



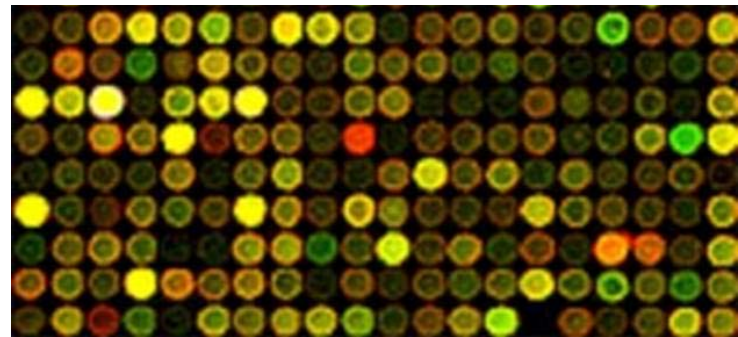
# Tissue Microarray in GRID

Federica Viti

Institute for Biomedical Technologies, CNR, Milan, Italy

University of Genoa - DIST

- TMA Introduction
- GRID Overview
- TMA on GRID
- Results
- Conclusion
- Acknowledgement



Genetic Diseases

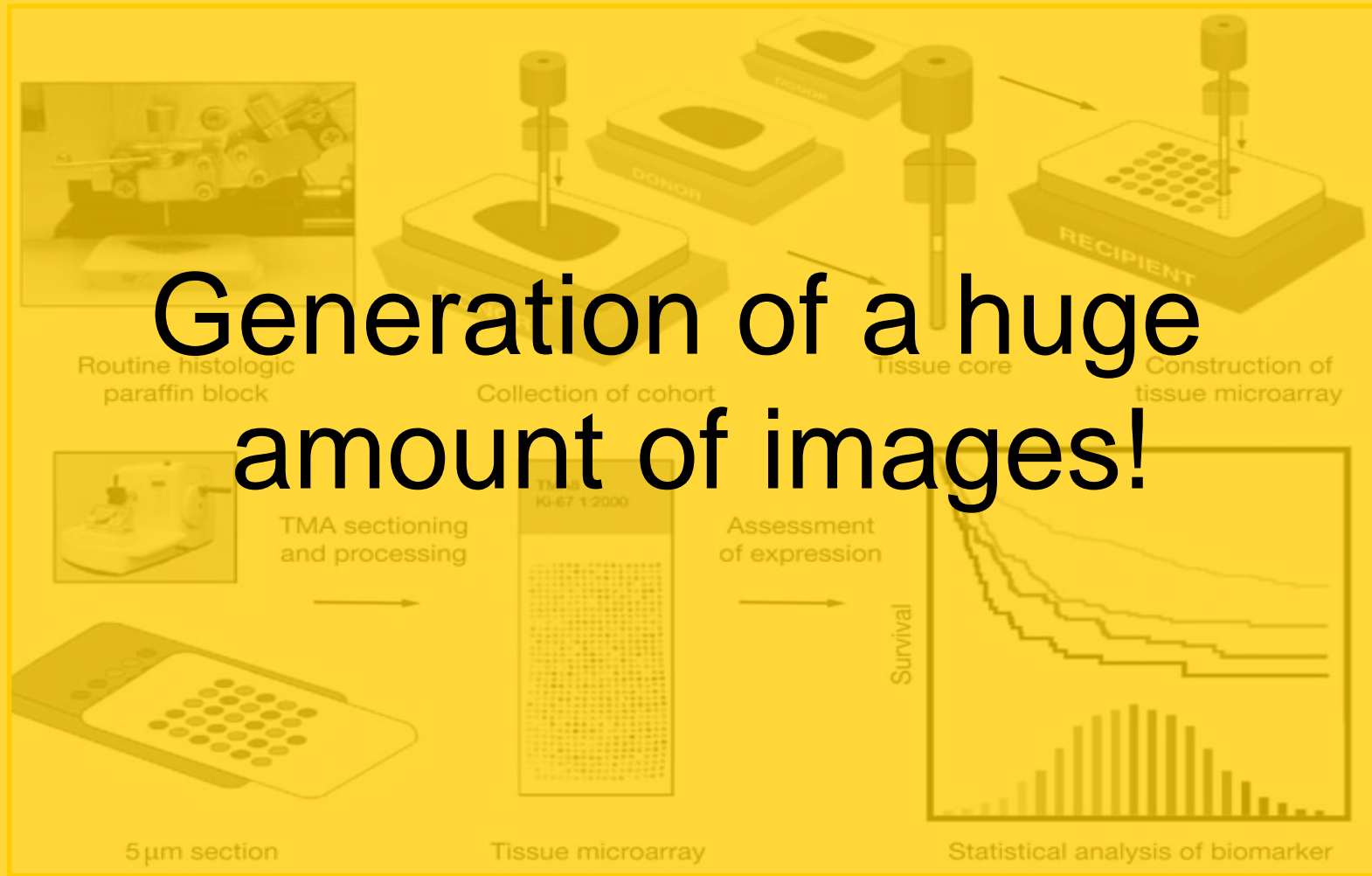
```
graph TD; A[Genetic Diseases] --> B[High throughput techniques (i.e. DNA microarray) to screen the whole genome]; B --> C[Low reliability]; C --> D[Validation through TMA];
```

High throughput techniques (i.e. DNA microarray)  
to screen the whole genome

Low reliability

Validation through TMA

## Genes and proteins detection





## Possible *analysis*:

- **ImmunoHistoChemistry (IHC)**: reactions on biological tissues using immunoglobulins
- **In Situ Hybridization (ISH)**: uses oligonucleotides probes to highlight the presence of nucleic acids (in particular RNA transcripts) 'in situ'
- **Fluorescent In Situ Hybridization (FISH)**: for searching genes on chromosomes through antibodies coupled with fluorochromes



To be informative each image must be associated to additional data:

- *patient*: patient\_id, age, gender, ethnicity, risk\_factors, familiarity, therapy, etc.
- *sample*: surgical\_operation\_date, disease, organism\_part, weight, conservation\_method, etc.
- *block*: general\_aim, donor\_block\_id, etc.
- *spot*: image\_id, coordinates, analysis, etc.

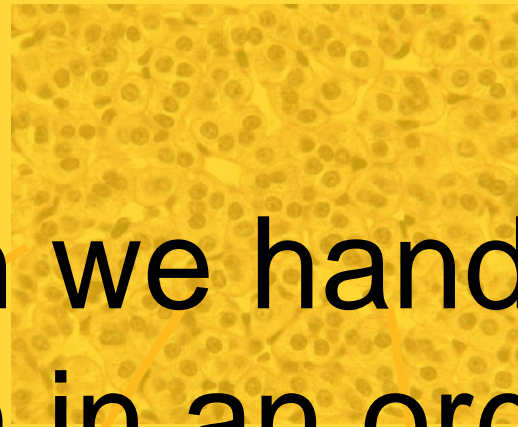
Metadata = data which describe data



# TMA Requirements

Parallelization => Decrease of time, costs, tissues

BUT this technique requires



How can we handle such information in an organic way?

Large Storage

Elaboration -> Large Computing Resources

Data Security

Metadata



EGEE Grid can help researchers in handling TMA experiments because relies on:

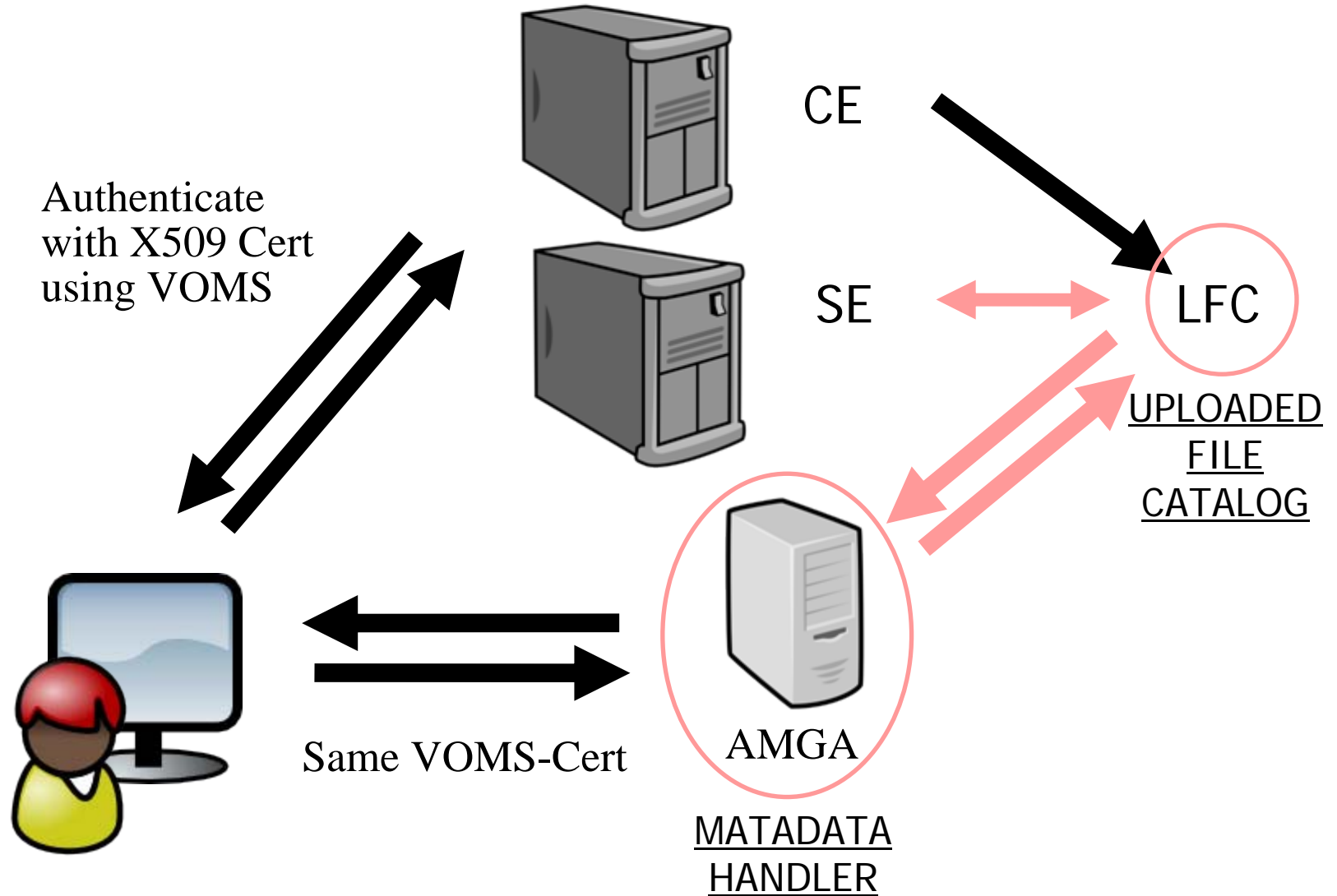
- CE e WN (computational resources)
- SE e LFC (for distributed storage)
- Security protocols (keys and codes)
- API GFAL: this library allows a file elaboration on remote SE
- AMGA (metadata handling)



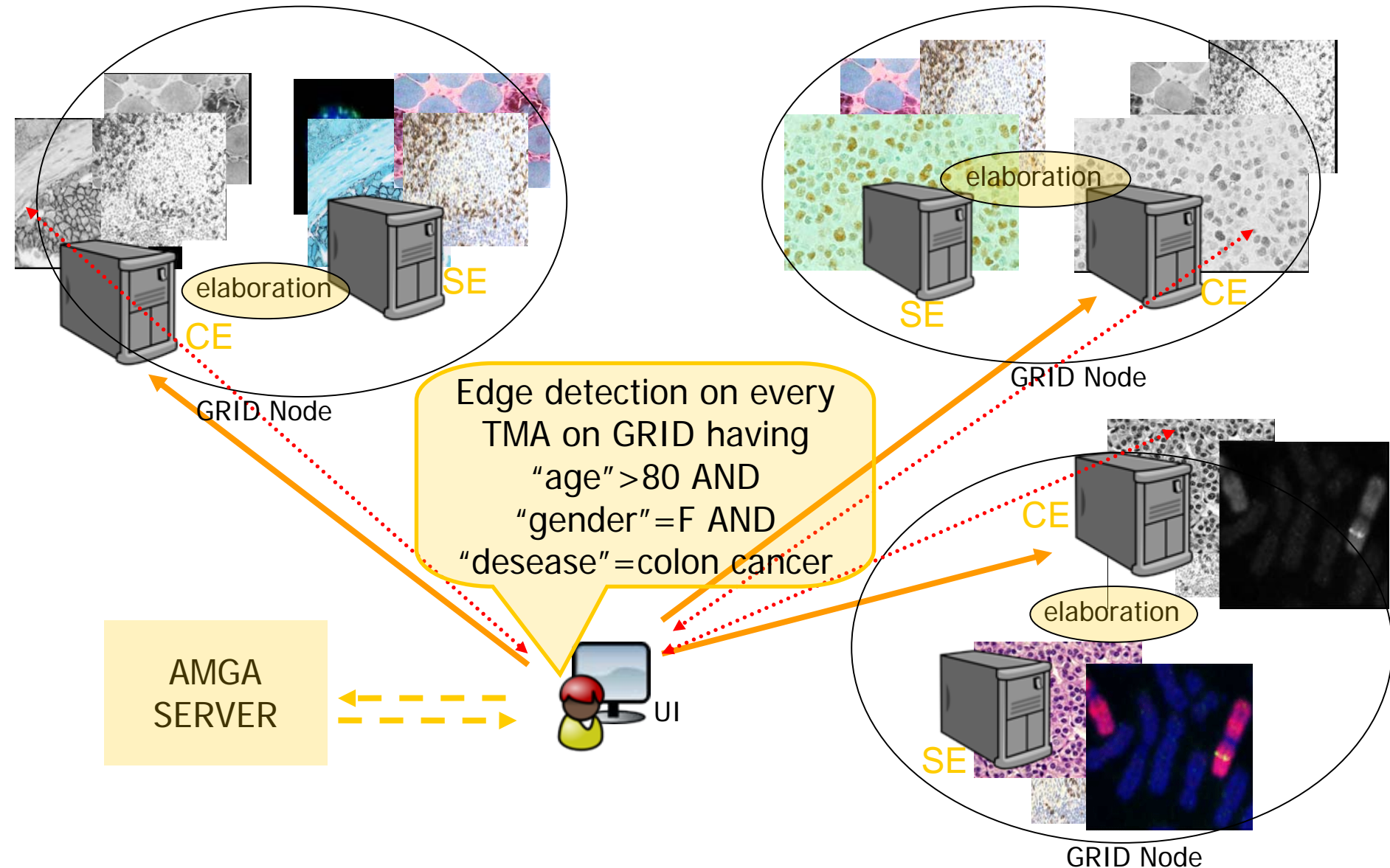
- ⇒ **PIMA(GE)<sup>2</sup>**: image elaboration library Image Algebra oriented, developed by Dr. A. Clematis staff.
- ⇒ We ported the library in GRID applying the **I/O GFAL** (Grid File Access Library) functions (`gfal_read`, `gfal_write`) to allow basic file handling on remote SE.
- ⇒ The security of data is guaranteed because GFAL file access is implemented **without copying** the file on machines different from the storage where owner stored it: in this way also sensible data (like TMA patient information) can be accessed without decreasing privacy.

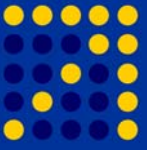


- ⇒ Our application needs a huge amount of metadata.
- ⇒ The **A**rda **M**etadata **G**rid **A**pplication (**AMGA**) is the only way available at the moment to store metadata on Grid.
- ⇒ AMGA is organised in schemas (= directories) and entries (=files) in a hierarchical way. It is accessible through a SQL-like query language.
- ⇒ AMGA client can be installed on UI to access an AMGA server.
- ⇒ Metadata are joined to related data (stored in the LFC) through a key, which is the GUID of the data file (created on the LFC catalog system and copied on an AMGA table).
- ⇒ It preserves security of data because it supports VOMS authentication.

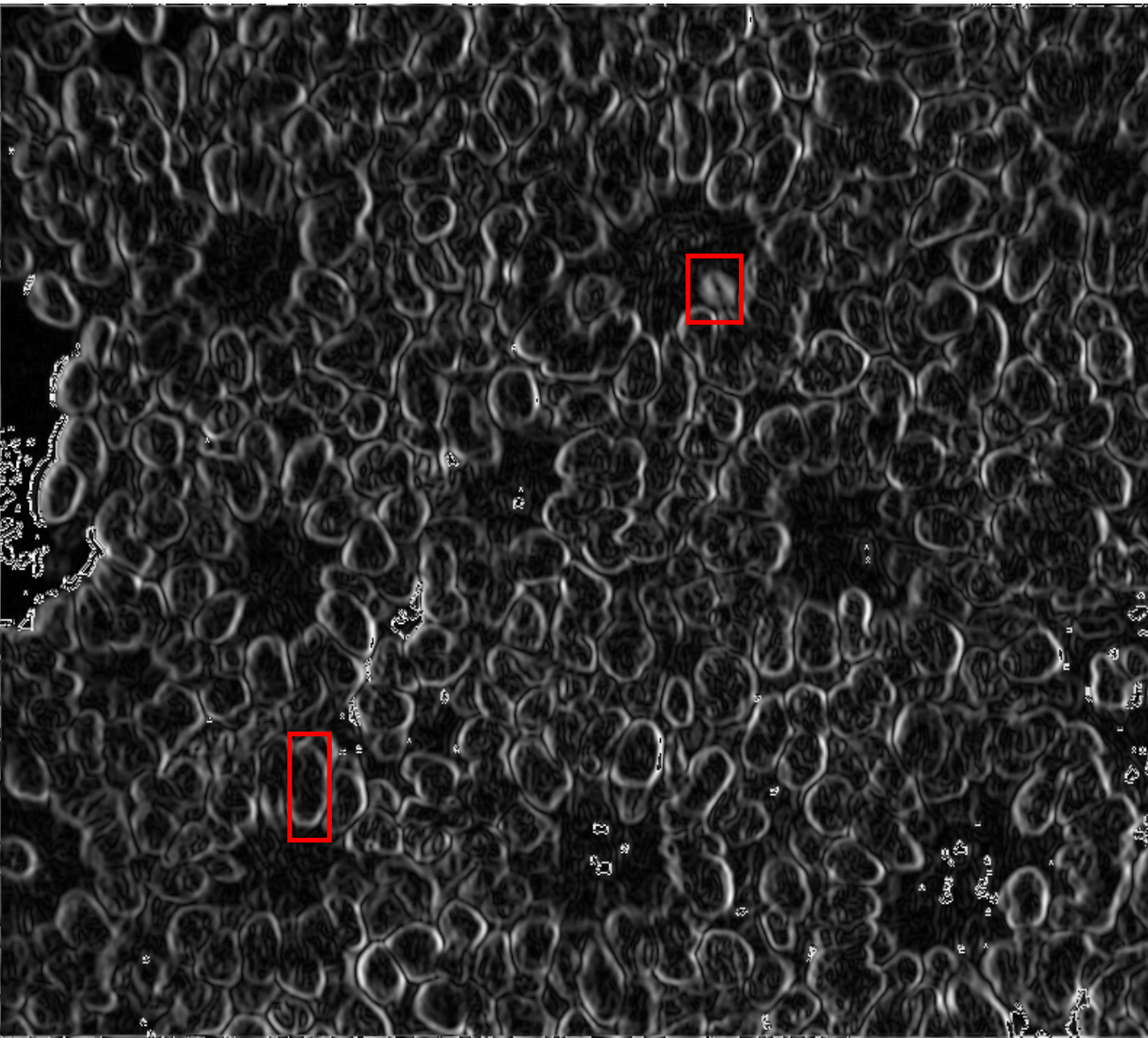


# TMA on GRID



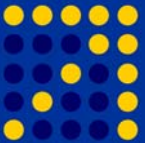


# Results



In cancer colon cells tend to change their shape becoming longer and going nearer one another.

For these reasons it is important the use of image analysis to identify **shape** and **position** of cells.



- TMAs are useful for pathologists because allow the screening of hundreds of human tissues in a parallel way, speeding up studies on genetic diseases (i.e. evaluations of new tumoral markers)
- TMA technology needs large storage and computing to handle the amount of produced data and a sicure structure to preserve patient privacy => GRID platform



# Acknowledgments

- Thank you very much to:  
my supervisor Luciano Milanese  
Ivan Merelli,  
Daniele D'Agostino,  
Antonella Galizia,  
Andrea Clematis,  
for their precious work and support.
- This work has been supported by the European Project N.026808 – Bioinfogrid
- University of Genoa – DIST and CNR-Bioinformatics