

# Development of New Statistics / Bioinformatics Methods for the Detection of CNVs

**Mireia Vilardell**  
**Genetics Unit**  
**Pompeu Fabra University**



# INTRODUCTION

- **Dosage changes in DNA** are usually related to **diseases** (i.e. Down Syn, Prader-Willi, Angelmann, WBS, DiGeorge,...).
- Segmental Duplications are **hotspot regions** for non-allelic homologous recombination.
- The term Copy Number Variations (**CNV**) is applied to define copy changes among **healthy subjects** from the same population.

# OBJECTIVES

- To **develop array-based technologies**:
  - to detect CNV and dosage changes related to diseases with unknown ethiology
  - to use in genetic diagnosis

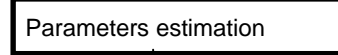
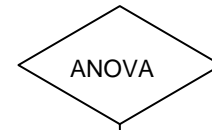
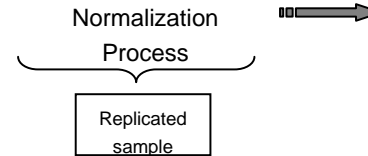
# MATERIAL AND METHODS

## SOURCES OF VARIATION

**Optimal Lab  
Conditions**

**Image  
Analysis**

**QUALITY AND RELIABILITY**



To get the theoretical normal distribution for each probe

Assess residual normality  
Assess homocedasticity

- Loess + BK subtraction
- Loess - BK subtraction
- Print-tip loess + BK subtraction
- Print-tip Loess - BK subtraction
- Loess loc + BK subtraction
- Loess Loc - BK subtraction
- Loess Loc scale + BK subtraction
- Loess loc scale - BK subtraction

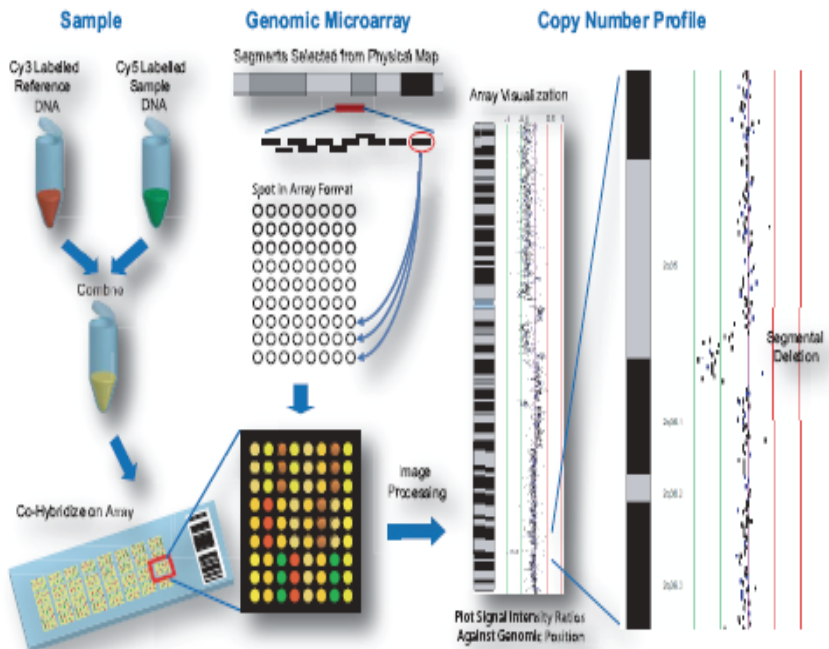
The best normalization procedures criteria:

- \*Min slide variability
- \*Min  $n^0$  clones affected by DB effect
- \*Min absolute values of DB
- \*Automatic procedure

Patient sample is used to preserve the true biological signal:

- \*Max sensitivity
- \*Max positive validation ratio
- \*Min false positive ratio

## DOSAGE CHANGE DETECTION



# MAIN RESULTS

## SOURCES OF VARIATION

Systematic dye problems exist that involve:

- Spatial distribution
- Intensity
- Probe content: GC, CpG, Alu and SD

Print tip without BK subtraction corrects better these artefacts

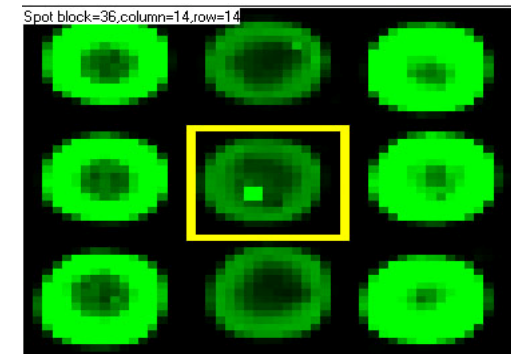
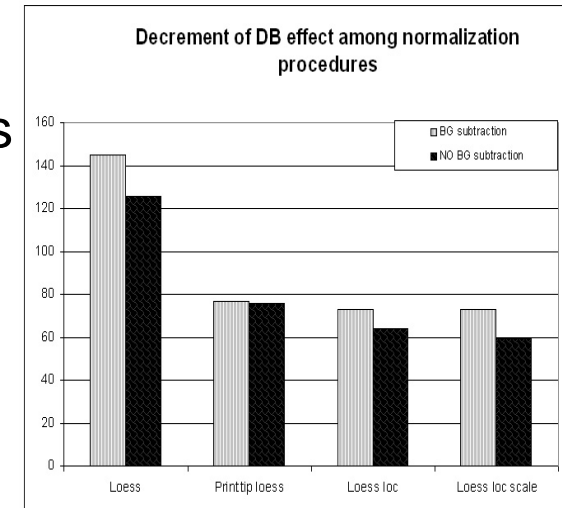
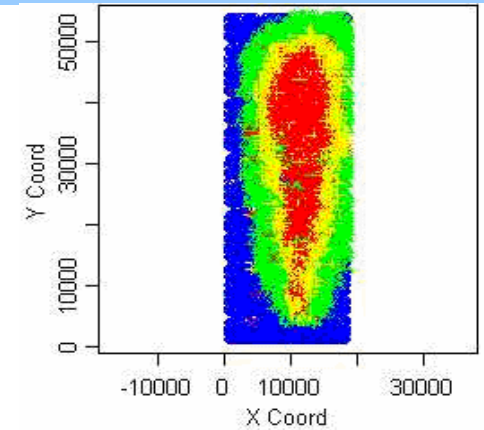
## QUALITY AND RELIABILITY

False positive results are associated to:

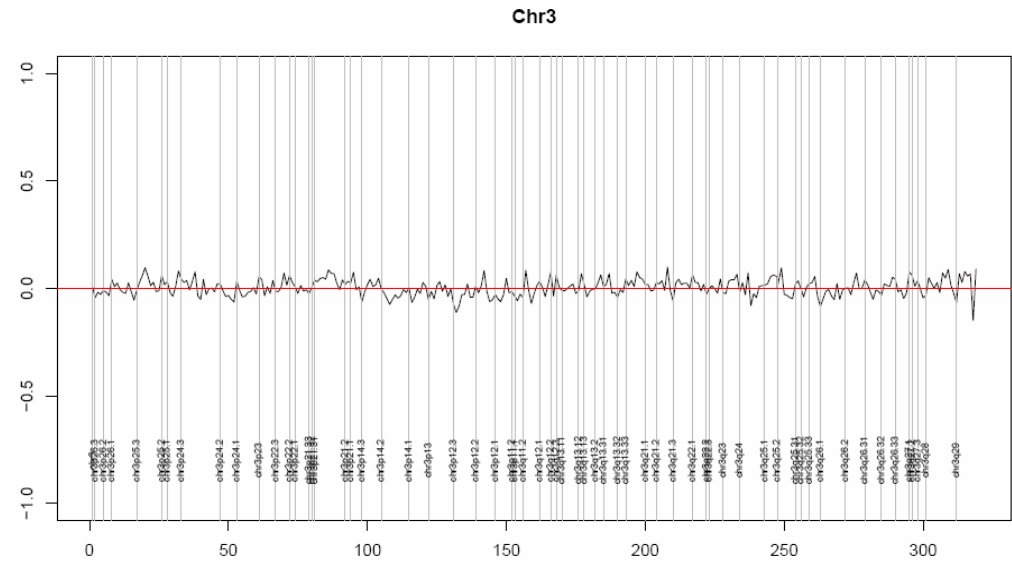
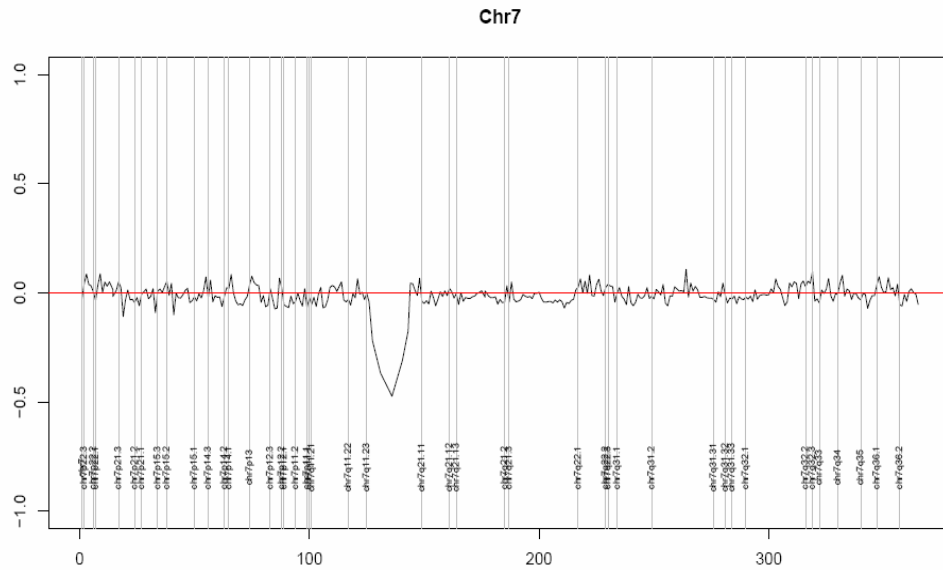
- Artefacts in spot image
- DNA quality

The best range for 260/280 is between 1.6 to 1.85

The best range for 260/230 is less than 2.0



# GROUP DOSAGE CHANGE DETECTION

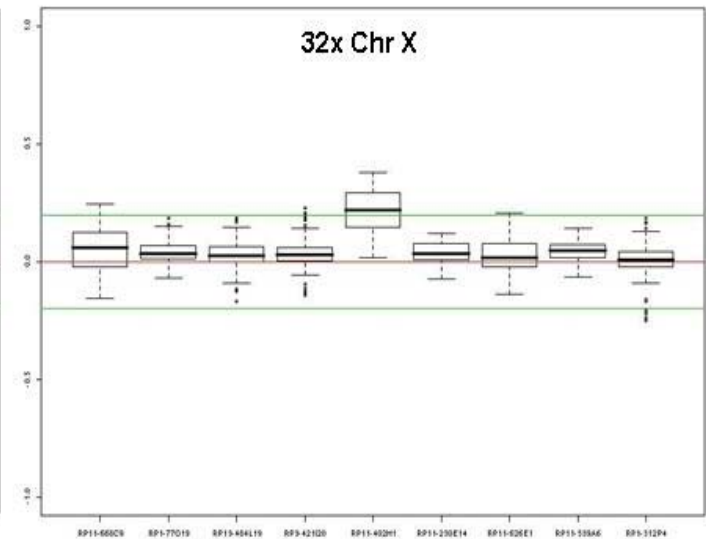
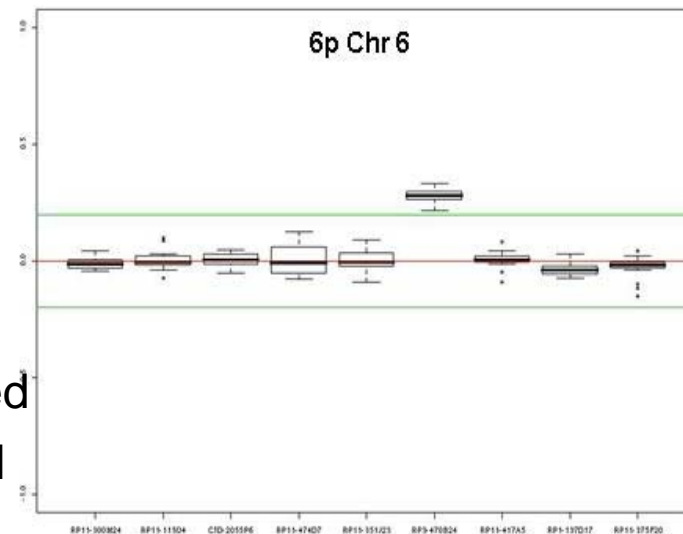


Different models allow to detect changes involving disease groups of patients or single clone changes

# SINGLE DOSAGE CHANGE DETECTION

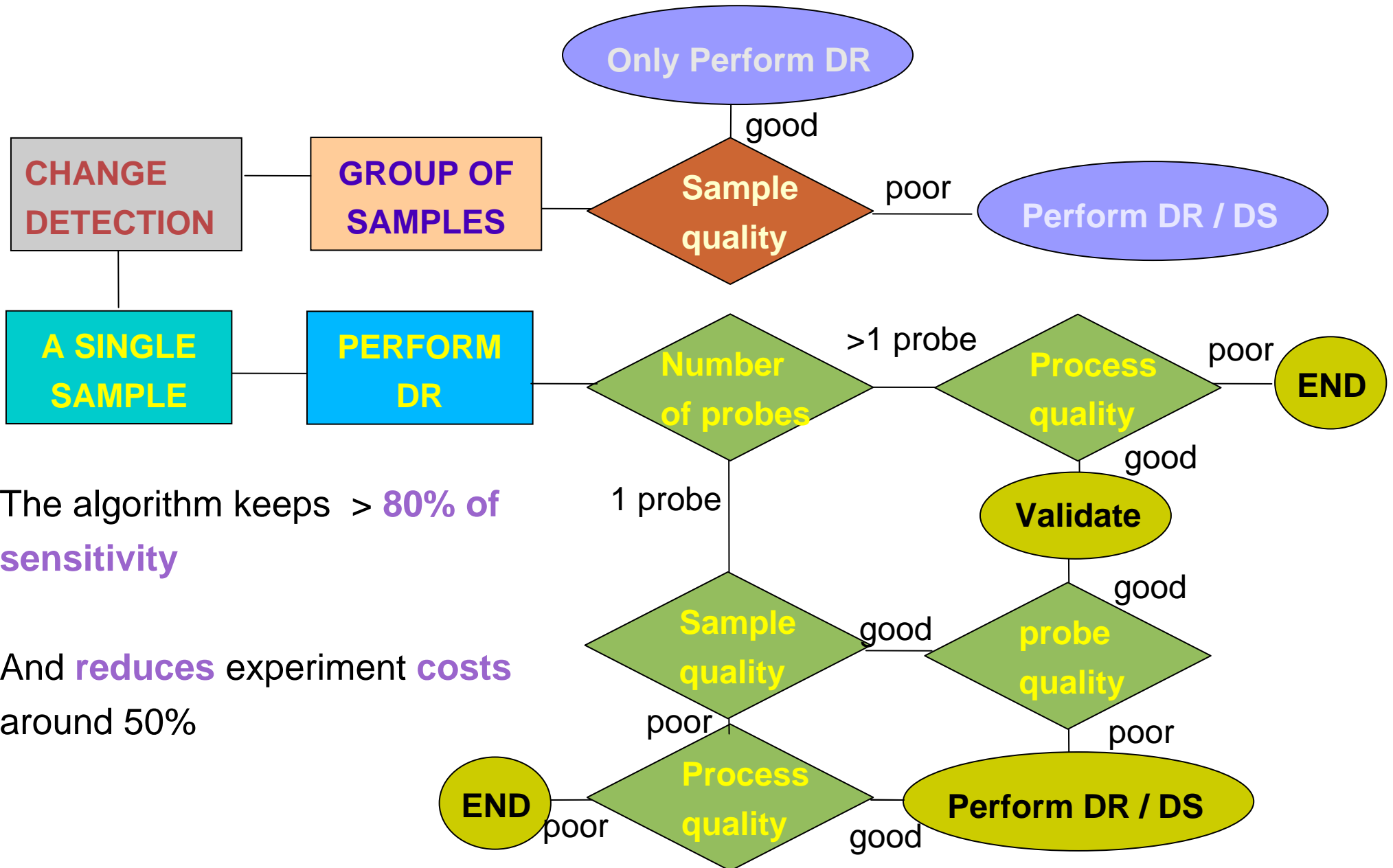
Sensitivity for fully-altered probes is > 80%

Sensitivity for partially-altered probes (min 30%) is around 50%



# CONCLUSIONS

- (1) There is a **good correlation** between technical and biological replicates.
- (2) Data from **different experiments** could be analyzed together

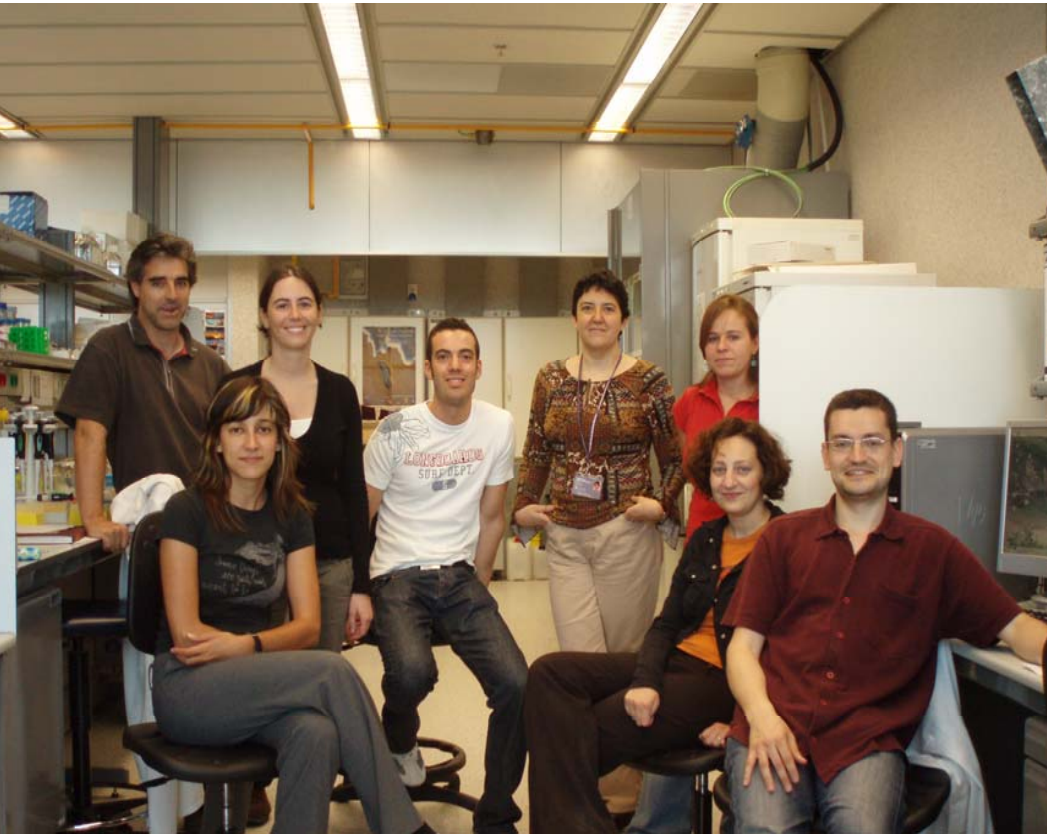


The algorithm keeps **> 80% of sensitivity**

And **reduces** experiment **costs** around 50%

THANKS FOR SHARING WITH ME YOUR KNOWDELEDGE!!!!

GENETICS UNIT FROM UPF



MICROARRAY'S LAB FROM CRG

