Goals and accomplishments from WP10 point of view

J. Montagnat, CNRS, CREATIS
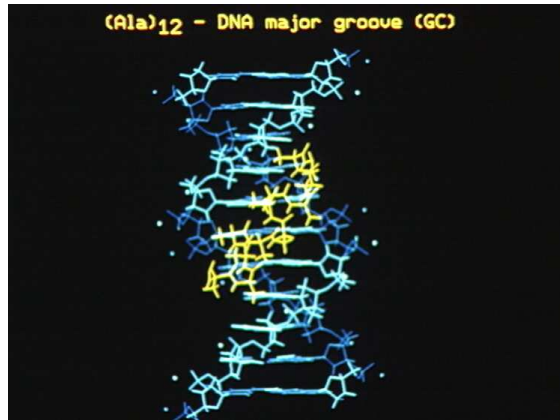
V. Breton, CNRS/IN2P3

Initial goals
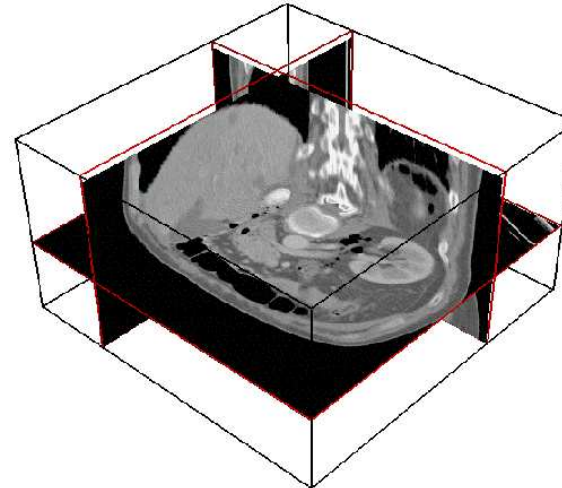

- **Make a list of requirements for biomedical grids (Task 10.1 ending PM6)**
  - Genomics   and   medical imaging

- **Identify and characterize a first biomedical application to deploy on DataGrid testbed (Task 10.2 ending PM12)**

- Deploy this first application on DataGrid testbed2 and testbed3 (Task 10.3 ending PM 36)

Challenges

- The biomedical community has NO strong center of gravity in Europe
  - No equivalent of CERN (High-Energy Physics) or ESA (Earth Observation)
  - Many high-level laboratories of comparable size and influence without a practical activity backbone (EMB-net, national centers,…) leading to:
    - Little awareness of common needs
    - Few common standards
    - Weak common long-term investment

- The biomedical community is very large

- The biomedical community is often distant to computer science issues

The WP10 policy during 2001

⊙ Because of the specific challenges, WP10 initial goals were less ambitious than the goals of the other application working groups (WP8-9) at the beginning of the project

⊙ We chose however to follow closely their path and rhythm :
  ○ To demonstrate ASAP to the biomedical community that grids are useful for genomics and medical imaging
  ○ To convince more people to join the Working Group
  ○ To participate actively to the design of the grid architecture and services from the beginning

# WP10 first technical accomplishments

- ⦿ Deliverables on time
  - ○ D10.1 « Requirements for grid-aware biology applications" in August 2001
  - ○ D10.2 "Grid-aware biomedical applications for DataGrid testbed assessment" in January 2002
  - ○ 7 applications identified (5 in genomics, 2 in medical imaging)

- ⦿ Demonstrator « DataGrid for human health » (shown at IST2001)
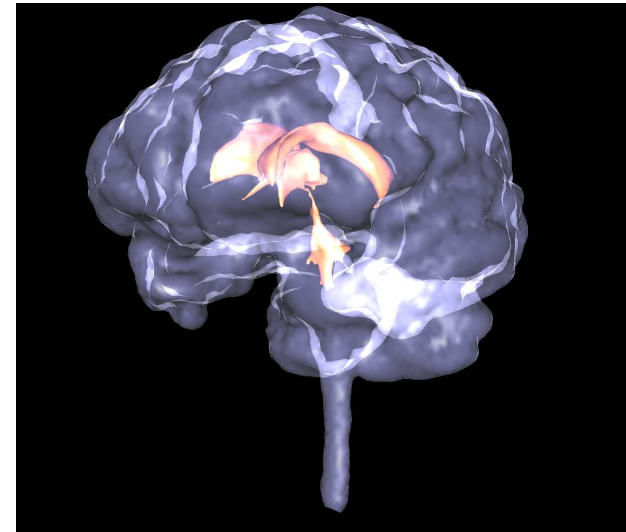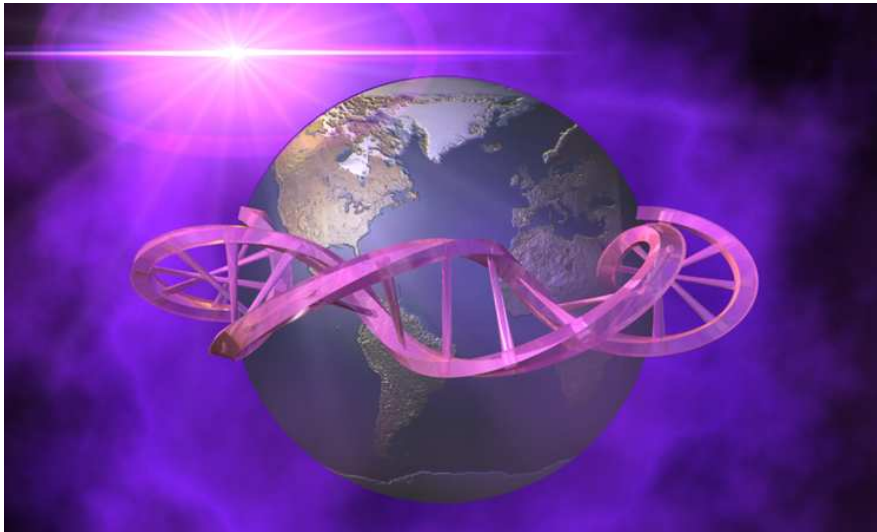
WP10 first technical accomplishments

◉ Creation of the Biomedical Virtual Organization on DataGrid testbed1

◉ « Gridification » of a first representative algorithm for genomics comparative analysis

    ○ See demo this afternoon
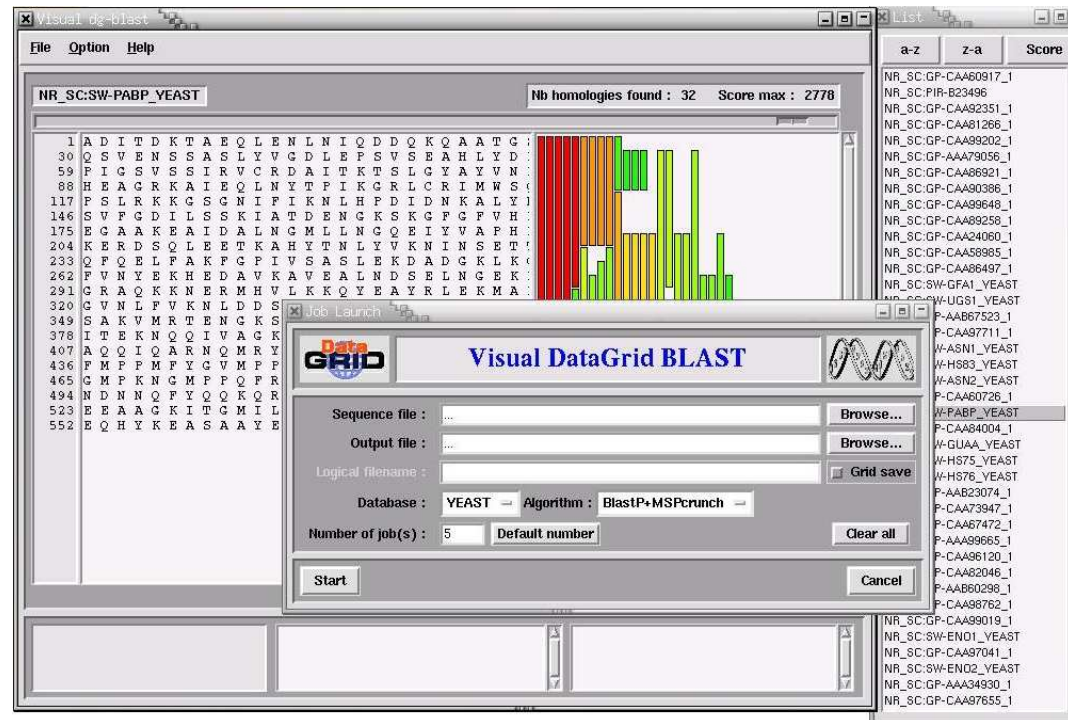
◉ Participation to testbed1 testing

3 layers for Biomedical applications

WP10 applications	Cooperative Framework
	Grid Service Portals
	Distributed Algorithms
EDG Middleware	

- **Distributed Algorithms**. New distributed "grid-aware" algorithms (c.f. this afternoon's demonstration).

- **Grid Service Portals**. Service providers taking advantage of the DataGrid computational power and storage capacity.

- **Cooperative Framework**. Use the DataGrid as a cooperative framework for sharing resources, algorithms, and organize experiments in a cooperative manner.

Biomedical requirements

WP10

- ◉ **Large user community**
  - ○ anonymous/group login
- ◉ **Data management**
  - ○ data updates and data versioning
- ◉ **Security**
  - ○ disk / network encryption
- ◉ **Limited response time**
  - ○ fast queues

- ◉ **High priority jobs**
  - ○ privileged users
- ◉ **Interactivity**
  - ○ communication between user interface and CE's
- ◉ **Parallelization**
  - ○ MPI site-wide / grid-wide
- ◉ **Pipeline processing**
  - ○ pipeline description language / scheduling

These requirements were discussed with Middleware Work Packages : some should be addressed by testbed2.

Conclusions

- **Key achievements**
  - We have started to demonstrate the relevance of the grid concept for biomedical applications (see demo)
  - We have established a fruitful dialog with
    - Biology and medical imaging (25 laboratories involved)
    - The other applications and middleware work packages of DataGrid
  - The group has found cohesion and coherence despite the various backgrounds of the actors (biologists, computer scientists, physicists)

- **Key challenges for the lifetime of the DataGrid project**
  - Deploy a true biomedical testbed involving genomics, bio-informatics and medical imaging laboratories (5 sites on testbed1)
  - Run the applications described in D10.2 to demonstrate the use of a grid in genomics and medical imaging

The growing interest of the biomedical community …