylated peptides in positive or negative ion mode.

DHB (8) - 10 mg/mL in 20/80 acetonitrile/water - for the analysis of peptides in positive ion mode.

DHBs (DHB/MSA) (9) - 9 parts DHB (10 mg/mL in 20/80 acetonitrile/water) and 1 part MSA (10 mg/mL in 50/50 ethanol/water) - for the analysis of proteins in positive ion mode.

1. Add 4  $\mu$ L matrix solution to an Eppendorf tube.
2. Add 1  $\mu$ L sample (0.5 - 50 pmol) and vortex.
3. Apply 1  $\mu$ L of the mixture to the MALDI target and allow to dry.



## Digest Calibration Mix Preparation

For calibrating PMF and PSD spectra, the procedure detailed in the AXIMA Reference Guide - Calibration, Reference Guide – MALDI-MS or Reference Guide – MALDI-sPSD should be followed. The calibrations standards recommended and the method of preparation of these samples are detailed below.

### 8 Peptide Calibration Mixture

Contains:

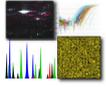
- Bradykinin (fragment 1-7) – RPPGFSP; monoisotopic [M+H] = 757.3997
- Angiotensin II – DRVYIHPF; monoisotopic [M+H] = 1046.5423
- Angiotensin I – DRVYIHPFHL; monoisotopic [M+H] = 1296.6853
- [Glu]1-fibrinopeptide – EGVNDNEEGFFSAR; monoisotopic [M+H] = 1570.6774
- N-acetyl Renin substrate – Ac-DRVYIHPFLLVYS; monoisotopic [M+H] = 1800.9437
- ACTH (fragment 1-17) – SYSMEHFRWGKPVGKKR; monoisotopic [M+H] = 2093.0867
- ACTH (fragment 18-39) – RPVKVYPNGAEDESAEAFPLEF; monoisotopic [M+H] = 2465.1989
- ACTH (fragment 7-38) – FRWGKPVGKKRRPVKYPNGAEDESAEAFPLE; monoisotopic [M+H] = 3657.9294

Start with 1 mg/mL or 1000 pmol/uL stock solutions for the individual peptides.

1. Add the necessary volumes required to add the following amounts of each peptide to the same sample tube:

Bradykinin (fragment 1-7)	11250 pmol
Angiotensin II	5000 pmol
Angiotensin I	5000 pmol
[Glu]1-fibrinopeptide	5000 pmol
N-acetyl Renin substrate	5000 pmol
ACTH (fragment 1-17)	5000 pmol
ACTH (fragment 18-39)	5000 pmol
ACTH (fragment 7-38)	7500 pmol

2. Make up the volume to 100 uL using 0.1% (v/v) TFA. This stock solution should be stored at -20 C and can be used for many months.
3. To use, add 1 µL of the stock solution prepared above to 74 µL 30/70 acetonitrile/0.1% TFA (total volume = 75 µL). This solution can be stored at room temperature and used for 1 week. A larger dilution may be required if you find the signal intensities are too high.



**Reference file:**

Bradykinin 1-7	[M+H] <sup>+</sup> = 757.3997
Angiotensin II	[M+H] <sup>+</sup> = 1046.5423
Angiotensin I	[M+H] <sup>+</sup> = 1296.6853
[Glu]1-fibrinopeptide B	[M+H] <sup>+</sup> = 1570.6774
N-acetyl Renin substrate	[M+H] <sup>+</sup> = 1800.9437
ACTH 1-17	[M+H] <sup>+</sup> = 2093.0867
ACTH 18-39	[M+H] <sup>+</sup> = 2465.1989
ACTH 7-38	[M+H] <sup>+</sup> = 3657.9294

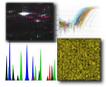
## Cleaning Target

1. Surface wiped with tissue soaked in methanol
2. Sonicate in acetone (15 min) – all cleaning steps target placed face down in cleaning solution
3. Rinse target with water
4. Sonicate (15 min) in 5% Micro90<sup>+</sup> prepared at 60 °C
5. Rinse and sonicate in water (5 min)
6. Sonicate in methanol (15 min)
7. Sonicate in water (15 min)
8. Rinse the target with methanol to remove residual water.
9. Allow target to dry overnight (preferably in vacuum). Typically, the visibly dry target is placed into the instrument overnight. The high vacuum removes any residual solvent resulting in a target which should not cause sample spreading during spotting. NB. only insert a dry target into the instrument.

**Note:**

An additional 1% formic acid wash (sonicate for 10-15 minutes) may be performed between steps 2. and 3. for new targets or targets which are heavily contaminated. Discard solvent to acid waste bottle following washing.

For most applications, steps 1, 2, 6-8 in the above protocol will be sufficient for target cleaning to remove samples. However, if evidence of carryover is observed, the additional Micro90 cleaning steps can be included.



## MALDI-MS Analysis using an AXIMA Series Mass Spectrometer

1. In the Experimental Technique (Exp. Tech.) tab of the Acquisition window (see figure 1), click Open Door. At the prompt click Yes to open the stage door. When the stage door opens, insert the MALDI target (face up with the cut-off corner to the front). Push the target until the back edge is flush with the stage. In the Experimental Technique (Exp. Tech.) tab of the Acquisition window (see figure 1), click Close Door to close the stage door. Select the correct ‘Tuning mode’. For samples < 10,000 amu, use the appropriate reflectron mode. For samples > 10,000 amu, use the appropriate linear mode. Ensure the instrument is in Operate and has been for at least 20 min prior to starting the acquisition. Set the appropriate mass range (the lower mass in the range should be set to 1.0). Check that the ‘Max. Laser Rep. Rate’ is 50.

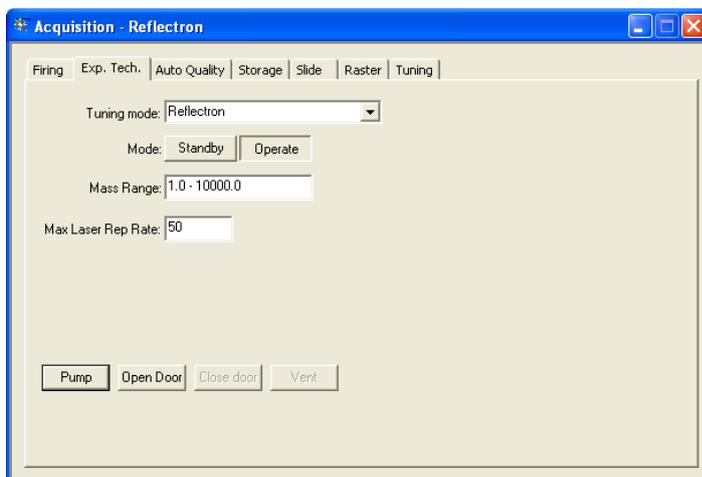


Figure 1: Experimental technique (Exp. Tech.) in the acquisition window (MALDI-MS software).

2. In the Firing tab (see figure 2), set the (laser) Power. The laser power required will depend on the type of matrix being used and the individual instrument. When using  $\alpha$ -cyano 4-hydroxycinnamic acid as matrix, a typical starting laser power is 30-50. The laser power can be optimised during the acquisition (see step 5). Set the number of Profiles to 1000 (this will ensure the acquisition will not stop automatically before satisfactory data has been acquired).

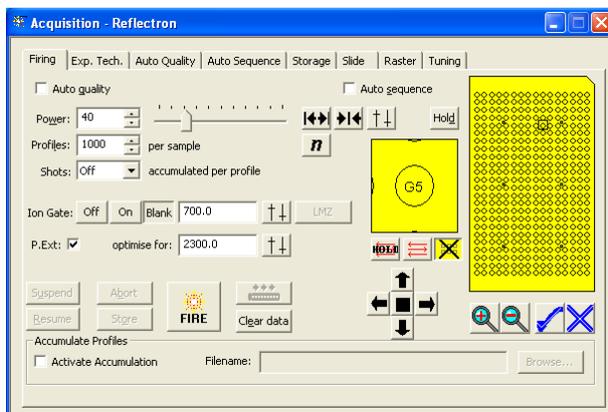
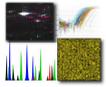


Figure 2: Firing tab in the acquisition window (MALDI-MS software).



3. Set the number of Shots (accumulated per profile) to 10-50. This is equivalent to 10-50 shots accumulated per profile. [Setting the number of shots per profile <10 will result in low repetition rates being achieved during acquisition].
4. Set the ion gate to Blank. Set the low mass cut-off for the ion gate (for peptide mass fingerprinting (PMF) experiments using peptides, this is typically 700 or 800).
5. Check/tick the (pulsed extraction) P.Ext option box. In the P.Ext. field, enter the mass of interest corresponding to the sample to be analysed. Instrument settings will automatically optimised to produce optimum sensitivity and resolution for the specified mass. For a sample containing a mixture of components (e.g. a protein digest), set the P.Ext. to approximately 2/3 of the maximum mass of interest. For the analysis of protein digest mixtures, this is typically set to 2300.
6. Ensure the correct sample plate file is selected corresponding to the MALDI target being used. In the Slide tab (see figure 3), the filename of the plate file currently loaded is displayed in the ‘Sample plate file’ field. If the filename does not match that on the bottom of the MALDI target being used, click Load... and select the correct plate file. When the correct plate file has been selected, click Apply.

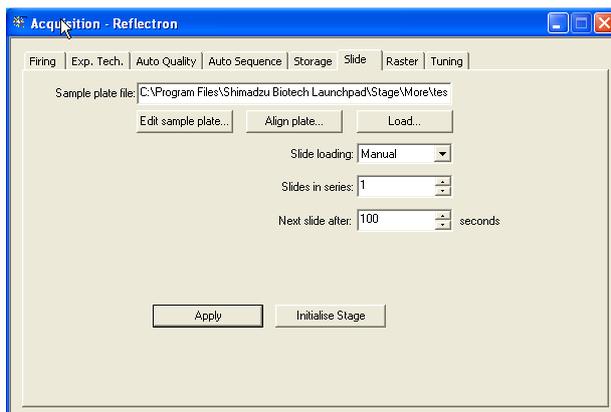
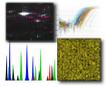


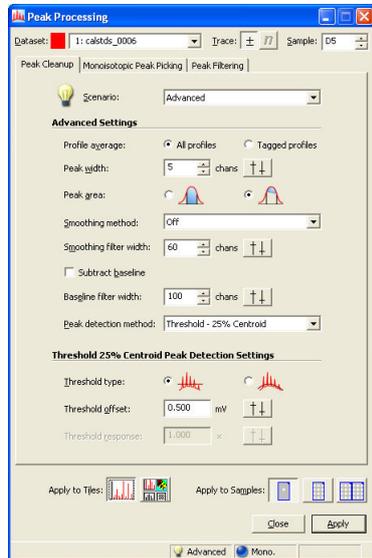
Figure 3: Slide tab in the acquisition window (MALDI-MS software).

7. Move the target position to that of the calibration standards. To do this, right click on the yellow target in the Acquisition window and select ‘Goto location...’. Type in the position of the calibration standards spot in the format row letter/column no. (e.g. A1).
8. Check that the appropriate peak processing parameters are applied (MALDI-MS main window: Processing » Peak Processing... (see figure 4)). Typically, no ‘Smoothing’ (OFF) or ‘Baseline Subtraction’ is used. The Peak detection method used is typically ‘Threshold – 25% centroid’ (peak width: 5-10 channels). Select the Threshold type (fixed or adaptive) and set the ‘threshold offset’ and ‘threshold reponse’ as required. Adjust the threshold parameters if necessary during the acquisition. In the monoisotopic peak picking tab, check/tick the use monoisotopic peak picking box. In the monoisotopic peak picking tab, ensure the mass range, minimum isotopes, maximum intensity variation and minimum peak percent are set correctly. The monoisotopic peak picking parameters and threshold may need adjusting during the acquisition in order to label all the peaks correctly. If



any changes have been made to the peak processing parameters, ensure the appropriate ‘Tile(s)’ and ‘Sample(s)’ options are selected click Apply.

(a)



(b)

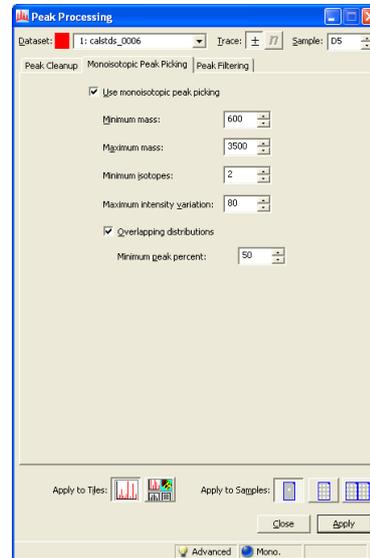
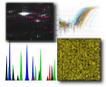


Figure 4: Peak processing window (MALDI-MS software) showing (a) Peak Cleanup and (b) Monoisotopic Peak Picking parameters.

9. In the Acquisition window, click the FIRE button to start the acquisition of the calibration standards. During the acquisition, check the resolution of the highest mass standard peak. To check the resolution during an acquisition, click Suspend to pause the acquisition (when the resolution has been measured, click Resume to continue the acquisition). Place the cursor over the labelled peak you wish to check the resolution of and click the middle mouse button. The resolution (M/dM) is displayed in the bottom left hand corner of the MALDI-MS main window. [The mass at which the resolution is calculated is also shown in parentheses]. For the peptide standard ACTH18-39 (m/z 2465.20), the resolution should be >11,000. The resolution will decrease at lower masses. If the resolution is much lower, try reducing the laser power to improve resolution. During the acquisition, ensure that signal intensity for each profile does not exceed 2000 mV (the signal intensity is shown next to % Int. of the spectrum being acquired (upper panel)). If necessary, move the target using the arrows in the Acquisition window to find a position at which a stable signal can be obtained. Use the Clear data button to restart the accumulation once a satisfactory position has been found on the sample spot or following optimisation of the laser power.

10. Check the name of the calibration file being used to acquire the data (refer to the Calibration using MALDI-MS Software reference guide). If the Default calibration (Factory Calibration) is being used, click Suspend to pause the acquisition and load a different calibration file which is suitable for the mode being used (i.e. reflectron or linear; positive or negative mode) (refer to the Calibration using MALDI-MS Software Reference Guide).



11. Once a satisfactory spectrum has been acquired, click Abort to stop the acquisition. Click Store to save the spectrum.
12. In the Calibration window (MALDI-MS main window: Processing » Calibration...), create an instrument calibration file (refer to the Calibration using MALDI-MS Software reference guide). Ensure that this instrument calibration file is loaded and is used to acquire the sample spectrum in step 13.
13. Move to the sample spot position as described in step 7. Acquire the sample spectrum (as described in step 9.) using the same P.Ext value and peak processing parameters as those used to acquire the calibration standards. During the acquisition, continually check the resolution of the higher mass peaks and the signal intensity of the base peak. Once a satisfactory spectrum has been obtained, click Abort and Store to stop and save the spectrum.

## MALDI-MS/MS Analysis using an AXIMA Series Mass Spectrometer

1. In the Experimental Technique (Exp. Tech.) tab of the Acquisition window, click Open Door. At the prompt click Yes to open the stage door. When the stage door opens, insert the MALDI target (face up with the cut-off corner to the front). Push the target until the back edge is flush with the stage. In the Experimental Technique (Exp. Tech.) tab of the Acquisition window, click Close Door to close the stage door. Select the appropriate ‘reflectron’ tuning mode. Ensure the instrument is in Operate and has been for at least 20 min prior to starting the acquisition. Set the appropriate mass range. Check that the ‘Max. Laser Rep. Rate’ is 50.
2. In the Firing tab (see figure 1), set the (laser) Power. The laser power required will depend on the type of matrix being used and the individual instrument. When using  $\alpha$ -cyano 4-hydroxycinnamic acid as matrix, a typical starting laser power is 30-50. The laser power can be optimised during the acquisition (see step 9). Set the number of Profiles to 1000 (this will ensure the acquisition will not stop automatically before satisfactory data has been acquired).

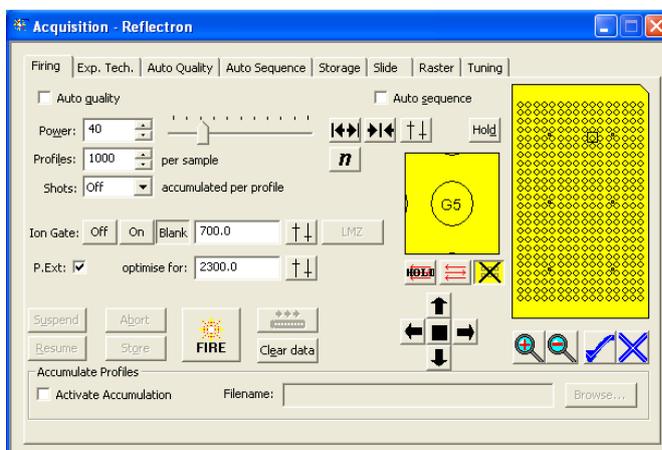
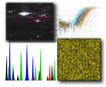
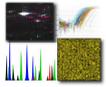


Figure 1: Firing tab in the acquisition window (MALDI-MS software).



3. Set the number of Shots (accumulated per profile) to 10 - 50. [Setting the number of shots per profile <10 will result in low repetition rates being achieved during acquisition].
4. Set the ion gate to OFF.
5. Check/tick the P.Ext option box. In the P.Ext. field, enter the mass of interest corresponding to the sample to be analysed. Instrument settings will be automatically optimised to produce optimum sensitivity and resolution for the specified mass. For a sample containing a mixture of components (e.g. a protein digest), set the P.Ext. to approximately 2/3 of the maximum mass of interest. For the analysis of protein digest mixtures, this is typically set to 2300.
6. Ensure the correct sample plate file is selected corresponding to the MALDI target being used. In the Slide tab of the Acquisition window, the filename of the plate file currently loaded is displayed in the 'Sample plate file' field. If the filename does not match that on the bottom of the MALDI target being used, click Load... and select the correct plate file. When the correct plate file has been selected, click Apply.
7. Move the target position to that of the calibration standards. To do this, right click on the yellow target in the Acquisition window and select 'Goto location...'. Type in the position of the calibration standards spot in the format row letter/column no. (e.g. A1).
8. Check that the appropriate peak processing parameters are applied (MALDI-MS main window: Processing » Peak Processing...). Typically, 'Average' smoothing method (width: 5-10 ch) is used. The Peak detection method used is 'Threshold – Apex' (width: 5 ch). Set the threshold type (fixed or adaptive) and set the threshold offset and threshold response and adjust if necessary during the acquisition. These peak processing parameters are only to be used as a guide and may need optimising depending on the quality of the spectrum obtained. In the Monoisotopic Peak Picking tab, check/tick the use monoisotopic peak picking box. In the Monoisotopic Peak Picking tab, set the 'Minimum mass' to 150 and the 'Minimum isotopes' to 2. If any changes have been made to the peak processing parameters, ensure the appropriate 'Tile(s)' and 'Sample(s)' options are selected click Apply. The threshold offset and/or threshold response may need adjusting during the acquisition in order to label all the peaks correctly.
9. Click the FIRE button (Acquisition window) to start the acquisition of the calibration standards. During the acquisition, check the resolution of the highest mass standard peak. To check the resolution during an acquisition, click Suspend to pause the acquisition (when the resolution has been measured, click Resume to continue the acquisition). Place the cursor over the labelled peak you wish to check the resolution of and click the middle mouse button. The resolution (M/dM) is displayed in the bottom left hand corner of the MALDI-MS main window. [The mass at which the resolution is calculated is also shown in parentheses]. For the peptide standard ACTH18-39 (m/z 2465.20), the resolution should be ~5000-6000 (for the smoothed data). The resolution will decrease at lower masses. If the resolution is much lower, try reducing the laser power to improve resolution. During the acquisition, ensure that signal intensity for each profile does not exceed 2000 mV (the signal intensity is shown next to % Int. of the spectrum being acquired (upper panel)). If necessary, move the target using the arrows in the Acquisition window to find a position at which a stable



signal can be obtained. Use the Clear data button to restart the accumulation once a satisfactory position has been found on the sample spot or following optimisation of the laser power.

10. Check the name of the calibration file being used to acquire the data (refer to the Calibration using MALDI-MS Software Reference Guide). If the Default calibration (Factory Calibration) is being used, click Suspend to pause the acquisition and load a suitable reflectron mode calibration file (refer to the Calibration using MALDI-MS Software Reference Guide).

11. Once a satisfactory spectrum has been acquired, click Abort to stop the acquisition. Click Store to save the spectrum.

12. In the Calibration window (MALDI-MS main window: Processing » Calibration...), create an instrument calibration file (refer to the Calibration using MALDI-MS Software reference guide). When generating an instrument calibration file for MALDI-MS/MS acquisitions, a low mass matrix peak should also be included in the reference file to extend the calibration range to lower masses so that the fragment ion masses are calibrated correctly. Ensure that this instrument calibration file is loaded and is used to acquire the sample spectrum in step 13.

13. Move to the sample spot position as described in step 7. Set the Ion Gate to Blank. Acquire the sample spectrum (as described in step 9.) using the same P.Ext value as that used to acquire the calibration standards. Acquire ~50 profiles.

14. Select the precursor for MALDI-MS/MS. Place the cursor at a mass slightly above the precursor mass. Click the middle mouse button (wheel) to set a cursor on the spectrum. Repeat, placing the cursor at a mass slightly below the precursor mass. The precursor ion should now be centrally gated by two peak cursors (see figure 2). Check the mass difference (dM) between the two cursors and resolution (M/dM) of the ion gate (this is shown in the bottom right hand corner of the MALDI-MS main screen). The maximum resolution of the ion gate is ~400 for the AXIMA-Performance and ~200 for the AXIMA-Confidence. This is equivalent to a dM of 2 at 1000 m/z. When selecting a precursor for MALDI-MS/MS, use the widest ion gate setting possible to isolate the precursor of interest whilst at the same time ensuring that other unrelated species are not transmitted. Using the widest ion gate setting possible will help to maximise sensitivity for the selected precursor. However, if other peaks are co-transmitted with the precursor of interest, reduce the ion gate width.

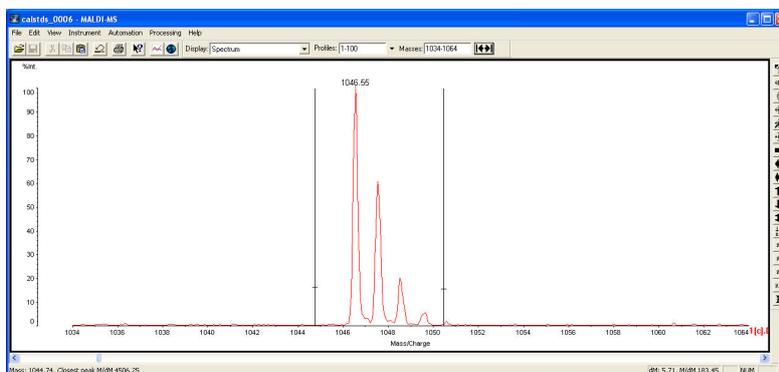
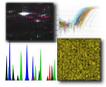


Figure 2: Setting the ion gate using the peak cursors.



15. In the Acquisition window, set the Ion Gate to ON. Click the button to update the Ion Gate field. Click FIRE. Ensure the isotopic profile of the selected precursor is the same as that with the Ion Gate set to Blank obtained in step 13. If the isotopic profile is significantly different, adjust the Ion Gate accordingly.

16. Insert a second column into the MALDI-MS main window (right click on the spectrum and select Insert » Column). In one column, zoom in on the region around the precursor. In the other column, zoom in on the region containing the fragment ions i.e.  $\sim 100$  m/z up to (but not including) the precursor ion (see figure 3). If using an AXIMA-Confidence, go to step 18. For AXIMA-Performance instruments, go to step 17.

17. In the Experimental Technique (Exp. Tech.) tab, click the CID Enable button to switch on the CID gas flow. Click OK at the message prompt. The instrument switches off all the voltages applied to the various lenses and temporarily switches off the ion gauge connected to the CID cell. In the Instrument Status window (View » Instrument Status...), ensure the pressure gauge reading the pressure at the collision cell re-stabilises at  $\sim 5 \times 10^{-6}$  mbar. If this pressure is not achieved, check that the supply of helium to the instrument (e.g. cylinder) is sufficient. After  $\sim 30$  sec, all voltages are reapplied and the instrument returns to the 'operate' state.

18. Increase the laser power to induce fragmentation. For MALDI-MS/MS analysis of peptides, increase the laser power by  $\sim 20\%$  above that used to acquire the MALDI-MS spectrum. Click FIRE. During the acquisition, continually check the precursor ion intensity and resolution of the fragment ions. It may be necessary to move position during the acquisition to maintain a high precursor ion signal as a result of sample ablation due to the higher laser power used during MALDI-MS/MS. The resolution of the fragment ions will be much lower than that obtained during MS-mode but it should still be possible to resolve the monoisotopic peak for some of the fragment ions. If the resolution is completely degraded, reduce the laser power slightly to try and improve this. If no fragment ions are detected, try increasing the laser power. Increasing the laser power will increase the level of the background noise and decrease resolution. Acquire a minimum of 200 profiles.

19. Once a satisfactory spectrum has been obtained, click Abort to stop the acquisition. Click Store to save the spectrum.

N.B. if the ion gate was set by manually typing in a mass range in the ion gate field (i.e. not as described in steps 14. and 15.), on clicking Abort, the fragment ions will shift towards lower mass as the fragment fit correction is applied. This does not occur if the ion gate is set as described above as the fragment fit correction is applied in real-time. The value used for the fragment fit correction is taken from the precursor mass in the spectrum used to set the ion gate using the cursors.

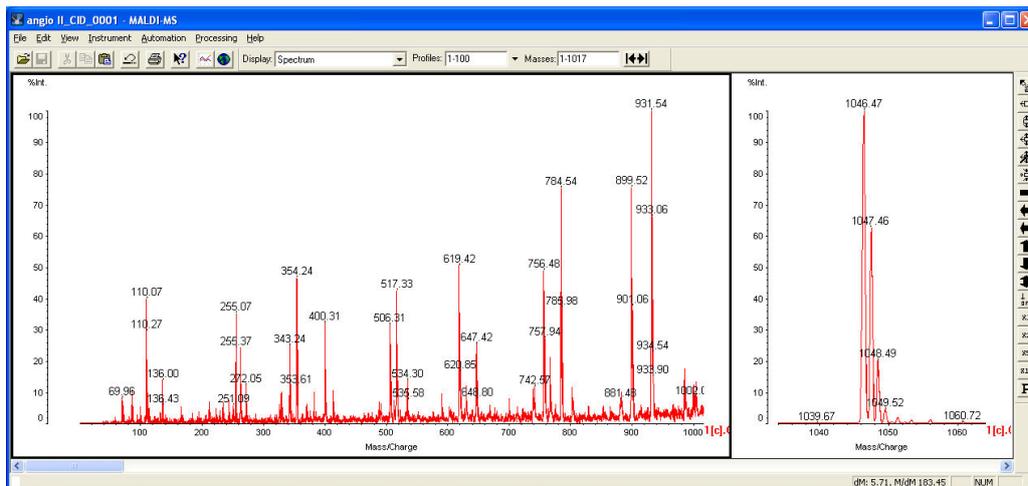
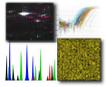


Figure 3: Spectrum view following MALDI-MS/MS acquisition. Left hand column: mass range showing fragment ions; Right hand column: mass range showing precursor ion.

## Calibration using MALDI-MS Software

1. Acquire a MALDI-MS spectrum of the calibration standards (refer to the MALDI-MS Analysis using an AXIMA Series Mass Spectrometer Reference Guide) using the same Pulsed Extraction (P.Ext) value as that which is to be used to acquire the sample spectrum. Click Store to save the spectrum. Ensure that all the peaks corresponding to calibration standards are labelled. If one or more of the peaks corresponding to calibration standards are not labelled, adjust the 'Threshold' in the Peak Processing parameters window (MALDI-MS main window: Processing » Peak Processing...).
2. In the Calibration window (see figure 1), click List references... and select the reference file corresponding to the calibration standards used in step 1. If a suitable reference file does not exist, use Reference editor... to create a new one (see Appendix). The references (typically, calculated  $[M+H]^+$  (for positive-ion mode) or  $[M-H]^-$  (for negative-ion mode) values for the different calibration standards) of the selected reference file are displayed in the 'Calibrant references' panel of the Calibration window (left hand side half of the Calibration window). In the example shown (figure 1), the reference file selected contains references for 10 standards. Displayed references which are not required can be removed from the Calibration window by selecting the corresponding reference and clicking Delete.

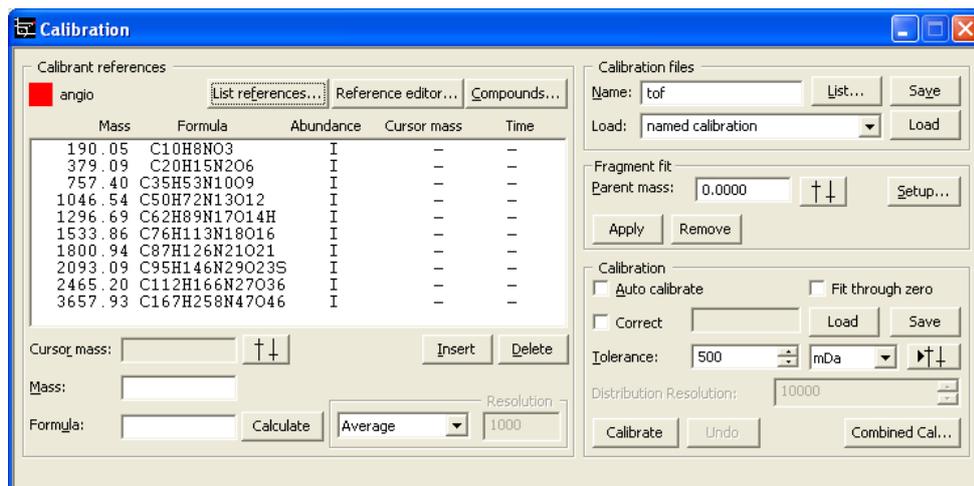
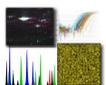


Figure 1: Calibration window (MALDI-MS software).

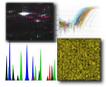
3. Set the 'Tolerance' (bottom right hand panel) to 500 mDa. Ensure 'Auto calibrate', 'Fit through zero', 'Correct' and 'Fragment fit' boxes are unchecked/unticked.

4. Click Calibrate. The 'Time' column in the 'Calibrant references' panel for each reference should contain a number (not zero). If the value in the 'Time' column is zero, a letter or – , press Calibrate again until a number appears for each reference.

If it is not possible to obtain a number in the 'Time' column for one or more of the standards or an error message is obtained (e.g. Error 24030: Calibration failed), it may be necessary to perform a manual calibration. To do this, double click on the first reference in the 'Calibrant references' panel. In the spectrum obtained for the calibration standards, place the cursor on the peak corresponding to the reference and click the middle mouse button. A cursor mark should appear on the peak. In the Calibration window, click the  button to update the 'Cursor mass' field. Repeat for all the references in the 'Calibrant references' panel so that each reference has a corresponding cursor mass. Click Calibrate twice. The 'Time' column in the 'Calibrant references' panel for each reference should contain a number (not zero).

5. To save the instrument calibration file, enter a filename in the 'Name' field (upper right hand panel of the Calibration window). Recheck/tick the 'Correct' option box and click Save. To load an instrument calibration i.e. so it can be used to calibrate data acquired in a subsequent acquisition, click List... and select the instrument calibration file from the list. Click Load to load the instrument calibration file. The instrument calibration file loaded in the 'Name' field will be used to calibrate data as it is acquired when the FIRE button is clicked.

To create a new reference file or to modify an existing file, click Reference editor... in the calibration window or select Reference editor... from the MALDI-MS main window (Processing » Reference Editor...). To create a new reference file, enter the empirical formula (typically [M+H]<sup>+</sup> or [M-H]<sup>-</sup>) of the standard. Alternatively, enter the calculated mass (monoisotopic or average; [M+H]<sup>+</sup> or [M-H]<sup>-</sup>) and a name for the standard. [The decision whether to enter a monoisotopic or average mass will depend on the type of experiment which is to be performed. For analyses where isotopic resolution can be achieved (e.g. reflectron mode), monoisotopic values should be used for the standards.]



Select whether the mass entered is Average, Monoisotopic or Most Abundant. Click Insert to update the ‘References’ field with the new standard information. Repeat if necessary until the reference file contains all the required standards. Enter a filename for the reference and click Save.

To modify an existing reference file, click List and select the reference file to be modified. Use the Insert or Delete buttons to update or remove standards from the reference file. To save the modified reference file, click Save.

## Database Searching (MALDI-PMF) using MALDI-MS Software

1. Acquire a MALDI-MS spectrum of the sample (refer to the MALDI-MS Analysis using an AXIMA Series Mass Spectrometer Reference Guide).
2. Check that the appropriate peak processing parameters are applied (MALDI-MS main window: Processing » Peak Processing...). Typically, no ‘Smoothing’ (OFF) or ‘Baseline Subtraction’ is used. The Peak detection method used is typically ‘Threshold – 25% centroid’ (width: 5 channels) and the ‘Threshold type’ is fixed. In the Monoisotopic peak picking tab, check/tick the Use monoisotopic peak picking box. In the Monoisotopic Peak Picking tab, ensure the mass range, minimum isotopes, maximum intensity variation and minimum peak percent are set correctly (typical values are shown in figure 1). If any changes have been made to the peak processing parameters, ensure the appropriate ‘Tile(s)’ and ‘Sample(s)’ options are selected click Apply. Check that all peptide peaks are correctly labelled. Peaks that are not labelled or are labelled incorrectly should be manually labelled using Manual Peak Assignment described in step 3.

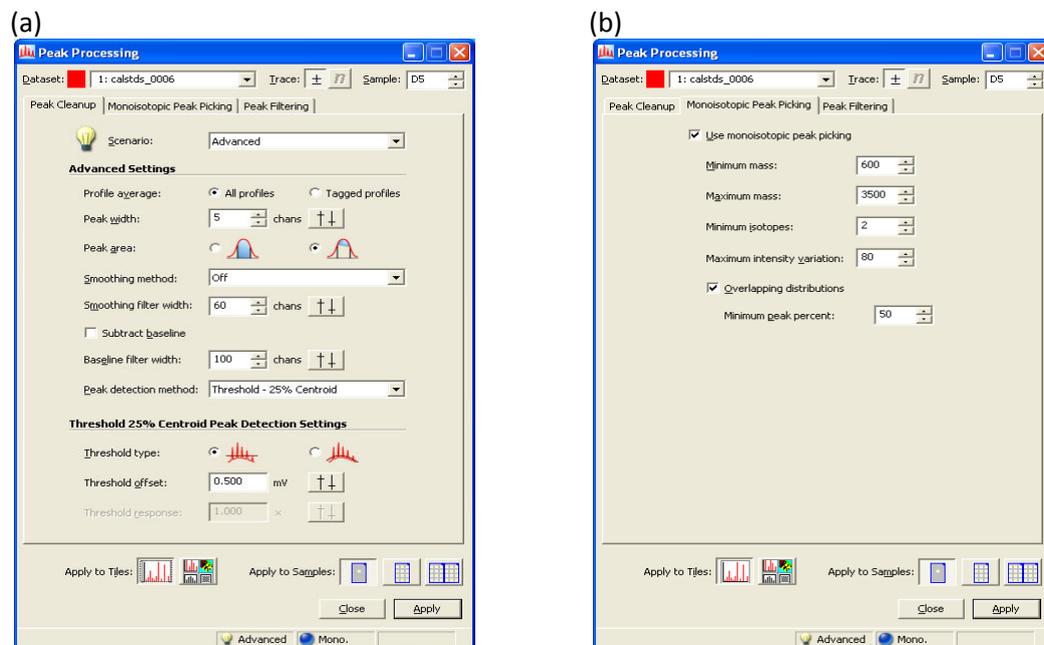
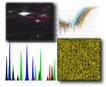


Figure 1: Peak processing window (MALDI-MS software) showing (a) Peak Cleanup and (b) Monoisotopic Peak Picking parameters.



3. Peaks that are not labelled or are labelled incorrectly should be manually labelled using the Manual Peak Assignment function (MALDI-MS main window: Processing » Manual Peak Assignment) (see figure 2).

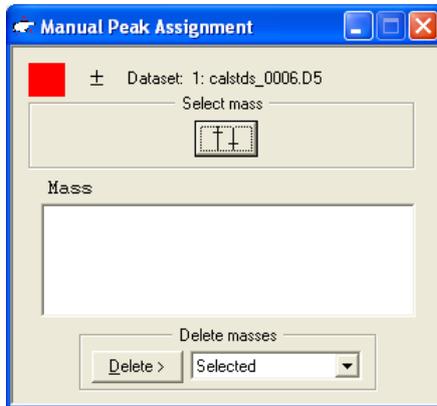


Figure 2: Manual Peak Assignment window (MALDI-MS software).

To manually label a peak, place the cursor on one side of the isotope peak to be labelled and press the middle mouse button to insert a peak cursor. Place a second peak cursor on the other side of the isotope peak. The peak to be labelled should now have two peak cursors either side of it. Click the button. The mass of the peak should appear in the ‘Mass’ field of the Manual Peak Assignment window (figure 2) and the peak label may also appear on the peak in the spectrum.

4. Once all the peaks are labelled correctly, open the Internet search window by clicking the button on the MALDI-MS tool bar (see figure 3).

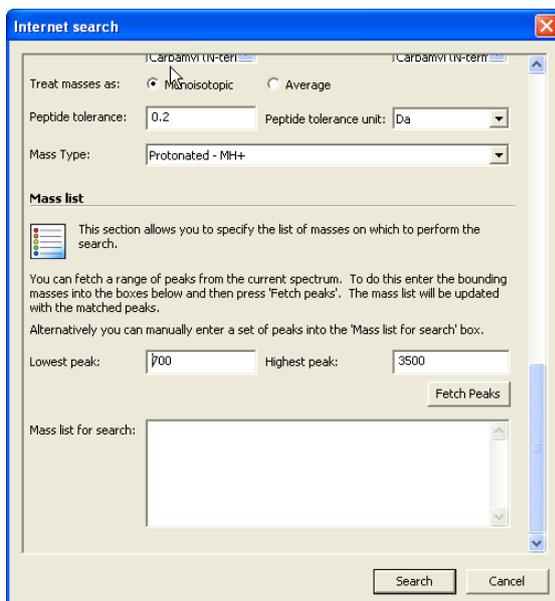
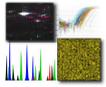


Figure 3: Internet search window (MALDI-MS software).

5. In the ‘Search engine’ field, select Mascot PMF. In the ‘Search URL’ field, type in the URL of the Mascot search engine to be used for the searches. [The searches can either be run directly from



the ‘Internet search’ window or by submitting the peak list generated into the Mascot search page i.e. from the Mascot home page in Internet Explorer (<http://www.matrixscience.com/cgi/nph-mascot.exe?1>). However, a valid URL must be entered so that the software is directed to a version of Mascot]. If the ‘Search URL’ is changed, click the Refresh button. The ‘Search URL’ field contains a valid URL when the databases can be viewed (white).

6. Complete the various fields (e.g. for a tryptic digest sample, set the enzyme as Trypsin, the number of missed cleavages = 1, peptide tolerance = 0.2 Da, mass values = MH+ and monoisotopic). In the ‘Mass list’ field (at the bottom of the Internet Search window), set the mass range to be used to generate the mass list and click Fetch peaks. The peptide masses are automatically entered into the ‘Mass list’ field and the Select Peaks window. The number of peaks and mass range can be modified in the Select Peaks window. If any changes are made, click List peaks to update the peak list. When all the parameters have been completed, click Search at the bottom of the Internet search window or copy and paste the list into the ‘Mascot Peptide Mass Fingerprint’ search page in Internet Explorer.

## Database Searching (MALDI-MS/MS data) using MALDI-MS Software

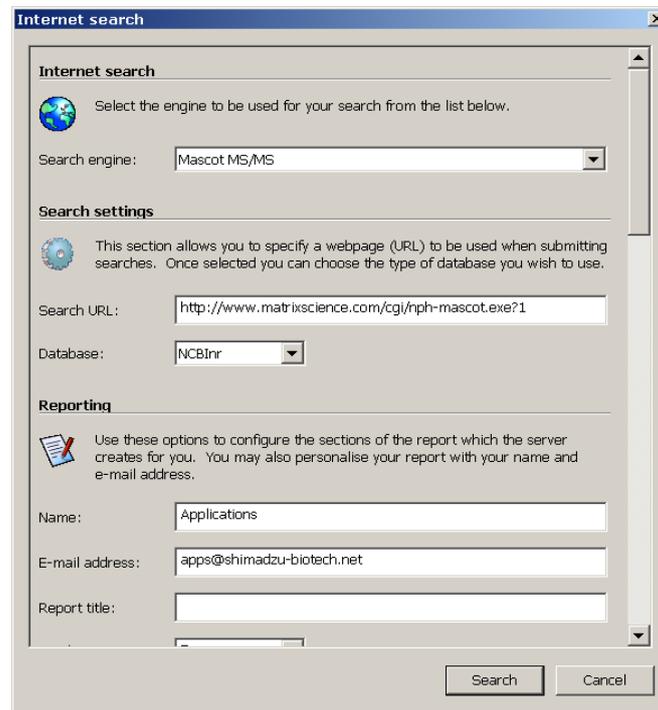
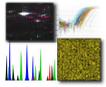
1. Acquire a MALDI-MS/MS spectrum of the sample (refer to the *MALDI-MS/MS Analysis using an AXIMA Series Mass Spectrometer Reference Guide*).

2. Check that the appropriate peak processing parameters are applied (*MALDI-MS* main window: *Processing* » *Peak Processing* » *Advanced view...*). Set the smoothing method and baseline subtraction to be used. Typically, ‘average’ smoothing (width: 50 ch) and ‘Subtract baseline’ (width: 50-150 ch) are used. Set the **Peak detection method** to ‘Threshold – Apex’ (width: 10 ch). Set the

threshold type to ‘Adaptive threshold’ i.e.  and set the ‘Threshold offset’ and ‘Threshold response’. Typically, the ‘Threshold offset’ is set to 0 mV and the ‘Threshold response’ is set to 1.5. It is possible to visualise the baseline to help achieve the correct threshold settings. [To visualise the baseline, in the *MALDI-MS* main window, select the *Options...* window (*View* » *Options*) and select the *Graphs* tab. In the *Graphs* tab, select/tick ‘Threshold’ and click **Apply**]. Adjust the ‘Threshold offset’ and ‘Threshold response’ until the level of the baseline is just above the level of the noise and that only real fragment ion peaks are above the baseline and are labeled.

3. In the *Monoisotopic Peak Picking* tab of the *Peak processing* window, ensure that the ‘Use monoisotopic peak picking’ option is not selected/unchecked. If any changes are made to the peak processing parameters, select which tiles and samples the processing is to be applied to and click **Apply**.

4. To create a Mascot generic format peak list, click the ‘Internet Search’ button  on the toolbar in the *MALDI-MS* main window (see figure 1).

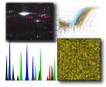


**Figure 1: Internet search window (MALDI-MS software).**

In the ‘Search engine’ field, select *Mascot MS/MS*. In the ‘Search URL’ field, type in the URL of the Mascot search engine to be used for the searches. [The searches can either be run directly from the ‘Internet search’ window or by submitting the peak list generated into the Mascot search page i.e. from the Mascot home page in Internet Explorer (<http://www.matrixscience.com/cgi/nph-mascot.exe?1>). However, a valid URL must be entered so that the software is directed to a version of Mascot]. If the ‘Search URL’ is changed, click the **Refresh** button. The ‘Search URL’ field contains a valid URL when the databases can be viewed (white).

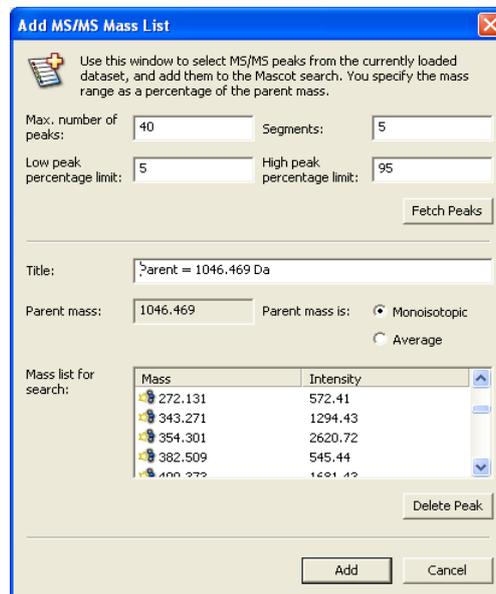
If the searches are to be run directly from the ‘Internet search’ window, complete the various search criteria fields according to the experimental conditions. Set whether the fragment masses are to be searched as *monoisotopic* or *average* (this will depend on the peak processing parameters used in 2.). Using the peak processing parameters described in step 2., fragment ion masses should be treated as *average*. Typical search tolerances are 0.5 Da (peptide tolerance) and 0.8 Da (MS/MS tolerance). Set the instrument type to ‘MALDI-TOF-TOF’.

5. To generate the peak list, at the bottom of the ‘Internet search’ window, click the **Add parent** button. Enter the maximum number of fragment ions to be included in the peak list. Enter the number of segments (typically 5 or 6) and the low- (typically 5%) and high- (typically 95%) percentage peak limits. The software only considers fragment ions within the low- and high-percentage peak limits. To facilitate peak picking, the mass range defined by the limits is divided into



*segments* and the total number of fragment ions requested is divided by the number of segments specified. When all parameters have been completed, click **Fetch peaks**.

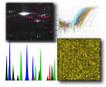
The fragment ions are listed as mass and intensity pairs at the bottom of the *Add MS/MS mass list* window (see figure 2).



**Figure 2: MS/MS peak list generation using MALDI-MS software.**

6. The ‘Title’ and ‘Parent mass’ fields show the precursor m/z which is taken from the mass used to apply the fragment fit correction. To the right of the parent mass, the user is asked to specify whether the precursor mass shown is a monoisotopic or average value. If the user has specified that the fragment ions are to be treated as average masses in step 4. and specifies that the parent mass corresponds to the monoisotopic value, the software converts the monoisotopic precursor value to an average value. Click **Add** to create the peak list and return to the ‘Internet search’ window.

7. The peak list can be manually modified in the ‘Internet search’ window by expanding the list by clicking + . Fragment ions to be detected can be highlighted and removed from the list by clicking **Delete Peak**. The final peak list can either be searched directly (by clicking **Search**) or saved as a text file (by clicking **Save List as...**) for searching at a later date.



## Overview of SARAMIS

SARAMIS is a software and database tool for the rapid automated identification of microorganisms on the basis of their characteristic MALDI-TOF fingerprint mass spectra. These mass spectra are compared to the SARAMIS database's SuperSpectra™ and spectra. SARAMIS is started from the Windows taskbar by selecting "Start -> Programs -> SARAMIS -> SARAMIS Premium". Alternatively, start the program by double-clicking the SARAMIS Premium shortcut icon on the computer's desktop.

When the program starts the "Presettings" window is automatically displayed.

Parameter	Value
mass accuracy:	0,08 %
graphical x-min:	2000
graphical x-max:	20000
massrange min:	3000
massrange max:	20000
exclusion list:	Sheep blood AGAR
identify-limit:	750
identify-count:	3

Quality of Spectra

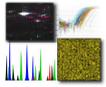
Parameter	Value
comparerange min.:	50
comparerange max.:	250

Parameter	Value
count of comparing results:	30

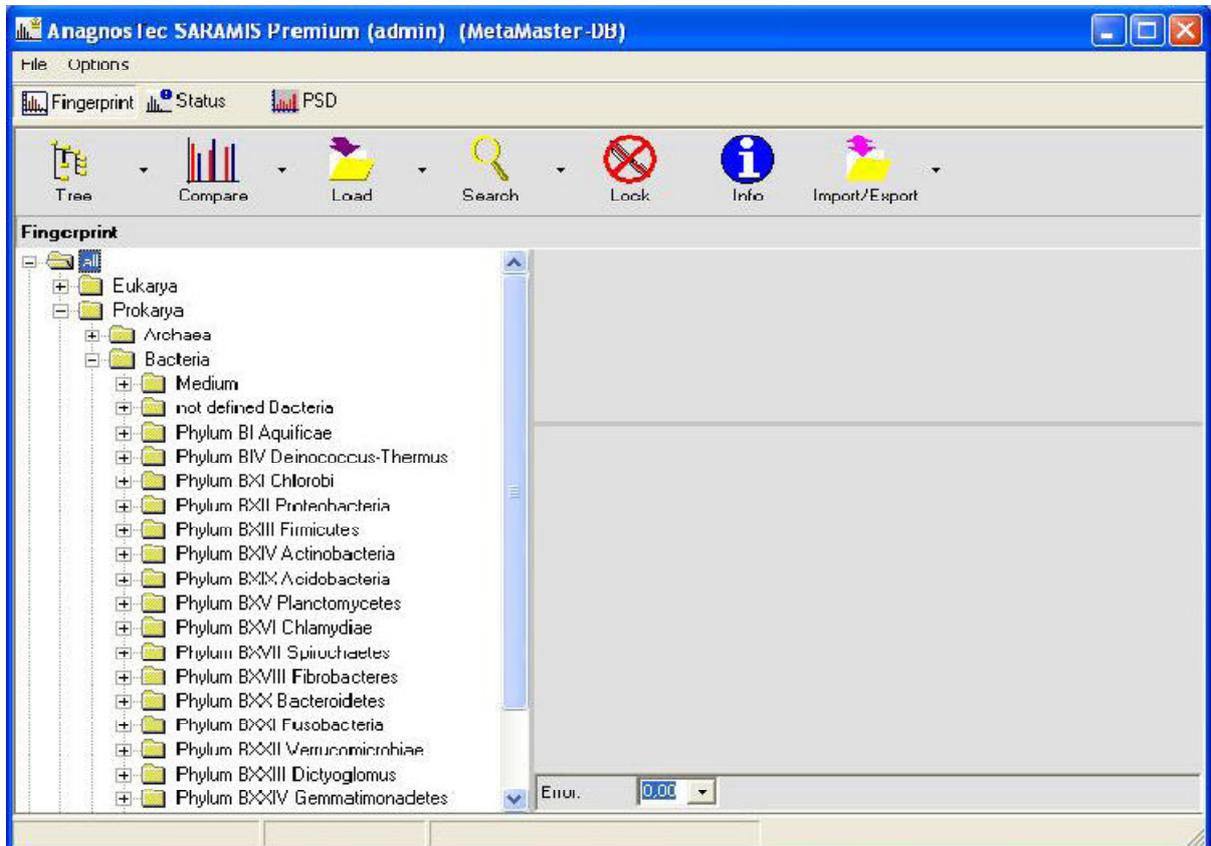
Parameter	Value
loading folder of files:	D:\AnagnosTec\Data\Files\ascii\
Folder for identified files:	D:\AnagnosTec\Data\Identified\
Folder for unidentified files:	D:\AnagnosTec\Data\Unidentified\
folder for unknown files:	D:\AnagnosTec\Data\Unknown\

Parameter	Value
savepath to LIMS:	D:\AnagnosTec\Data\LimsExport\

Close

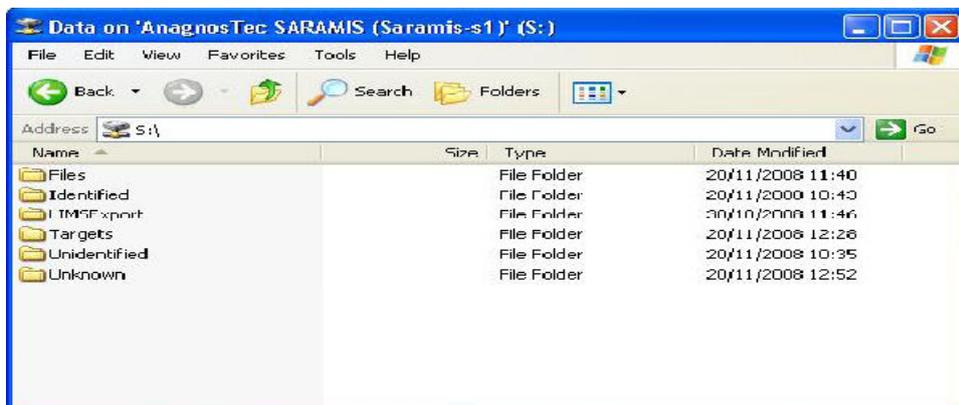


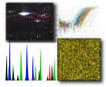
The SARAMIS Premium main application window opens. In the Fingerprint mode a split window is displayed, consisting of three sections: the Tree, the Compare/Search and the Results sections:



After the MALDI-TOF mass spectrometer has generated a spectrum the raw data is saved. The instrument's software processes the raw data, applying a peak detection algorithm with settings for minimal peak height and area, intensity threshold values and exclusion lists optimized for the downstream data handling. This results in a peak list that contains only individual masses and their intensities. The peak list is saved as a text file (\*.txt).

Immediately after a MALDI-TOF MS spectrum is recorded, SARAMIS can proceed with automatic identification. A peak list of each spectrum is automatically created, saved as a text file and transferred into the folder “\*\files” on the SARAMIS server.

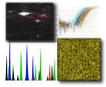




After the successful database comparison the window “Identify Details” opens and displays SARAMIS identification results.

Depending on the points that a spectrum scores for its similarity with a SuperSpectrumTM and on the user’s software settings (see 3.1), the line in the results table is displayed in either dark green, light green, yellow, white or red. The red color signifies that a spectrum has received the user-preset minimum number of points in comparisons with at least two SuperSpectraTM that differ on species level. In our example, this minimum level is 900 points, corresponding to a 90% confidence level. Thus the spectrum contains the specific masses of two different species, indicative of a mixed sample. It must be deleted from the database. The information contained in the window “Identify Details” can be saved as a Microsoft Excel file by selecting “Export to Excel” from the context menu.

Name	Sample	Identification	Points	%	Family	Genus	Species	datacount	Always in Database	Save in Tree	Export to LIMS
ANAC_001_0621_1E2[e]	18053559-1_02	by superspec						93	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ANAC_001_0621_1F2[e]	31106225-1_02	superspectru	952.0	95.2	Family I/St	Stenotrophomonas	loma/galkylinus	198	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1116.0	99.9	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1100.0	99.9	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1000.0	99.0	Family I Pse	Pseudomonas	aeruginosa		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1045.0	99.9	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1041.0	99.9	Family I Pse	Pseudomonas	aeruginosa		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1028.0	99.9	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	987.0	98.7	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	980.0	98.0	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	978.0	97.8	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	950.0	95.0	Family I Pse	Pseudomonas	aeruginosa		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	929.0	92.3	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	912.0	91.2	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	906.0	90.6	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	904.0	90.4	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	901.0	90.1	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	900.0	90.0	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	896.0	89.6	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	875.0	87.5	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	852.0	85.2	Family I Ent	Escherichia	coli		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ANAC_001_0621_1F4[e]	31106236-1_02	superspectru	1320.0	99.9	Family I/Stan	Stenotrophomonas	maltophila	181	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1342.0	99.9	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1196.0	99.9	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1157.0	99.9	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1075.0	99.9	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1041.0	99.9	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1023.0	99.8	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1015.0	99.9	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	940.0	94.0	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	050.0	05.0	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	801.0	80.1	Family I/Stan	Stenotrophomonas	maltophila		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ANAC_001_0621_1G1[e]	03116256-1	superspectru	1118.0	99.8	Family I Ent	Providencia	rettgeri	196	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1110.0	99.9	Family I Ent	Providencia	rettgeri		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1098.0	99.9	Family I Ent	Providencia	rettgeri		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1004.0	99.9	Family I Ent	Providencia	rettgeri		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	980.0	98.0	Family I Ent	Providencia	rettgeri		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	972.0	97.2	Family I Ent	Providencia	rettgeri		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	846.0	84.6	Family I Ent	Providencia	rettgeri		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	770.0	77.0	Family I Ent	Providencia	rettgeri		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ANAC_001_0621_1H3[e]	03116268-1	superspectru	1195.0	99.9	Family I Ent	Klebsiella	oxytoca	208	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	1131.0	99.9	Family I Ent	Klebsiella	oxytoca		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	889.0	88.9	Family I Ent	Klebsiella	oxytoca		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		superspectru	800.0	80.0	Family I Ent	Klebsiella	oxytoca		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ANAC_001_0621_1U4[e]	0403116671R-1	superspectru	1152.0	99.9	Family II/Mo	Acinetobacter	ghisborgii	179	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>



## SARAMIS Auto Experiment setup

1. In the MALDI-MS main window, open the Auto Experiment window (MALDI-MS main window: Automation » Auto Experiment...) (see figure 1).

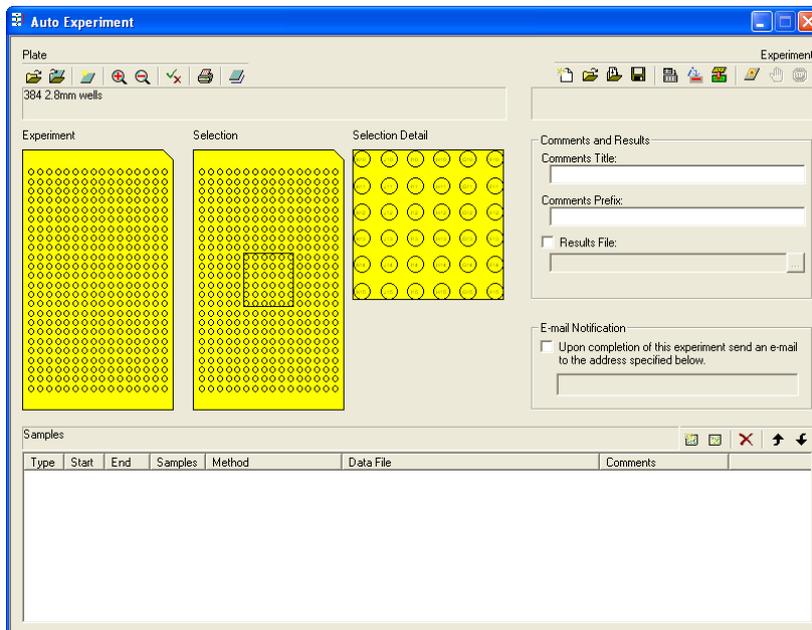
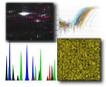


Figure 1: Auto experiment window (MALDI-MS software).

2. Ensure the correct plate file is loaded corresponding to the MALDI target to be used for the auto experiment. If the currently loaded plate file is incorrect, click the  button on the 'Plate' panel toolbar (left hand side of the screen). In the Select Plate window, select the correct plate file and click Open.
3. In the 'Selection' plate view (middle yellow plate view), select (click) the well positions containing the calibration standards. Selected wells are highlighted in black. To simplify selecting wells, right click on the 'Selection' plate file view and select 'Expand experiment plate' (once all the positions required have been selected, click OK to close the expanded view). If a group of wells are to be selected, click-and-drag over the required wells to select them. Once all the wells containing calibration standards have been selected, click the  button to create a new group in the 'Samples' table (bottom panel of the Auto Experiment window). A sample line containing the positions and number of selected wells should appear in the 'Samples' table. The selected wells should also appear in the 'Experiment' plate view (left yellow plate view). If it is necessary to add/remove samples from an existing sample group, in the 'Samples' table, click the line containing the sample group to be modified. In the 'Selection' plate view, add/remove wells as required by clicking the wells to select/deselect them. Once the changes have been made, click the  button to update the changes in the 'Samples' table.



4. In the calibration standards sample group created in step 3., double-click the method file currently loaded (in the column headed ‘Method’). Click the button which appears at the end of the field. Select a suitable method for the analysis of the calibration standards (refer to the IntelliMarque™ – Method Editor reference file). Still in the calibration standards sample group created in step 3., double-click the data file currently loaded (in the column headed ‘Data File’). Click the button which appears at the end of the field. Select the location where the calibration standards data files are to be saved. Enter a suitable filename for the calibration standards. Once the filename and data storage location have been specified, click Save. Samples analysed using a method in which the ‘Calibration’ option box was checked/ticked are identified as calibration standards by a red outline in the ‘Selection’ and ‘Experiment’ plate file views.

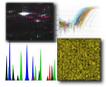
5. In the ‘Selection’ plate view (middle yellow plate view), select (click) the well positions containing the samples to be analysed (not calibration standards). Selected wells are highlighted in black. To simplify selecting wells, right click on the ‘Selection’ plate file view and select ‘Expand experiment plate’ (once all the positions required have been selected, click OK to close the expanded view). If a group of wells are to be selected, click-and-drag over the required wells to select them. Once all the wells containing samples have been selected, click the button to create a new group in the ‘Samples’ table (bottom panel of the Auto Experiment window). A sample line containing the positions and number of selected wells should appear in the ‘Samples’ table. The selected wells should also appear in the ‘Experiment’ plate view (left yellow plate view). If it is necessary to add/remove samples from an existing sample group, in the ‘Samples’ table, click the line containing the sample group to be modified. In the ‘Selection’ plate view, add/remove wells as required by clicking the wells to select/deselect them. Once the changes have been made, click the button to update the changes in the ‘Samples’ table.

6. In the samples sample group created in step 3., double-click the method file currently loaded (in the column headed ‘Method’). Click the button which appears at the end of the field. Select a suitable method for the analysis of the samples (refer to the IntelliMarque™ – Method Editor reference file). Still in the samples sample group created in step 3., double-click the data file currently loaded (in the column headed ‘Data File’). Click the button at which appears at the end of the field. Select the location where the samples data files are to be saved. Enter a suitable filename for the samples. Once the filename and location have been specified, click Save.

There should now be 2 lines in the ‘Samples’ table – the first line (group) containing the calibration standards information and the second line (group) containing the samples information.

7. Enter a ‘Comments Title’ and ‘Comments Prefix’ if required. These will appear on all the data files acquired during the Auto Experiment.

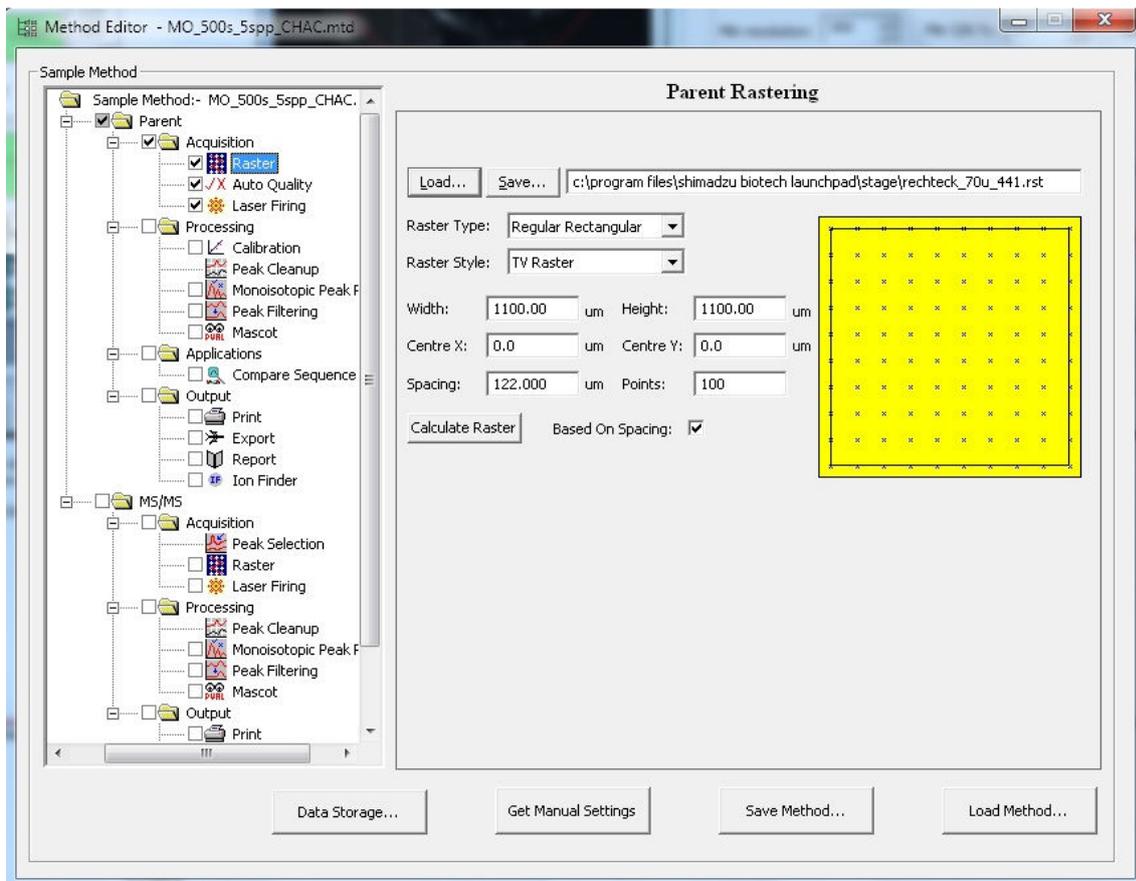
8. To save the sample table (Auto Experiment), click the button on the Experiment toolbar (right hand side of the Auto Experiment window). Select the location where the sample table is to be saved. Enter a suitable filename. Once the filename and location have been specified, click Save.

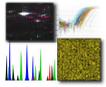


9. To ensure that the initial calibration (following acquisition of the first calibration spot) is successful, it is necessary to manually acquire a calibration spectrum (once) and update/create the initial instrument calibration prior to starting the Auto Experiment/IntelliMarque™ experiment. This is also a good opportunity to check that the laser power specified in the calibration method (refer to the IntelliMarque™ – Method Editor reference guide) is correct. By performing the initial calibration manually, this will ensure that (a) the first automated calibration will not fail e.g. due to tolerance (resulting in uncalibrated sample data) and (b) the acquisition parameters are set correctly.

10. To start the Auto Experiment, click the button. The progress of the Auto Experiment can be monitored in the Auto Experiment window – samples awaiting analysis are highlighted black, samples currently being analysed are highlighted grey and samples already processed/analysed are highlighted green.

The following values should be added into the Method for Saramis identification of bacterial samples:





The screenshot shows the 'Method Editor' window for 'MO\_500s\_5spp\_CHAC.mtd'. The left sidebar displays a tree view of the method structure, with 'Parent' expanded to show 'Acquisition' (Raster, Auto Quality, Laser Firing) and 'Processing' (Calibration, Peak Cleanup, Monoisotopic Peak F, Peak Filtering, Mascot). The main panel is titled 'Parent Auto Quality' and contains the following settings:

- Monitor:** Mass range: 3000-13000; Start noise: 17000; Noise width: 50.
- Auto power:** Power limits: 56; Start power: 56.
- Criteria:** Min intensity: 8 mV; Min S/N: 5; Min resolution: 300; Min S/N %: 50; Max rejects: 2; Max intensity: 800.
- Auto aim:** Prescan: 1 Profiles/Pt; Min points: 5; % Cutoff: 80.
- Auto calibration:** Lock Mass Calibration: ; Lock Mass: 0.0000; Tolerance: 10 Da.
- Performance measurement options:**  Abort experiment;  Prompt for user's intervention;  Skip to next sample.

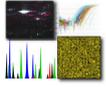
Buttons at the bottom include 'Data Storage...', 'Get Manual Settings', 'Save Method...', and 'Load Method...'.

The screenshot shows the 'Method Editor' window for 'MO\_500s\_5spp\_CHAC.mtd'. The left sidebar shows the method structure with 'Parent' expanded to 'Acquisition' (Raster, Auto Quality, Laser Firing) and 'Processing' (Calibration, Peak Cleanup, Monoisotopic Peak F, Peak Filtering, Mascot). The main panel is titled 'Parent Laser Firing' and contains the following settings:

- Setup:** Configuration: Linear\_SARAMIS; Mass Range: 2000.0 - 20000.0; Calibration: ECAL\_acyano.
- Laser Firing:** Power: 56; Profiles: 100 per sample; Shots: 5 accumulated per profile; Laser Rep Rate (Hz): 50.0; Ion Gate (Da): Off; Pulsed Extraction optimised at (Da): 8330.0;  Restrain peak cleanup.

Buttons at the bottom include 'Data Storage...', 'Get Manual Settings', 'Save Method...', and 'Load Method...'.





## **Genomics**

- I. Designing and performing an MLPA for DNA detection and Quantification**
- II. Perform an HRM Mutation Scanning Experiment**
- III. Real-time gene expression analysis**
- IV. DNA methylation arrays**
- V. Microarray Data Analysis**

### **General MLPA Protocol for DNA Detection & Quantification**

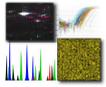
#### **Introduction**

Multiplex Ligation-dependent Probe Amplification (MLPA<sup>®</sup>) is a high-throughput method developed to determine the copy number of up to 50 genomic DNA sequences in a single multiplex PCR-based reaction. The MLPA reaction results in a mixture of amplification fragments between 100 to 500 nt in length which can be separated and quantified by capillary electrophoresis. In contrast to a standard multiplex PCR, only one pair of PCR primers is used in the MLPA PCR reaction, resulting in a more robust system.

Fundamental for the MLPA technique is that it is not sample DNA that is amplified during the PCR, but MLPA probes that hybridised to the sample DNA. Each MLPA probe consists of two oligonucleotides which have to hybridise to adjacent DNA targets in order to be ligated. Since only ligated probes can be amplified exponentially during the subsequent PCR, the amplification of a given probe depends on the presence of its target DNA. Moreover, the relative amount of each probe's amplification product is a measure for the relative quantity of that probe's target sequence in the DNA sample. To quantify this relation, two calculations need to be done:

- (1) establishing the relative signal of each probe as compared to that of the other probes (intra-sample normalisation);
- (2) establishing the relative signals in the sample under study compared to those of reference samples (inter-sample normalisation).

Upon normalisation, it is possible to establish whether any deletions or duplications in the target DNA have occurred.



## Experimental set up

Slight differences between different experiments may affect the MLPA peak pattern, it is therefore necessary to include reference samples in each MLPA experiment. Reference samples are samples in which the target sequences are assumed to have a ‘normal’ copy number. Usually, reference samples are DNA samples obtained from healthy individuals. It is recommended to use reference samples that are derived from the same type of tissue and are extracted using the same method as the test samples. Multiple reference samples are needed to estimate each MLPA probe’s reproducibility within each experiment.

Follow these steps to set up an MLPA experiment:

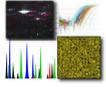
1. Select test and reference samples.
2. Use at least 3 reference samples in each MLPA run. When using more than 21 samples, add 1 additional reference sample for each 7 samples. Reference samples should be distributed randomly over the sample plate to minimise variation.
3. Include a ‘no DNA control’ reaction. This is an MLPA reaction on dH<sub>2</sub>O or TE (10 mM Tris-HCl pH 8.2 + 0.1 mM EDTA) instead of DNA which will reveal possible contamination of dH<sub>2</sub>O, TE, MLPA reagents, electrophoresis reagents or the capillaries.
4. When available, inclusion of positive control samples is recommended. However, this is not essential.

When using DNA from cell lines as positive control sample, please take into account that cell lines may have acquired additional copy number changes, including gains or losses of complete chromosomes.

## 3. Sample treatment

It is recommended using a total of 50-100 ng of human chromosomal DNA in 5 µl for each MLPA reaction. The treatment of DNA (storage and extraction method) may affect MLPA results. As the effect of contaminants is generally consistent and reproducible, this effect will be corrected when the samples have undergone identical treatment. In addition, it is recommended to use similar DNA concentrations in the test/reference samples to minimise the variation caused by possible contaminants.

1. MLPA does not require a special method or kit for DNA extraction. However, it is important that the extraction method does not leave a high concentration of contaminants, such as salt. The use of heparinised blood is not recommended. Traces of heparin are very difficult to remove from DNA preparations and can distort the MLPA PCR reaction. RNase treatment is required for some probemixes. For samples derived from paraffin-embedded tissues, we recommend to use reference samples derived from similarly treated healthy tissue.



2. Dissolve the DNA in TE (10 mM Tris-HCl pH 8.2 + 0.1 mM EDTA). The pH of the DNA preparation should be between 8.0 - 8.5 in order to prevent depurination during the initial heat treatment at 98 °C
3. Aliquot precious reference / control samples and store the DNA at -20 °C. Despite the fact that DNA is very stable when dissolved in TE, storage at 4 °C for several years is not recommended. Contamination with micro-organisms and/or moulds can deteriorate your samples.
4. If needed, dilute the DNA in TE to a working stock of 10-20 ng/μl. Reducing the amount of sample DNA also reduces the amounts of impurities and often has a beneficial effect on the MLPA results.
5. When the DNA concentration is below 4 ng/ μl, an ethanol precipitation is recommended. Glycogen can be used as a carrier in ethanol precipitations. Concentrating samples by evaporation, e.g. using SpeedVac, may result in high EDTA concentrations that can influence MLPA results.

#### **MLPA Reaction - DNA detection/quantification**

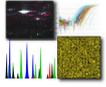
The use of a thermocycler with heated lid (99-105 °C) is essential.

##### **DNA denaturation (day 1)**

1. Label 0.2 ml tubes, strips or plates.
2. Add 5 μl of DNA working stock to each tube (or TE/dH<sub>2</sub>O for a ‘no DNA control’ reaction).
3. Place the tubes in a thermocycler and start the MLPA program (see 4.5 PCR program for the MLPA reaction): denature sample DNA for 5 minutes at 98 °C and cool the samples to 25 °C before opening the thermocycler.

##### **Hybridisation reaction (day 1)**

1. Prepare a hybridisation master mix containing for each reaction: 1.5 μl MLPA buffer (yellow cap) + 1.5 μl probemix (black cap). Mix well by pipetting or vortexing.
2. When the thermocycler reaches 25 °C, add 3 μl of the hybridisation master mix to each sample tube. Mix well by pipetting up and down.
3. Continue the thermocycler program: incubate for 1 minute at 95 °C, then for 16 – 20 hours at 60 °C.



### **Ligation reaction (day 2)**

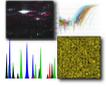
1. Prepare a ligase master mix, containing for each reaction: 3  $\mu$ l Ligase-65 buffer A (transparent cap) + 3  $\mu$ l Ligase-65 buffer B (white cap) + 25  $\mu$ l dH<sub>2</sub>O. Mix by pipetting or vortexing.
2. Add 1  $\mu$ l of Ligase-65 (green cap) for each reaction to the ligase master mix. Mix well by pipetting gently up and down. Never vortex enzyme solutions.
3. Continue the thermocycler program: pause at 54 °C.
4. When the samples are at 54 °C, add 32.0  $\mu$ l of the ligase master mix to each reaction tube. Mix by gently pipetting up and down.
5. Continue the thermocycler program: 15 minutes incubation at 54 °C (for ligation), followed by 5 minutes at 98 °C for heat inactivation of the Ligase-65 enzyme and then pause at 15 °C.

### **PCR reaction (day 2)**

1. Label new tubes for the PCR reaction.
2. Prepare a PCR buffer mix containing for each reaction: 4  $\mu$ l SALSA PCR buffer + 26  $\mu$ l dH<sub>2</sub>O. Mix briefly by vortexing.
3. Add 30  $\mu$ l of the PCR buffer mix to each new tube.
4. At room temperature, transfer 10  $\mu$ l of each ligation product to its corresponding PCR tube.
5. Prepare a polymerase master mix by adding for each reaction: 2  $\mu$ l SALSA PCR-primers (brown cap) + 2  $\mu$ l SALSA Enzyme dilution buffer (blue cap) + 5.5  $\mu$ l dH<sub>2</sub>O + 0.5  $\mu$ l SALSA Polymerase (orange cap). Mix well by pipetting up and down; do not vortex. Store on ice until use.
6. Continue the thermocycler program: pause at 60 °C and place the PCR tubes in the thermocycler.
7. While these tubes are in a thermocycler at 60 °C, add 10  $\mu$ l polymerase mix to each tube. Mix by pipetting gently up and down and continue the thermocycler program immediately: 35 cycles: 30 seconds 95 °C; 30 seconds 60 °C; 60 seconds 72 °C. End with 20 minutes incubation at 72 °C and then pause at 15 °C.

### **Fragment separation**

The amount of MLPA PCR reaction, run conditions and fluorescent label required for analysis all depend on the type of capillary electrophoresis instrument or gel electrophoresis instrument used.



For proper fragment separation, the recommended initial instrument settings often require optimisation.

#### *Equipment and reagents*

1. Capillary electrophoresis instrument with fragment analysis software
2. High quality formamide (e.g. Hi-Di Formamide, Applied Biosystems nr.4311320).
3. Labelled size standard: ROX-500, TAMRA-500, LIZ-500 or Beckman D1-600.
4. Polymers POP-4 (preferred), POP-7, POP-6 (not recommended) or Beckman gel.

#### *Fluorescent label, sample preparation and recommended initial instrument settings*

*ABI-Prism 3100, ABI-Prism 3100 Avant , ABI-3130, ABI-3130XL , ABI-3730, ABI-3730XL, ABI-3500*

MLPA PCR Primer label: **FAM**. Recommended capillaries: 36 or 50 cm.

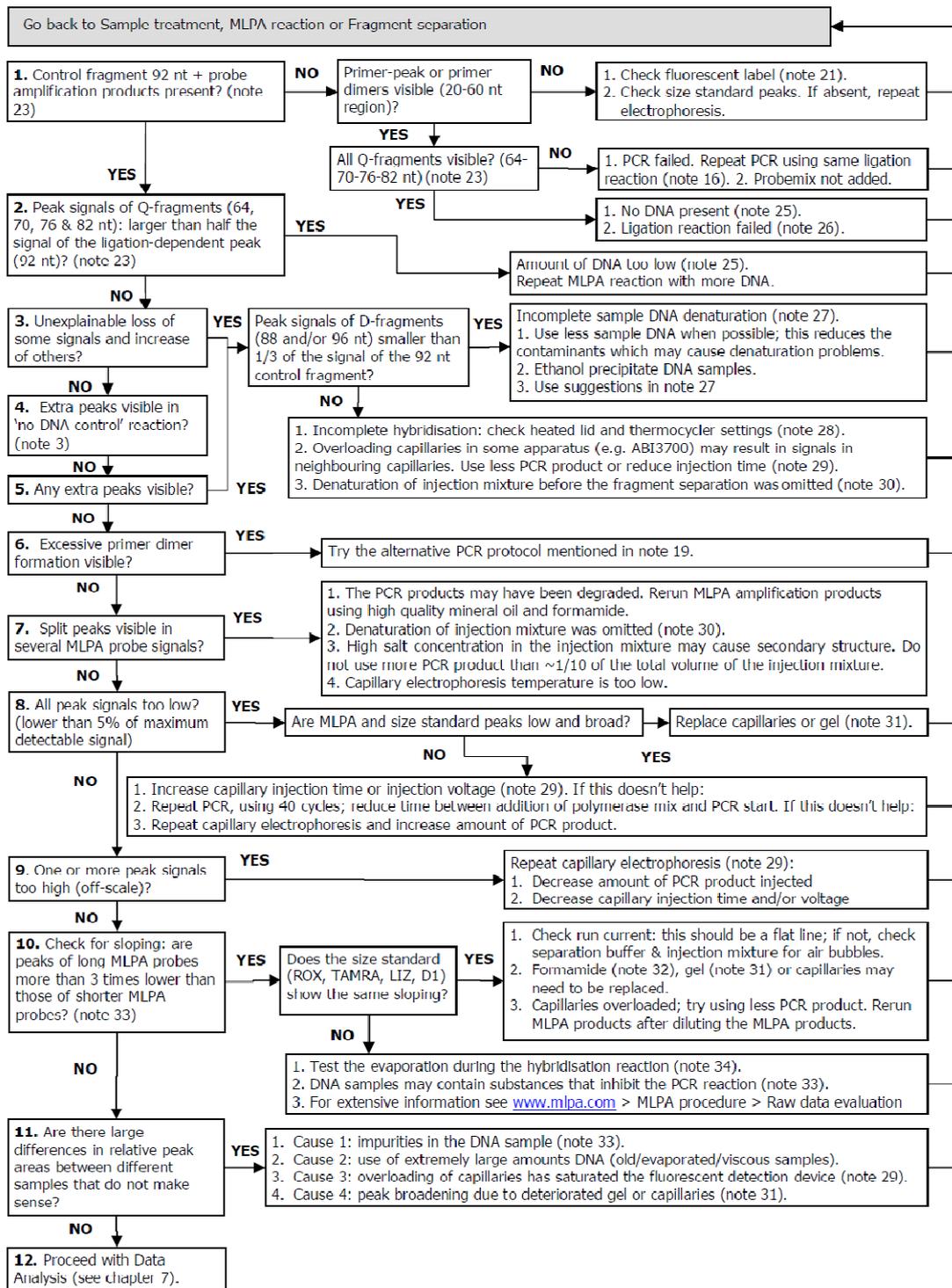
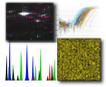
Method: Following the PCR reaction, mix: 0.7 µl of the PCR reaction + 0.3 µl ROX or 0.2 µl LIZ size standard + 9 µl HiDi formamide. Seal the injection plate and incubate for 2 minutes at 80 °C.

Initial settings: injection voltage: 1.6 kV; injection time: 15 sec.; run voltage: 15 kV; polymer: POP4. Polymer POP7 can be used with a run voltage of 10 kV and increased run time.

#### **Peak pattern evaluation**

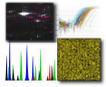
Visual examination of the peak patterns, using the Raw data checklist and the peak pattern flow chart, is essential to prevent false positive results. Based on this evaluation it may be necessary to repeat (parts of) the MLPA procedure, e.g. run samples again on the capillary sequencer when some peaks are above the fluorescence maximum of the sequencer or when most peaks are very low. Coffalyser software can be used for rapid evaluation of peak patterns before and after size calling. Please note that automatic analysis by Coffalyser will not always identify problems such as split peaks and excessive broadening of longer peaks which are rapidly identified by visual examination of the peak patterns. Only data passing the peak pattern evaluation should be used for the actual MLPA data analysis.

The peak pattern flow chart should be used for a more extensive evaluation of the peak patterns. Visual examination of the peak patterns of the sized-called data, including the size standard, will identify problems e.g. due to the use of deteriorated gel, old capillaries or low quality formamide. Part of the peak pattern evaluation is the evaluation of the control fragments with amplification products between 64 nt and 105 nt and of the ‘no DNA control’ reaction. Please note that only data that has passed the raw data checklist and the peak pattern flow chart is suitable for the actual MLPA data analysis.



## Data analysis

Data that has passed raw data and peak pattern evaluation can be used for data analysis. The absolute fluorescence intensities detected dependent on many factors, including bleaching of the fluorescent label, variability between capillaries and pipetting accuracy. Intra sample normalisation of electrophoresis results, resulting in data with the relative fluorescence intensity of each peak, is



therefore essential for obtaining useful MLPA data. Coffalyser software, developed at MRC-Holland, is recommended for the analysis of MLPA data. All data normalisation steps are built-in functions of Coffalyser.

#### *Data normalisation*

Many different methods exist for the normalisation of MLPA data and it is impossible to single out one strategy that works best for all probemixes. Each product description of a SALSA MLPA kit purchased from MRC-Holland describes the best method to normalise your data. The majority of SALSA MLPA probemixes can be normalised following the method described below (see notes 35, 36).

1. Intra-sample normalisation: divide the peak area of each probe's amplification product by either
  - (a) the total peak area (or peak height) of only the reference probes (block normalisation) or
  - (b) the combined peak areas of all probes (population normalisation).
2. Inter-sample normalisation: divide these intra-normalised probe ratios of a sample by the average intranormalised probe ratio of all reference samples. This provides a probe ratio for each probe per sample.

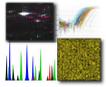
#### **Result interpretation**

To determine whether the produced ratios are reliable, sufficient knowledge about the MLPA technique and the studied application is essential.

Arranging probes according to chromosomal location facilitates interpretation of the results and may reveal more subtle changes such as those observed in mosaic cases. Please take the following in mind:

–Sequence changes (SNPs, point mutations) in the target sequence detected by a probe can cause false positive MLPA results. Mutations / SNPs close to the ligation site (1-5 nt) can reduce or prevent the probe signal by preventing ligation of the two probe oligonucleotides. Mutations / SNPs at further distance (at least up to 20 nt from the probe ligation site) can result in a lower probe signal by destabilisation of the binding of the probe oligonucleotide to the sample DNA. MRC-Holland avoids SNPs when designing probes if possible. However, the number of known SNPs in the databases is growing rapidly and it is impossible to completely avoid SNPs. Furthermore, most SNPs in the SNP databases are not validated, some SNPs have only been detected once and the reported SNPs are not always reliable.

Please notify us when a polymorphism or a frequent pathogenic mutation influences one of the probe signals.



–Most MLPA probes detect a sequence with a length between 55 and 76 nt long. The majority of the product descriptions only contain a part of the sequence detected by the probes, e.g. 12 nt at each side of the ligation sites. The complete sequences detected by the probes, as well as the identity of the reference probes, is available upon request: [info@mlpa.com](mailto:info@mlpa.com).

–Copy number changes detected by only a single probe therefore always require confirmation by other methods. Sequencing is often used to determine whether a mutation or polymorphism in the probe target sequence is present. Long Range PCR and qPCR are often used to confirm (single) exon deletions.

–The frequency of single exon deletions is in general higher in genes with large introns than in small genes with very short introns.

–Note that not all deletions and duplications detected by MLPA will be pathogenic. Partial gene duplications within the gene may disrupt that copy of the gene thereby resulting in a disease, while a duplication of the complete gene at another chromosomal location might not be pathogenic.

–For most applications, the major cause of genetic defects will be small (point) mutations, most of which will not be detected by SALSA MLPA kits. MLPA will not detect most inversions, translocations, nor copy number changes outside the sequence detected by the SALSA MLPA probes.

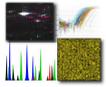
–Probe ratios below 0.7 are often regarded as an indication of a heterozygous deletion (copy number change from two to one allele) and probe ratios above 1.3 of a duplication (copy number change from two to three alleles).

MLPA results are more plausible when:

- o The overall standard variation per probe in the reference samples is low (< 10 %).
- o Few or none abnormalities are found in a patient cohort for a disease in which copy numbers changes are known to be rare.
- o Probes for adjacent exons show a decreased or increased signal, e.g. exons 2 to 5, indicating a multi-exon deletion or duplication.
- o The same result is obtained in a new experiment using less (minimum 20 ng) DNA and/or using different reference samples. When less DNA is used, possible contaminants which may influence the probe signal are diluted.

A result is unlikely to be reliable when:

- o Probes for non-neighbouring exons show a decreased or increased signal, e.g. exon 3 and 17.
- o In the same sample one or more reference probes show an abnormal copy number.
- o Many abnormalities are found in a patient cohort of a disease in which copy number changes are known to be rare.



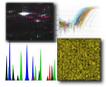
### **SALSA MLPA probemix P376-B2 BRCA1ness**

This probemix contains 2 probes for the BRCA1 gene, 2 probes for the BRCA2 gene and 34 probes for 11 different chromosomal regions that are suggested to be clinically relevant in prediction for BRCA1- association and for PARP inhibitors benefit. In addition, 10 reference probes have been included in this probemix, detecting 6 different autosomal chromosomal locations, which are relatively quiet in breast cancer genome and which have been validated in a large collaborative study (Lips E et al., 2011, Breast Cancer Res. 13:R107). However, it should be noticed that breast cancer karyotypes can harbour multiple numerical and structural aberrations, which can complicate interpretation of these reference probes.

#### **Interpretation of the results - Data analysis**

The P376-B2 BRCA1ness probemix contains 48 MLPA probes with amplification products between 124 and 500 nt. In addition, it contains 9 control fragments generating an amplification product smaller than 120 nt: four DNA Quantity fragments (Q-fragments) at 64-70-76-82 nt, three DNA denaturation control fragments (D-fragments) at 88-92-96 nt, one X-fragment at 100 nt and one Y-fragment at 105 nt. More information on how to interpret observations on these control fragments can be found in the MLPA protocol. Data generated by this probemix should be normalised with a more robust method, as the target sites of the reference probes maybe gained or lost.

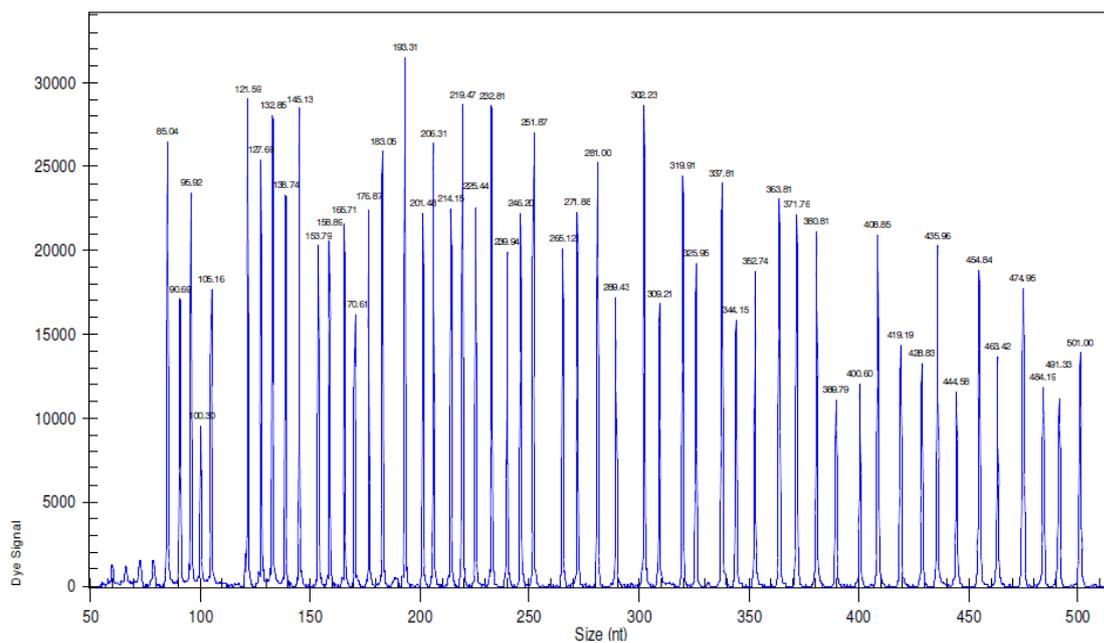
1. Intra-sample normalisation should be performed by dividing the signal of each target-specific probe by the signal of every single reference probe in that sample, thus creating as many ratios per target-specific probe as there are reference probes. Subsequently, the median of all these produced ratios per probe should be taken; this is the probe's Normalisation Constant.
2. Secondly, inter-sample comparison should be performed by dividing the Normalisation Constant of each probe in a given sample by the average Normalisation Constant of that probe in all the reference samples. Data normalisation should be performed within one experiment. Always use sample and reference DNA extracted with the same method and derived from the same source of tissue. Confirmation of deletions, duplications and amplifications can be done by e.g. Southern blotting, long range PCR, qPCR, FISH. An alternative normalization method used in validation study by Lips et al., 2011, BCR: To automate the interpretation of the fragment analysis, the relative quantity of the amplified probes in each sample should be determined using a template in Excel (Microsoft). For this purpose, the relative peak areas for each probe should be calculated as fractions of the total sum of peak areas in each sample. Subsequently, the fraction of each peak should be divided by the average peak fractions of the corresponding probe in control samples (Promega female reference DNA). These ratio data should be then used for consequent BRCA1-like class prediction.



### BRCA1ness classification:

To classify a sample as BRCA1-like or non-BRCA1-like a classifier in the statistical programming language R should be used. This classifier was developed using Prediction analysis for microarrays (PAM) (Tibshirani R et al., 2002, PNAS). Data values from Excel for all the 34 target-specific probes should be used. The classifier is based on a training set of 84 samples. Class prediction of new samples is also performed in PAM. The R code is available upon request by the authors (Lips et al., 2011, BCR). The cut-off value to classify a sample as ‘BRCA1-like’ should be set at 0.5. Below this score, a sample should be classified as ‘non-BRCA1-like’.

### SALSA MLPA probemix P376-B2 BRCA1ness sample picture



**Figure 1.** Capillary electrophoresis pattern from a sample of approximately 50 ng human female control DNA analysed with SALSA MLPA probemix P376-B2 BRCA1ness (lot B2-0911).

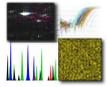
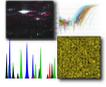


Table 1. SALSA MLPA P376-B2 BRCA1ness probemix

Length (nt)	SALSA MLPA probe	Chromosomal position				
		reference	BRCA1	BRCA2	Lost in BRCA1-like aCGH profile	Gained in BRCA1-like aCGH profile
64-70-76-82	Q-fragments: DNA quantity; only visible with less than 100 ng sample DNA					
88-92-96	D-fragments: Low signal of 88 or 96 nt fragment indicates incomplete denaturation					
100	X-fragment: Specific for the X chromosome					
105	Y-fragment: Specific for the Y chromosome					
124	Reference probe 15370-L13762	7q11				
130	<b>BRCA2 probe</b> 12289-L13282		<b>13q13</b>			
136	<b>GATA3 probe</b> 07632-L07317				<b>10p14</b>	
140	<b>SEMA3B probe</b> 15373-L02625			<b>3p21</b>		
148	<b>CAPN3 probe</b> 05795-L05242			<b>15q15</b>		
155	<b>BTG1 AREA probe</b> 14116-L17758			<b>12q21</b>		
160	Reference probe 00453-L17937	7q31				
167	<b>BRCA1 probe</b> 02610-L02239		<b>17q21</b>			
172	<b>FBN1 probe</b> 02449-L01893			<b>15q21</b>		
178	Reference probe 05506-L04929	10q11				
184	<b>GPR98 probe</b> 03994-L03261			<b>5q14</b>		
193	<b>HLTF probe</b> 02758-L02207				<b>3q24</b>	
202	<b>PAH probe</b> 02332-L01819			<b>12q23</b>		
208	Reference probe 05731-L05170	14q11				
214	<b>LMNB1 probe</b> 09223-L09541			<b>5q23</b>		
220	Reference probe 01119-L17938	10q22				
226	<b>APC probe</b> 14902-L16646			<b>5q22</b>		
233	<b>PTEN probe</b> 13685-L17760			<b>10q23</b>		
241	<b>DEPDC1B probe</b> 09847-L16616			<b>5q12</b>		
247	<b>ABCC4 probe</b> 15371-L17942				<b>13q32</b>	
253 *	Reference probe 06236-L01311	21q11				
265	<b>TPMT probe</b> 09451-L09707				<b>6p22</b>	
274	<b>FANCE probe</b> 02431-L01878				<b>6p21</b>	
283	<b>GCH1 probe</b> 04404-L02610			<b>14q22</b>		
291	<b>PCCA probe</b> 08688-L08700				<b>13q32</b>	
301	<b>HFE probe</b> 14344-L16013				<b>6p22</b>	
310	Reference probe 15372-L00892	2p25				
319	<b>RDH12 probe</b> 08892-L08854			<b>14q22</b>		
328	<b>RASSF1 probe</b> 02248-L01734			<b>3p21</b>		
337	<b>BRCA2 probe</b> 12313-L13306		<b>13q13</b>			
346	<b>THBS1 probe</b> 01678-L01258			<b>15q14</b>		
355	Reference probe 00965-L00552	2p13				
364	<b>LOC254312 probe</b> 01234-L12680				<b>10p14</b>	
373	<b>ATR probe</b> 02560-L02023				<b>3q23</b>	
382	<b>PDCD10 probe</b> 04397-L03795				<b>3q26</b>	
390	<b>BRCA1 probe</b> 02827-L02256		<b>17q21</b>			
400	<b>CEP290 probe</b> 08886-L08848			<b>12q21</b>		
409	<b>MCCC2 probe</b> 06551-L17944			<b>5q13</b>		
418	<b>PCCA probe</b> 08682-L08694				<b>13q32</b>	
427	<b>PIK3CA probe</b> 03827-L21157				<b>3q26</b>	
436	<b>FGF23 probe</b> 07116-L06696				<b>12p13</b>	
445 *	Reference probe 15368-L05185	2p11				
454	<b>BMP4 probe</b> 12473-L13488			<b>14q22</b>		



### **SALSA MLPA KIT P078-B1 Breast Tumour**

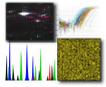
This probe mix contains 39 probes for 10 different genes / chromosomal regions, which are suggested to be prognostically and therapeutically important in breast cancers. In addition, 11 reference probes are included in this probemix detecting 11 different autosomal chromosomal locations, which are relatively quiet in breast tumours. However, it should be noticed that breast cancer karyotypes may harbour complex numerical and structural aberrations, which can complicate interpretation of these reference probes. This SALSA MLPA kit is designed to detect amplifications, gains and losses of one or more genes / chromosomal regions in breast tumours. Heterozygote deletions of probe recognition sequences should give a 35-50% reduced relative peak area of the amplification product of that probe. However, a mutation or polymorphism (e.g. SNP) in the sequence detected by a probe can also cause a reduction in relative peak area, even when not located exactly on the ligation site! In addition, some probe signals may be more sensitive to sample purity and small changes in experimental conditions. Therefore, deletions / duplications and amplifications detected by MLPA should always be confirmed by other methods or by MLPA probemixes with higher resolution in the gene or chromosomal area of interest.

#### ***Data analysis***

The P078-B1 Breast tumour probemix contains 50 MLPA probes with amplification products between 113 and 500 nt. In addition, it contains 9 control fragments generating an amplification product smaller than 110 nt: four DNA Quantity fragments (Q-fragments) at 64-70-76-82 nt, three DNA denaturation control fragments (D-fragments) at 88-92-96 nt, one X-fragment at 100 nt and one Y-fragment at 105 nt. More information on how to interpret observations on these control fragments can be found in the MLPA protocol.

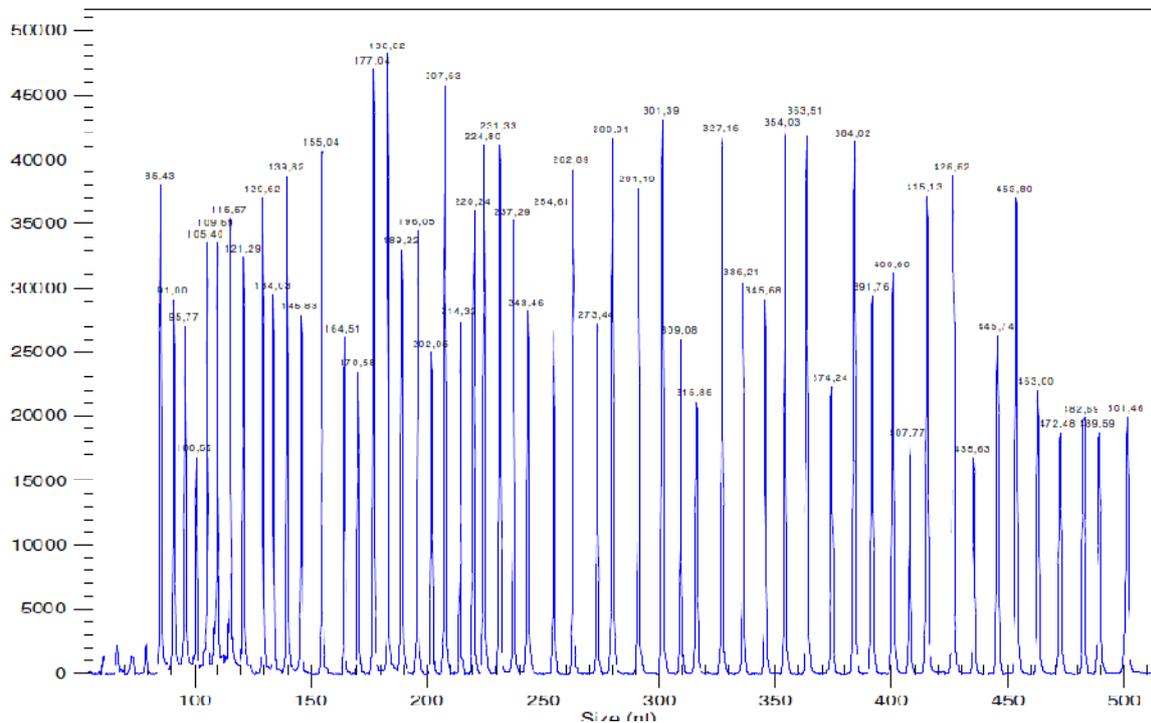
Data generated by this probemix should be normalised with a more robust method, as the target sites of the reference probes may be gained or lost.

1. Intra-sample normalisation should be performed by dividing the signal of each target specific probe by the signal of every single reference probe in that sample, thus creating as many ratios per target-specific probe as there are reference probes. Subsequently, the median of all these produced ratios per probe should be taken; this is the probe's Normalisation Constant.
2. Secondly, inter-sample comparison should be performed by dividing the Normalisation Constant of each probe in a given sample by the average Normalisation Constant of that probe in all the reference samples. Data normalisation should be performed within one experiment. Always use sample and reference DNA extracted with the same method and derived from the same source of tissue. Confirmation of deletions, duplications and amplifications can be done by e.g. FISH, Southern blotting, long range PCR, qPCR. MLPA analysis on tumour samples provides information on the average situation in the cells from which the DNA sample was purified. Gains or losses of genomic regions or genes may not be

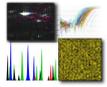


detected if the percentage of tumour cells is low. Reference probes are located in ‘silent’ regions that are not frequently altered in copy number in breast cancer, but there is always a possibility that one or more reference probes have a copy number alteration in a sample. Normal copy number variation in healthy individuals is described in the database of genomic variants: <http://projects.tcag.ca/variation>. Reference probes are carefully selected not to have normal copy number variation, but when in doubt, users should always verify the latest updates of the database and scientific literature when interpreting their findings.

### SALSA MLPA kit P078-B1 Breast Tumour sample picture

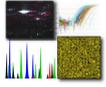


**Figure 1.** Capillary electrophoresis pattern from a sample of approximately 50 ng human male control DNA analyzed with SALSA MLPA kit P078-B1 Breast tumour (lot 0210). Please note a non-specific peak is present at 310 nt in the no DNA control reaction.



**Table 2. SALSA MLPA P078-B1 Breast Tumour probemix**

Length (nt)	SALSA MLPA probe	Chromosomal position reference target gene
64-70-76-82	Q-fragments: DNA quantity; only visible with less than 100 ng sample DNA	
88-92-96	D-fragments: Low signal of 88 or 96 nt fragment indicates incomplete denaturation	
100	X-fragment: Specific for the X chromosome	
105	Y-fragment: Specific for the Y chromosome	
113	<b>ERBB2</b> probe S0393-L12911	<b>17q12</b>
118	<b>MYC</b> probe S0247-L08464	<b>8q24</b>
124	<b>TRAF4</b> probe 09176-L09350	<b>17q11</b>
132	<b>C11orf30 (EMSY)</b> probe 09173-L09347	<b>11q13</b>
136	<b>ADAM9</b> probe 11992-L12820	<b>8p11</b>
142	<b>ERBB2</b> probe 00675-L00146	<b>17q12</b>
148	<b>IKBKB</b> probe 11993-L12821	<b>8p11</b>
157	<b>MYC</b> probe 00580-L00625	<b>8q24</b>
166	<b>CCNE1</b> probe 02881-L02348	<b>19q12</b>
172	<b>TOP2A</b> probe 11994-L12822	<b>17q21</b>
178	<b>CDH1</b> probe 02410-L02237	<b>16q22</b>
184	Reference probe 11350-L12075	12p13
190	<b>CCNE1</b> probe 09170-L09344	<b>19q12</b>
196	<b>CDC6</b> probe 08611-L13204	<b>17q21</b>
202	Reference probe 05730-L06767	13q11
208	Reference probe 10223-L10704	3q26
215	<b>ESR1</b> probe 11996-L12824	<b>6q25</b>
220	Reference probe 09570-L10024	22q13
226	<b>CPD</b> probe 09628-L09913	<b>17q11</b>
232	<b>ESR1</b> probe 11998-L12826	<b>6q25</b>
238	<b>MYC</b> probe 00672-L00169	<b>8q24</b>
244	<b>ERBB2</b> probe 12048-L12913	<b>17q12</b>
256	<b>C11orf30 (EMSY)</b> probe 09175-L09349	<b>11q13</b>
264	<b>EGFR</b> probe 05969-L05386	<b>7p11</b>
274	Reference probe 00978-L00565	15q25
281	<b>MTDH</b> probe 04151-L03506	<b>8q22</b>
292	<b>CCND1</b> probe 00583-L00148	<b>11q13</b>
301	Reference probe 00971-L09490	10q22
310	<b>ERBB2</b> probe 00986-L00406	<b>17q12</b>
316	<b>BIRC5</b> probe 03717-L02410	<b>17q25</b>
328	<b>TOP2A</b> probe 11999-L13177	<b>17q21</b>
337	<b>MTDH</b> probe 04152-L03507	<b>8q22</b>
346	<b>MED1</b> probe 09963-L13205	<b>17q12</b>
355	<b>CDH1</b> probe 02860-L01849	<b>16q22</b>
364	<b>TOP2A</b> probe 12000-L12828	<b>17q21</b>
373	<b>FGFR1</b> probe 01046-L00624	<b>8p11</b>
384	<b>BIRC5</b> probe 03025-L14708	<b>17q25</b>
391	Reference probe 08964-L09059	5q33
400	<b>FGFR1</b> probe 04440-L03826	<b>8p11</b>
409	Reference probe 00963-L00550	2p16
416	<b>MAPT</b> probe 08358-L08211	<b>17q21</b>
427	<b>EGFR</b> probe 02063-L03283	<b>7p11</b>
436	<b>BIRC5</b> probe 03189-L02540	<b>17q25</b>
445	<b>PRDM14</b> probe 12002-L12830	<b>8p21</b>
453	<b>IKBKB</b> probe 12003-L12831	<b>8p11</b>
463	<b>CCND1</b> probe 05402-L04808	<b>11q13</b>
472	Reference probe 09205-L09581	18p11
481	<b>AURKA</b> probe 10236-L10717	<b>20q13</b>
490	Reference probe 10224-L10705	9q21
500	Reference probe 10218-L14675	7q22



## Perform an HRM Mutation Scanning Experiment

Perform an HRM mutation scanning experiment to screen DNA samples for new single-base changes, insertions/deletions, or other unknown mutations.

### Perform an HRM Mutation Scanning Experiment

Design the HRM experiment

Prepare the HRM reactions

Amplify and melt the DNA

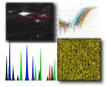
Review the high-resolution melting data

### Design and order the primers

Design attribute	Design guidelines
Amplicon	Length is less than 250 basepairs
Primer length	~20 bases each
T <sub>m</sub>	58 °C to 60 °C (Optimal T <sub>m</sub> is 59 °C)
% GC content	30–80% GC content in each primer
3' end	No more than 2 G+C residues in the last 5 nucleotides at the 3' end
Repeating oligonucleotides	Avoid consecutive identical nucleotides. If you are unable to avoid consecutive identical nucleotides, make sure that each primer contains fewer than 4 consecutive Gs.

If you want to use M13F and M13R primers in the sequencing reaction, add the appropriate M13 tail to the 5' end of the primers:

- M13F (add to the 5' end of the forward primer): TGAAAACGACGGCCAGT
- M13R (add to the 5' end of the reverse primer): CAGGAAACAGCTATGACC



## Select controls

Include controls for each target sequence in your HRM mutation scanning experiment:

- At least one negative control
- At least one wild type control

Run up to 5 replicates for each wild type control to improve your results. Running multiple wild type controls allows you to more effectively define the natural spread or variation within the normal population.

## Prepare the HRM reactions

### **Materials needed:**

Microcentrifuge tubes 0.2 ml

Optical reaction plate appropriate for your Real-Time PCR instrument

MicroAmp™ Optical Adhesive Film

MeltDoctor™ HRM Master Mix

For each target sequence:

– Forward and reverse primers (5 µM each)

DNA samples

Deionized water

Pipettors and pipette tips

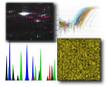
Vortexer

Centrifuge

Prepare the reactions for each replicate group separately, then transfer the reactions to a reaction plate appropriate for your instrument.

1. Prepare the **negative control** reactions in an appropriately sized, labeled tube:

Components	96-well standard reaction plate	
	Volume for one 50-µL reaction	Volume for three 50-µL replicates plus 10% excess
MeltDoctor™ HRM Master Mix	25.0 µL	82.5 µL
Primer 1 (5 µM)	3.0 µL	9.9 µL
Primer 2 (5 µM)	3.0 µL	9.9 µL
Deionized water	19.0 µL	62.7 µL
<b>Total reaction volume</b>	<b>50.0 µL</b>	<b>165.0 µL</b>



2. Prepare **positive control** reactions and unknown reactions in separate appropriately sized, labeled tubes:

Components	96-well standard reaction plate	
	Volume for one 50- $\mu$ L reaction	Volume for three 50- $\mu$ L replicates plus 10% excess
MeltDoctor™ HRM Master Mix	25.0 $\mu$ L	82.5 $\mu$ L
Primer 1 (5 $\mu$ M)	3.0 $\mu$ L	9.9 $\mu$ L
Primer 2 (5 $\mu$ M)	3.0 $\mu$ L	9.9 $\mu$ L
Genomic DNA (20 ng/ $\mu$ L)	2.5 $\mu$ L	8.25 $\mu$ L
Deionized water	16.5 $\mu$ L	54.45
<b>Total reaction volume</b>	<b>50 <math>\mu</math>L</b>	<b>165 <math>\mu</math>L</b>

3. Vortex the reactions to mix, then spin the tubes briefly.
4. Prepare a reaction plate appropriate for your instrument:
- a. Pipet each reaction replicate to the appropriate wells of the optical reaction plate:

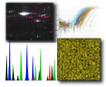
Reaction plate	Reaction volume
Fast 384-well plate	20 $\mu$ L
Fast 96-well plate	20 $\mu$ L
Standard 96-well plate	50 $\mu$ L

- b. Seal the reaction plate with optical adhesive film, then spin the reaction plate.
- c. Confirm that the liquid is at the bottom of the wells in the reaction plate

### Amplify and melt the DNA

After you have prepared the reaction plate, run the plate to amplify and melt the DNA and to generate HRM fluorescence data.

Create and set up the HRM run file:



Run file setting		7500 Fast System with 7500 Software v2.0 or later
Document/experiment		<ul style="list-style-type: none"> <li>Instrument: <b>7500 Fast (96 Wells)</b></li> <li>Experiment type: <b>Quantitation - Standard Curve</b></li> <li>Reagents: <b>Other</b>, then select the <b>Include Melt Curve</b> checkbox</li> <li>Ramp speed: <b>Standard (~ 2 hours to complete a run)</b></li> </ul>
Detector/target and plate layout		<ul style="list-style-type: none"> <li>Reporter: <b>MeltDoctor</b></li> <li>Quencher: <b>None</b></li> </ul>
Plate layout		<ul style="list-style-type: none"> <li>Task for negative control wells: <b>N</b></li> <li>Passive Reference: <b>None</b></li> </ul>
Thermal profile/run method		<ul style="list-style-type: none"> <li>Reaction Volume Per Well: <b>20 µL</b></li> <li>Expert Mode: Select the checkbox</li> <li>Click <b>Select/View Filters</b>, then select only <b>Filter-1</b></li> </ul>

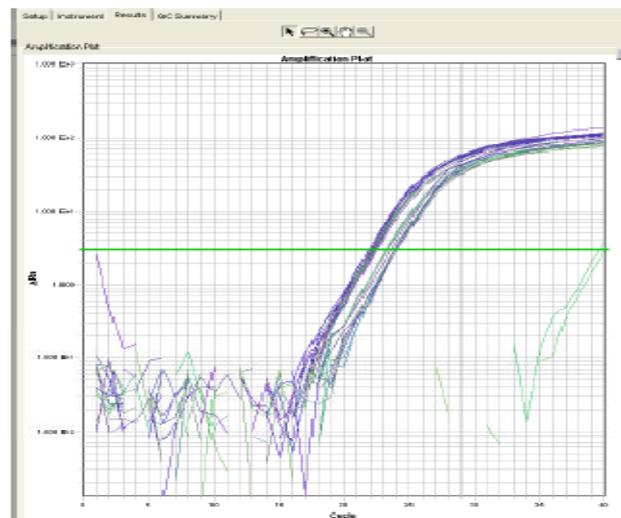
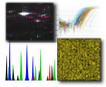
Run the plate:

Stage	Step	Temp	Time	Ramp rate (7900HT only)
Holding	Enzyme activation	95 °C	10 min	100%
Cycling (40 cycles)	Denature	95 °C	15 sec	100%
	Anneal/extend	60 °C	1 min	100%
Melt curve/dissociation	Denature	95 °C	10 sec	100%
	Anneal	60 °C	1 min	100%
	High resolution melting	95 °C	15 sec	1%
	Anneal	60 °C	15 sec	100%

Verify that the samples amplified and review the peaks in the melt curve

1. Review the Amplification Plot for normal characteristics:

- Fluorescence levels that exceed the threshold between cycles 8 and 35
- An exponential increase in fluorescence

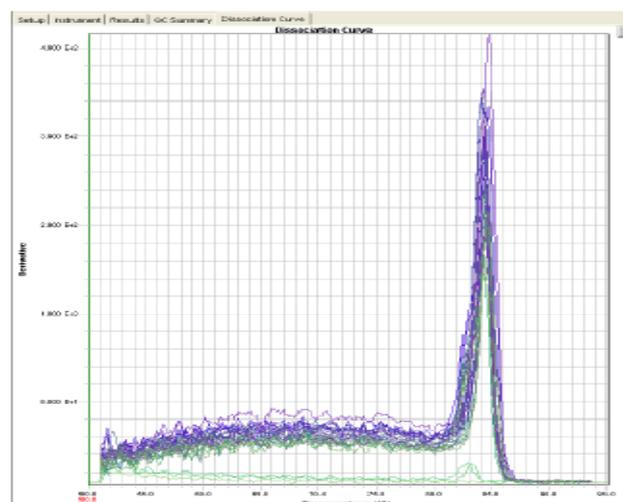


2. Verify that the Dissociation Curve/Melt Curve shows no unexpected T<sub>m</sub> peaks:

If the sequence you amplified contains more than 1 variant or a more complex mutation, you may see more than 1 T<sub>m</sub> peak.

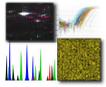
**Note:** Unexpected peaks may indicate contamination, primer dimers, or nonspecific amplification.

**Note:** The data appear noisy because more data is collected during a high resolution melt curve than during a standard melt curve. The extra data are required for analysis with the High Resolution Melting Software.



### Review the high-resolution melting data

After you create, run, and analyze the \*.eds or \*.sds file on 7500 Fast system, use the Applied Biosystems High Resolution Melting Software (HRM software) to perform high resolution melting analysis of the data and screen the samples for mutations.



## Example HRM experiments

To view an example of an HRM mutation scanning experiment, use the example files that are installed with the HRM software:

- 384 Mutation Scanning Example.hrm
- 96well Mutation Scanning Example.hrm

The files are located in  $X:\backslash$ Applied Biosystems\HRM\experiments, where  $X$  is the drive where you installed the HRM Software.

## Create and set up the HRM experiment

1. Create an HRM experiment in the HRM Software using the \*.eds or \*.sds run file from your 7500 Fast system.

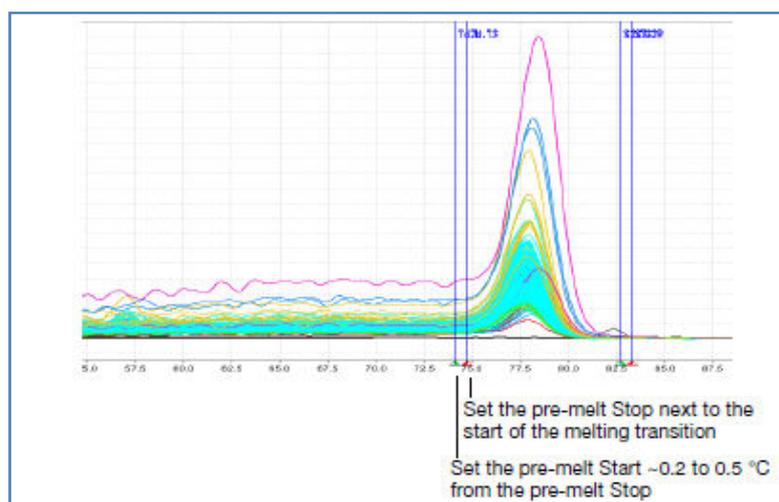
**Note:** If this is your first time creating an HRM experiment with the HRM software, select the default HRM calibration file.

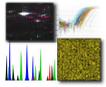
**IMPORTANT!** Use the HRM calibration file name to make sure that the HRM calibration file that is assigned to the HRM experiment is correct. The HRM calibration file must be:

- Run on the same instrument system (the same instrument, the same block type, and the same software version) as the run file
- Run using the same HRM dye and master mix used in the HRM calibration plate

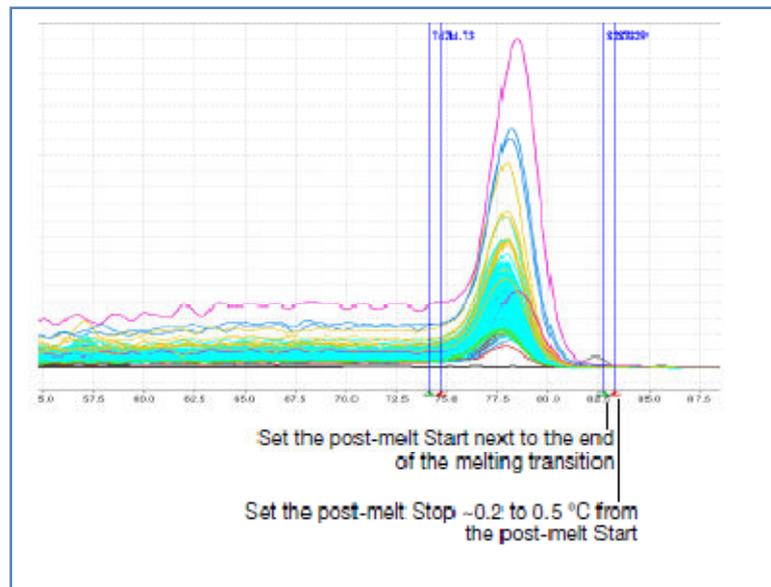
2. In the Derivative Melt Curves plot, review and adjust the pre- and post-melt regions to optimize your separation and variant calls. For most experiments, set the pre- and post-melt regions as close as possible to the melting transition region:

The pre-melt Start and Stop temperature lines (green and red arrows on the left) should be approximately 0.2 to 0.5 °C apart from each other.





The post-melt Start and Stop temperature lines (green and red arrows on the right) should be approximately 0.2 to 0.5 °C apart from each other.



3. For each control sample, enter information about that control in the HRM Software and assign the control to the appropriate wells.

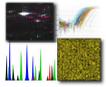
### About the melting profiles

The melting profile of a PCR product depends on its GC content, length, sequence, and heterozygosity. High-resolution melting analysis calls variants based on the differences in the shape of the melt curves and the differences in the  $T_m$  values. In mutation scanning experiments, the variants have a different curve shape or  $T_m$  compared to the wild type.

### Review the populations in the Aligned Melt Curves plot

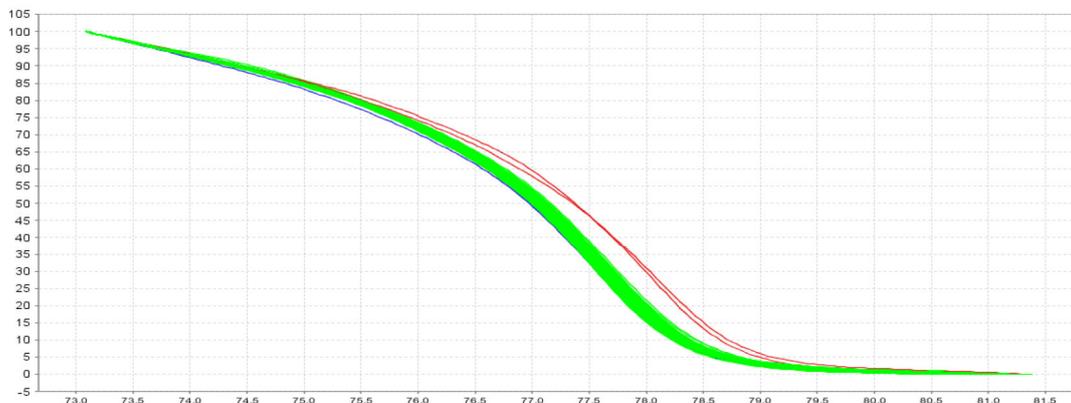
The Aligned Melt Curves plot displays the melt curves as % melt (0 - 100%) over temperature. The melt curves are aligned to the same fluorescence level using the pre and post-melt regions that you set.

1. In the Analyzed Data pane, select the **Aligned Melt Curves** tab.
2. Review:
  - Wild type controls – Do the melt curves for the wild type controls cluster well? Are there any outliers?
  - Possible mutations – Are there any samples with melt curves that are different from the wild type melt curves?



### Aligned Melt Curves example

In the example below, there is 1 distinct variant group for the wild type samples. There are 2 samples that vary from the wild type samples and may contain mutations.



### Review the Difference Plot for outliers

The Difference Plot displays the aligned data as the difference in fluorescence between the melt curve for a reference sample and the other melt curves. You can select a control or any well as a reference. After you select the reference, the software subtracts the reference curve from the other curves.

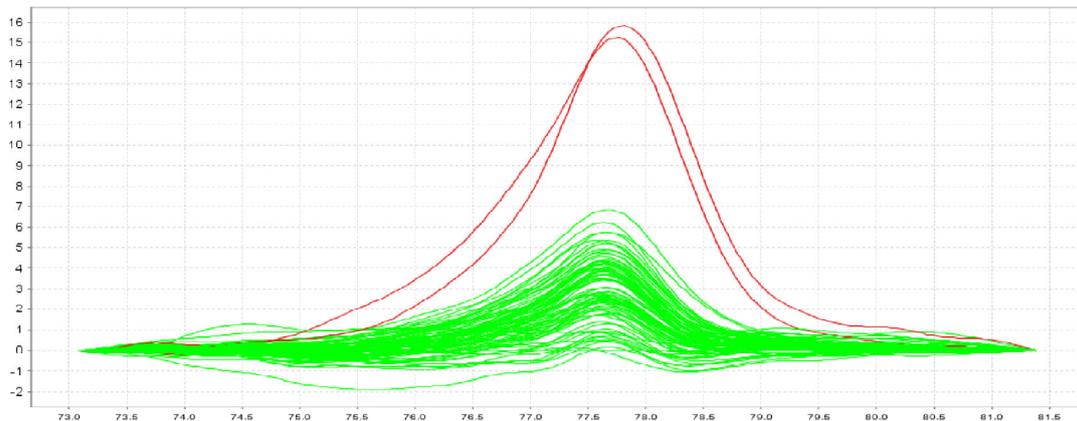
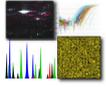
The Difference Plot allows you to more easily see small differences between curves and identify outliers.

1. In the Analyzed Data pane, select the **Aligned Data - Difference Plot** tab.
2. From the **Reference** dropdown menu, select a control or any well as the reference, then review:
  - Variant clusters – How many distinct clusters are displayed?
  - Outliers – How tight are the curves within each variant cluster?

**Note:** Try selecting different reference samples to find the optimal display of the clusters.

### Difference Plot example

In the example below, there is 1 distinct variant group for the wild type samples. There are 2 samples that vary from the wild type samples and may contain mutations.



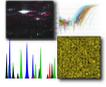
### Review the software calls

The High Resolution Melting Software automatically makes a call for each sample according to the shape of the aligned melt curves and the T<sub>m</sub>. Review the software calls, then omit outliers or change calls.

1. In the Results pane, click the **Well** column header to sort the results according to the well position.
2. For each replicate group, review:
  - Variant Call column – Do all replicates have the same call?
  - Confidence column – Are there any outliers within the replicate group? Do the values for the replicate group differ greatly from the confidence values for other replicate groups in the plate?
3. To view the fluorescence data for certain wells, select the rows in the Results table.
4. In the Results pane, click the **Variant Call** column header to sort the results according to the variant call. Scan the results for samples that were not assigned the same call as the wild type control.

### Omit outliers or change calls

After you review the software calls, you can omit outliers or change calls. Remember to click to reanalyze the data after you omit outliers or change calls.



## REAL-TIME (QUANTITATIVE) PCR

Real-time PCR, also known as kinetic PCR, qPCR, qRT-PCR and RT-qPCR, is quantitative PCR method for the determination of copy number of PCR templates such as DNA or cDNA in a PCR reaction. This method is used for a broad range of applications including quantitative gene expression analysis, SNP genotyping, copy number analysis, pathogen detection, drug target validation and for measuring RNA interference. Frequently, real-time polymerase chain reaction is combined with reverse transcription to quantify messenger RNA (mRNA) and microRNA (miRNA) in cells or tissues. RT-PCR or reverse transcription-polymerase chain reaction is a sensitive method for the detection of mRNA expression levels.

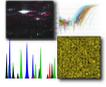
In real-time PCR, nucleic acid quantity is measured through the detection of a fluorescent signal proportional to the concentration of double-stranded DNA in the PCR tube. There are two flavors of real-time PCR: probe-based and intercalator-based. Both methods require a special thermocycler equipped with a sensitive camera that monitors the fluorescence in each well of the 96-well plate at frequent intervals during the PCR Reaction. Probe-based real-time PCR, also known as TaqMan PCR, requires a pair of PCR primers as regular PCR does, an additional fluorogenic probe which is an oligonucleotide with both a reporter fluorescent dye and a quencher dye attached. Intercalator-based method, also known as SYBR Green method, requires a double-stranded DNA dye in the PCR reaction which binds to newly synthesized double-stranded DNA and gives fluorescence. TaqMan method is more accurate and reliable than SYBR green method, but also more expensive. The quantity of the target sequence is then calculated automatically.

In quantification using the Standard Curve Method, one can quantify unknowns based on a known quantity. First a standard curve is created; then unknowns are compared to the standard curve and values are extrapolated.

Comparative Ct ( $\Delta\Delta C_t$ ) method is method for determining relative target quantity in samples. With this method amplification of the target and of the endogenous control in samples and in a reference sample is measured. Measurements are normalized using endogenous control. The software determines the relative quantity of target in each sample by comparing normalized target quantity in each sample to normalized target quantity in the reference sample.

## QUANTIFICATION OF ERBB2 mRNA BY REAL-TIME PCR

The *ERBB2* gene (also known as *HER2* or *neu*) encodes a 185-kDa transmembrane glycoprotein with intrinsic tyrosine kinase activity, and it has extensive sequence homology with the other members of the epidermal growth factor receptor family. Amplification/overexpression of *ERBB2* is present in 15–25% of breast cancers and is associated with a poor prognosis. Many studies have suggested that the poorer outcomes of such patients may be explained by a link between *ERBB2* protein and responses and/or resistance to both chemotherapy and endocrine therapy.



ERBB2 expression status could thus be a useful marker to predict the response to cancer therapy. *ERBB2* is also an outstanding candidate therapeutic target gene.

The protocol consists of three separate procedures: Isolation of total RNA from tissue, Reverse transcription and Real-time PCR assay.

## Isolation of total RNA using RNeasy Mini Kit, Qiagen

### **Materials needed:**

RNeasy Mini Kit (Qiagen, 74104)

Pipetors with range: 0 - 20  $\mu$ l; 0-100  $\mu$ l; 0-1000  $\mu$ l.

Eppendorf tubes 1.5 ml

Before starting:

Add 10  $\mu$ l  $\beta$ -mercaptoethanol per 1 ml Buffer RLT

Make sure ethanol has been added to Buffer RPE

70% Ethanol 50 ml

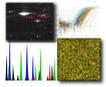
96.6% Ethanol 38.8 ml

Follow instructions for isolation of RNA using the RNeasy<sup>®</sup> procedure (Qiagen RNeasy Mini Kit, catalogue No 74104).

Perform all steps including centrifugations at room temperature. Ensure the centrifuge does not cool below 20°C. During the procedure, work quickly to minimise the risk of RNA degradation.

### **Procedure:**

1. Pour a few millilitres of liquid nitrogen into a mortar and drop a piece of frozen tissue. Wait a few seconds until the nitrogen is just about to evaporate fully and using a pestle crush the tissue into fine powder. If necessary, add more liquid nitrogen and repeat. Do not let thaw. Using a scalpel, quickly transfer powder to a 1.5 ml tube.
2. Immediately add **Buffer RLT** to tissue sample and homogenise further by pipetting up and down several times. Depending on extent of homogenisation, use 300  $\mu$ l to 4 ml buffer per 100 mg tissue. If using more than 600  $\mu$ l buffer per sample, distribute homogenate into 600  $\mu$ l aliquots.
3. Centrifuge the homogenate for 3 min at 12,000  $\times$  g (>13,000 rpm).
4. Transfer the supernatant to a fresh tube, add one volume of 70% ethanol and mix well by pipetting.
5. Transfer up to 600  $\mu$ l of the sample to an RNeasy spin column placed in a 2 ml collection tube. Centrifuge for 15 s at >8,000  $\times$  g (>10,000 rpm). Discard the flowthrough. Place the column back in the same collection tube. Load the rest of the sample onto the same RNeasy spin column discarding the flow-through after each centrifugation.
6. Add 700  $\mu$ l **Buffer RW1** to the RNeasy spin column. Close the lid gently and centrifuge for 15 s at >8,000  $\times$  g (>10,000 rpm) to wash the spin column membrane. Discard the flow-through and the collection tube. Transfer column with RNA into a clean 2 ml collection tube.



7. Add 500  $\mu\text{l}$  **Buffer RPE** to the RNeasy spin column. Close the lid gently and centrifuge for 15 s at  $>8,000 \times g$  ( $>10,000$  rpm) to wash the spin column membrane. Discard the flow-through. Place the column back in the same collection tube.
8. Once again, add 500  $\mu\text{l}$  **Buffer RPE** to the RNeasy spin column. Centrifuge for 2 min at  $>8,000 \times g$  ( $>10,000$  rpm) to wash and dry the spin column membrane. Discard the flow-through and the collection tube.
9. Place the column in a new 2 ml collection tube. Centrifuge at full speed for 1 min.
10. Transfer the RNeasy spin column to a 1.5 ml collection tube. Add 30-50  $\mu\text{l}$  RNasefree water directly to the spin column membrane. Incubate for 10 min at room temperature. Centrifuge for 1 min at  $>8,000 \times g$  ( $>10,000$  rpm) to elute the RNA.
11. Optional: Repeat step 10 using another 30-50  $\mu\text{l}$  RNase-free water and collect the eluate in the same collection tube.
12. Analyse RNA using Agilent 2100 Bioanalyzer (see instructions *Agilent Bioanalyzer*). Alternatively, analyse by agarose electrophoresis and quantify on UV spectrophotometer (see instructions *Spectrophotometers*)
13. Store RNA samples at  $-80^\circ\text{C}$  or proceed with reverse transcription.

## Reverse transcription

### **Materials needed:**

Pipetors with range: 0 - 20  $\mu\text{l}$ ; 0-100  $\mu\text{l}$

Eppendorf tubes 0.2 ml

### **Reagents:**

RNase-free water, 10xPCR Gold Buffer (ABI/Roche L00026-01), 25mM  $\text{MgCl}_2$  (ABI/Roche), 20mM dNTP mix (Solis BioDyne 02-31-00100), 500  $\mu\text{M}$  Random hexanuclotide mix (Amersham random priming kit RPN3540), RNase Inhibitor (20U/ $\mu\text{l}$ ) (ABI/Roche N808-0119) MuLV Reverse Transcriptase (ABI/Roche N808-0018, 50U/ $\mu\text{l}$ ).

Before starting:

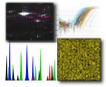
Random hexanuclotide mix preparation from Amersham random priming kit RPN3540

Concentration  $\sim 1\mu\text{g}/\mu\text{l}$  by absorption at  $A_{260}$  (NanoVue Spectrometer, GE Healthcare)

5  $\mu\text{M}$  Random hexamers: 198  $\mu\text{l}$  RNase-free water + 2 $\mu\text{l}$   $\sim 1\mu\text{g}/\mu\text{l}$  stock (concentration  $\sim 10$  ng/ $\mu\text{l}$  or 5 $\mu\text{M}$ )

Reaction mixture:

10xRb Gold Buffer	1.0 $\mu\text{l}$ (final concentration 1x)
25 mM $\text{MgCl}_2$	0.6 $\mu\text{l}$ (final concentration 1.5 mM)
20 mM dNTP	0.25 $\mu\text{l}$ (final concentration 500 $\mu\text{M}$ )



5 µM Randonne Hexanucleotide mix	1.0 µl (final concentration 500 nM)
RNase Inhibitor	0.4 µl (final concentration 8 U)
MuLV Reverse Transcriptase	0.8 µl (final concentration 40 U)
Total RNA	0.5-3.5 µl (200 ng)
RNase-free water	up to 10 µl total volume
Final Volume	10 µl

Program: 42 °C for 60' (incubation)

75°C for 10' (stop reaction)

4°C (for a short time before proceeding with real-time PCR or at -15°C- -25°C for storage).

Use 4 µl pre real-time reaction.

## Real-time PCR

### Materials needed:

Pipetors with range: 0 - 20 µl; 0-100 µl

microAmp™ Optical Adhesive Film

Optical reaction plate appropriate for your Real-Time PCR instrument

### Reagents:

TaqMan® Universal PCR Master Mix (Applied Biosystems, for example Cat. No.4304437)

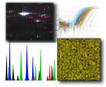
TaqMan® Gene Expression Assay (Applied Biosystems, Part No. 4331182, Assay Type: Erbb2 - hs01001580\_m1, 20x mix, probe dye – FAM, lot 799863; Human GAPDH - 4326317E-0908031, 20x mix, probe dye - VIC)

RNase-free water

### Procedure:

1. Design the layout of your plate. Example with 2 assays, Erbb2 and GAPDH, and 12 RT reactions from tissues 1 to 5, RT1-Erbb2 to RT6-Erbb2, RT1-R to RT6-R, is included below:

	1	2	3	4	5	6	7	8	9	10	11	12
A	Erbb2 RT-1	Erbb2 RT-1	Erbb2 RT-2	Erbb2 RT-2	Erbb2 RT-3	Erbb2 RT-3	Erbb2 RT-4	Erbb2 RT-4	Erbb2 RT-5	Erbb2 RT-5	Erbb2 RT-6	Erbb2 RT-6
B	GAPDH RT-1	GAPDH RT-1	GAPDH RT-2	GAPDH RT-2	GAPDH RT-3	GAPDH RT-3	GAPDH RT-4	GAPDH RT-4	GAPDH RT-5	GAPDH RT-5	GAPDH RT-6	GAPDH RT-6
C	NTC- Erbb2	NTC- GAPDH										



2. Mix the following on ice

	Single reaction		for 13 samples
H <sub>2</sub> O	5.0 µl		65 µl
20 × TaqMan® Gene Expression Assay	1.0 µl (final 1 ×)		13 µl
20 × TaqMan® Gene Expression Assay	1.0 µl (final 1 ×)	x 13	13 µl
RT-reaction	4 µl		
Final volume	20 µl		

Distribute 16.0 µl master mix into individual wells. Add 4.0 µl template/product from RT reaction or 4 µl H<sub>2</sub>O in NTCs.

3. In the 7500 real-time PCR system software, create an experiment or plate document using the following parameters:

#### Experiment Properties

Type of experiment: Quantitation – Comparative CT ( $\Delta\Delta CT$ )

Type of reagents: TaqMan® Reagents

Ramp speed: Standard

#### Plate Setup; Define Targets and Samples

Targets: Erbb2, GAPDH; assign to individual wells; define NTCs

Samples: RT1, RT2, RT3, RT4, RT5, RT6; assign to individual wells

#### Run Method

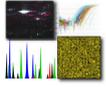
Thermal cycling conditions: Enzyme activation and initial denaturation at 95°C for 10 min followed by 40 cycles: 95°C, 15 s; 60°C, 60 s.

## DNA Methylation arrays

DNA Methylation is in the forefront of discovery in Epigenetics. Differential methylation is known to be associated with cancer, but recent studies have shown other complex diseases have been linked to methylation states. The Agilent Human DNA Methylation Microarray is designed to interrogate known CpG islands and two additional groups of methylated DNA. The array is specifically designed for analysis of methylated DNA derived from affinity-based isolation methods such as methylated DNA immunoprecipitation (MDIP).

Specification of the chip that are used for DNA Methylation arrays are following:

- 27,627 expanded CpG islands and 5081 UMR regions
- Probes annotated against UCSC HG18
- 237,227 biological probes, median probe spacing of 97bp
- 244K slide format printed using Agilent's 60-mer SurePrint technology



- One slide per kit. Slide contains one 244K microarray.

Methylated DNA lets investigators capture DNA sequences and study their relative methylation levels across an entire genome. The protocol requires an antibody to 5-methyl cytosine that will immunoprecipitate methylated DNA from a DNA sample.

The Methylated DNA Immunoprecipitation protocol consists of five general steps:

- DNA shearing
- Methylated DNA immunoprecipitation
- DNA labeling
- Microarray hybridization and washing
- Microarray scanning and storage

## **1. Chromatin immunoprecipitation (ChIP)**

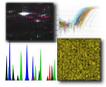
### **Goal**

Use selective antibody bound to magnetic beads to specifically capture the Methylated DNA.

1. Mix antibody bound to magnetic beads (Dynal) with DNA sample.
2. Place at 4°C overnight on a rotating platform.
3. Isolate the beads containing the antibody bound to the methylated DNA.
4. Wash 2 times with buffer to remove non-specific contaminants.
5. After the wash, heat the complexes for a few minutes with detergent to elute the methylated DNA from the antibody and beads.

### **Key variables**

- Beads, type, and quantity
- Time
- Temperature
- Immunoprecipitation buffer, volume, and composition



- Wash buffer composition
- Number of washes

### QC Metrics

After the DNA is isolated, you can run the Bioanalyzer or gel equivalent to assess fragment size after shearing. Use 1  $\mu$ L for the Bioanalyzer, or 5  $\mu$ L on a 1.5% agarose gel.

**Notes:** Magnetic beads coated with protein G are routinely used due to their ease- of- use and ability to bind a variety of antibodies. Other coatings (e.g. protein A) and bead types (e.g. agarose) are available but have not been validated by Agilent.

## 2. DNA isolation

### Goal

Purify DNA from associated proteins and RNA and protein contaminants.

1. Purify the DNA via organic extraction and ethanol precipitation.

### Key variables

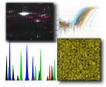
- Temperature
- Time
- SDS concentration

## 3. DNA labeling

### Goal

Incorporate fluorescent- tagged nucleotides into the IP and reference DNA for hybridization.

1. Use the Agilent Genomic Enzymatic Labeling Kit and cyanine dyes.
2. Do 1 Labeling reaction for each sample. Approximately 1  $\mu$ g input per reaction for the reference channel Cyanine 3, and all IP DNA for the Cyanine 5.
3. Anneal random primers to the DNA.
4. Extend primers using high concentration exo- Klenow enzyme and fluorescent- labeled nucleotides.
5. Purify labeled DNA using the Amicon 30 kDa columns.



### Key variables

- Reaction size
- Reagent quantity (input DNA material, Cy dye, enzyme) per reaction

### QC Metrics

Nanodrop measurement of total DNA yield (expect >2.5 µg per reaction); Nanodrop measurement of pmol/µL dye (expect >2 pmol/µL with Cy5- dUTP and >3 pmol/µL Cy3- dUTP).

## 4. Microarray hybridization and washing

### Goal

Hybridize material to and wash excess/nonspecific material from Agilent 60-mer oligo arrays to yield low background and high signal (“flat” background with high peaks)

1. Hybridize for 40 hours at 67°C in hybridization oven rotating at 20 rpm. Hybridization buffer contains a proprietary wetting agent (that keeps bubbles moving freely), approximately 5 µg labeled DNA per channel (10 µg total) and competitor nucleic acids.
2. Wash slides in a series of two buffers with an optional 3rd wash that contains ozone-scavenging reagents to help prevent premature dye degradation.

### Key variables

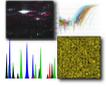
- Hybridization duration
- Quantity of labeled material
- Temperature
- Type and quality of detergent
- Type and quantity of nucleic acid competitors

## 5. Microarray scanning and storage

### Goal

Extract data from microarray; store microarray for possible future analysis

1. Use default settings on Agilent scanner.
2. Store used slides in N2 box.



## Microarray Data Analysis

### The R environment

R is an integrated suite of software facilities for data manipulation, calculation and graphical display. Among other things it has

- \_ an effective data handling and storage facility,
- \_ a suite of operators for calculations on arrays, in particular matrices,
- \_ a large, coherent, integrated collection of intermediate tools for data analysis,
- \_ graphical facilities for data analysis and display either directly at the computer or on hardcopy, and
- \_ a well developed, simple and effective programming language (called ‘S’) which includes conditionals, loops, user defined recursive functions and input and output facilities.

### AgiMicroRNA

AgiMicroRna provides useful functionality for the processing, quality assessment and differential expression analysis of Agilent microRNA array data. The package uses a limma-like structure to generate the processed data in order to make statistical inferences about differential expression using the linear model features implemented in limma. Standard Bioconductor objects are used so that other packages could be used as well. AgiMicroRna reads into R the scanned data exported by the Agilent Feature Extraction (AFE) image analysis software. Standard graphical utilities can be used to evaluate the quality of the data. AgiMicroRna includes a full data example that can be loaded into R in order to illustrate the capabilities of the package.

The Human microRNA microarray v3.0 contains 866 human and 89 human viral microRNAs. The array contains also a set of positive and negative controls that are replicated a different number of times. For the statistical analysis we need an estimate of the expression measure for every microRNA that has to be normalized between arrays. This processed signal is going to be used to make statistical differences about the differential expression.

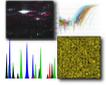
AgiMicroRna has been primarily designed to produce a processed data to be analyzed using the limma package. First, a target file is needed in order to assign each scanned data file to a given experimental group. The target file is a tab-delimited text format file created by the user where we specify the factors that are going to be included in the statistical model.

An example for miRNA microarray analysis of breast cancer patients, using R and AgiMicroRNA package, with certain modifications is presented. The commands are listed below:

```
AnalyseMA <- function(in_file, in_vector, out_file)
```

This function have three input parameters, input file for analysis, vector for generating design matrix, and output file for writing results;

```
library("AgiMicroRna")
```



**#importing the "AgiMicroRna" library;**

```
targets=readTargets(infile=in_file,verbose=TRUE)
```

**# this command reads the input file;**

```
dd<-
```

```
read.maimages(targets,source="agilent",columns=list(Rf="gTotalGeneSignal",Gf="gTotalProbeSignal",Gb="gProcessedSignal"),  
other.columns=list(IsGeneDetected="gIsGeneDetected",IsSaturated="gIsSaturated",IsFeatNonUnifOF="gIsFeatNonUnifOL",IsFeatPopnOL="gIsFeatPopnOL",BGKmd="gBGMedianSignal"),  
annotation=c("ControlType",  
"ProbeName","GeneName"),verbose=TRUE,sep="\t",quote="")
```

**# "limma" package function that reads the data from the files that are named in the input file,**

**columns that are imported are: gTotalGeneSignal, gTotalProbeSignal, gProcessedSignal, gIsGeneDetected, gIsSaturated, gIsFeatNonUnifOL, gIsFeatPopnOL, gBGMedianSignal, ControlType, ProbeName, GeneName);**

**# NOTE: In the files that are analysed the column SystematicName should be changed to GeneName;**

```
ddu=new("uRNAList")
```

**# New AgiMicroRna functions use "uRNAList" class, read.maimages returns "RGList". The next commands are for this conversion;**

```
ddu$TGS=dd$R  
ddu$TPS=dd$G  
ddu$procS=dd$Gb  
ddu$targets=dd$targets  
ddu$genes = dd$genes[, c(1, 2, 3)]  
ddu$other$IsGeneDetected=dd$other$IsGeneDetected  
ddu$other$gIsGeneDetected=dd$other$gIsGeneDetected  
ddu$other$gIsSaturated=dd$other$gIsSaturated  
ddu$other$gIsFeatNonUnifOL=dd$other$gIsFeatNonUnifOL  
ddu$other$gIsFeatPopnOL=dd$other$gIsFeatPopnOL  
ddu$other$BGKmd=dd$other$BGKmd
```

```
TGS=tgsMicroRna(ddu,half=TRUE,makePLOT=TRUE,verbose=TRUE)
```

**# function for generating "uRNAList" of the "gTotalGeneSignal" and plotting the signal before normalization;**

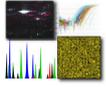
```
ddNORM=tgsNormalization(TGS,"quantile",makePLOTpre=TRUE,makePLOTpost=TRUE,targets,verbose=TRUE)
```

**# normalization of the "gTotalGeneSignal" using quantile approach, and plotting the normalized signal;**

```
proc=filterMicroRna(ddNORM,ddu,control=TRUE,IsGeneDetected=TRUE,wellaboveNEG=FALSE,limIsGeneDetected=75,limNEG=25,makePLOT=FALSE,targets,verbose=TRUE,writeout=FALSE)
```

**# function for filtering the normalized data, removing the non-detected genes, that are expressed below given threshold;**

```
eset=esetMicroRna(proc,targets,makePLOT=FALSE,verbose=TRUE)
```



```
# function for creating the expression set object, generated from the filtered data and the input file;

# the next command are used for generation of comparison model between normal and mutant
subjects;
levels.treatment=levels(factor(targets$Treatment))
treatment=factor(as.character(targets$Treatment), levels =
levels.treatment)
levels.subject=levels(factor(targets$Subject))
subject=factor(as.character(targets$Subject), levels = levels.subject)

desig=model.matrix(~-1+factor(in_vector))
# creation of the design matrix;

colnames(desig)=c("normal", "cancer")

fit=lmFit(eset, desig)
# fitting the expression set against the design matrix to obtain the treatment estimates for each
microRNA;

print(head(fit$genes))

contr=makeContrasts(cancer-normal, levels=desig)
# generating contrast matrix;

fit2=contrasts.fit(fit, contr)
fit2=eBayes(fit2)
# estimation of the contrasts using moderate Bayes statistics;

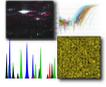
dt=getDecideTests(fit2, DEmethod="separate", MTestmethod="BH", PVcut=0.05, verb
ose=TRUE)
# the "getDecideTests" is the final command that selects the significant genes, summarizing the up
and down genes according to the given p-value;

write.fit(fit2, adjust="BH", file=out_file)
# writing the final results to output file;

significantMicroRna(eset, proc, targets, fit2, contr ,dt, DEmethod =
"separate", MTestmethod = "BH", PVcut = 0.5, Mcut = 0, verbose = TRUE)
# generation of text and html files of the genes and their relevant statistics.
```

The function `significantMicroRna` summarizes the results of the differential expression analysis extracting information from the `MArrayLM` and `TestResults` objects generated by the `limma` functions. `significantMicroRna` creates a list containing the genes with their relevant statistics.

When multiple contrasts are done, the method for the selection of the significant genes can be either separated or nestedF. When separated is plugged in into the `significantMicroRna` function then a list including all the genes that have been analyzed is given for each contrast. These lists contain the statistics obtained from the differential expression analysis. The information that is given for each gene is shown in Table below.



**Table 1:** Statistics given by significantMicroRna for separate

Variable	Data
PROBE	Probe name (one of the probes interrogating the gene)
GENE	microRNA name
M	Fold change
A	Mean of the intensity for that microRNA
t	moderated t-statistic
pval	p value of the t-statistic
adj.pval	p value adjusted by MTestmethod
fdr.pval	p value adjusted by fdr