DS5: Biological Storage and Retrieval – Bringing e-Science to Embryo Bioinformatics resources.

Mark K. Scott[1], Jano van Hemert[2], Yin Chen[2], Xunxian Wang[1], Steven Lisgo[1], Susan Lindsay[1], Malcolm Atkinson[2] and Richard A. Baldock[3].

[1] Institute of Human Genetics, Newcastle University; [2] National e-Science Centre, Edinburgh; [3] MRC Human Genetics Unit, Edinburgh.

Introduction

With the numerous changes in technologies (lab-based and *in silico*) for gathering information, there is a notable increase in the amount of data which needs to be stored. As part of the FP6 funded DGEMap project we are continuing to assess current architecture and further requirements associated with management, curation and analysis of large scale biological data sets, in particular gene expression data and morphological mapping. The collaborative effort has so far looked to focus on three main areas: a) Storage and access of internal experimental details and their associated results; b) 3D mapping and visualisation tools to increase both the accuracy and usefulness of disseminated data; and c) Establishing a web presence for the wider community complete with search and analysis tools.

Data Storage and Access:

Previous implementations of databases for data storage within our laboratory were not ideal and had little scope for development. After careful consultation with the laboratory staff who would use this, a prototype database has been set-up for the storage and easy access of internal laboratory details containing sample, experiment and collaborator information. The new prototype uses increased security and portability while offering easier scalability for future development.

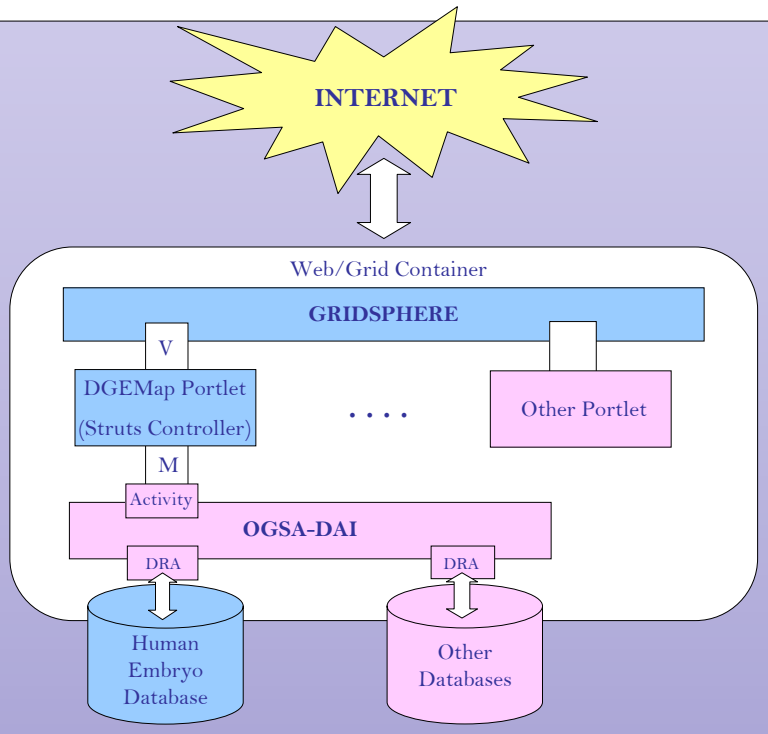


Figure 1 shows a schematic of the database which uses web-portal technology allowing secure access from any computer with a web browser. It is now in a test phase before further development. Parts already implemented are shown in blue while areas still in progress are shown in pink.

3D Mapping and Visualisation:

Development of newer and more accurate ways of mapping 'real' data onto 3D models of different stages of human development is an ongoing task but there is also a need for more effective visualisation of results. One such task is to facilitate comparisons in 3D, for example of two 3D models (e.g. Figure 2a - of human embryos with a normal and an abnormal karyotype) or gene expression patterns of two different genes at a single stage of development (e.g. Figure 2b) or of the same gene at two different stages of development.

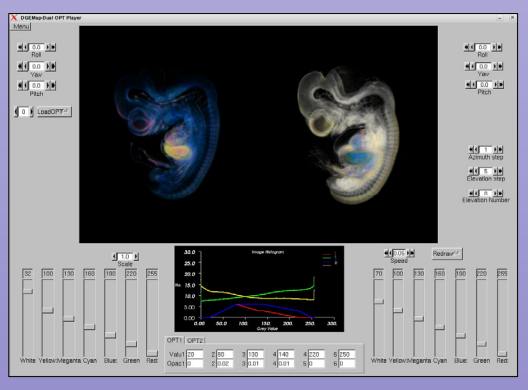


Figure 2a

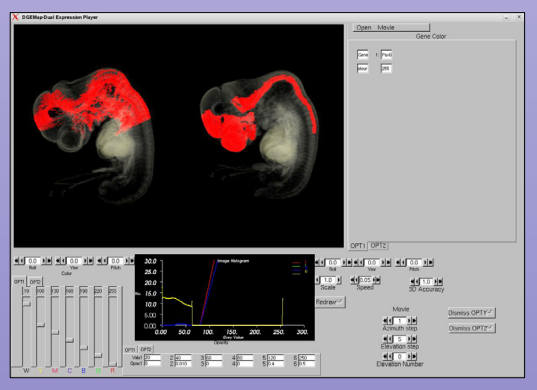
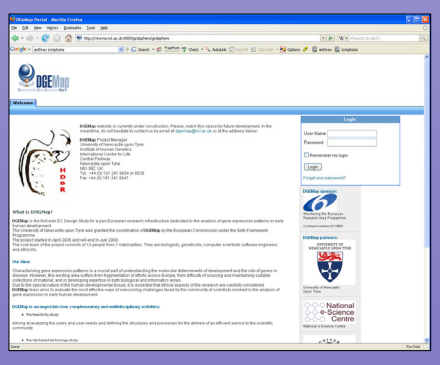


Figure 2b



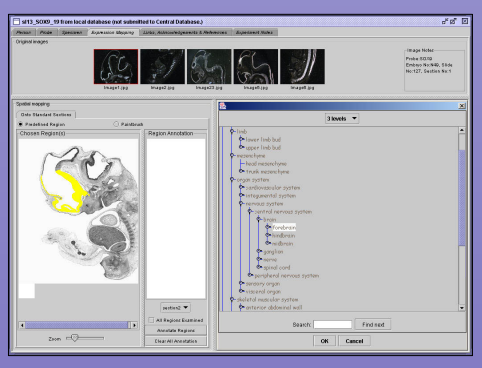


Figure 3

Establishing a Web Presence:

One of the provisions for the use of human material from the **MRC-Wellcome Trust HDBR** is that the results must be made publicly available upon publication, however this raises the problem of how to disseminate these results. The current implementation is a Java Web-start application built on Object Store software with a newly added web based search page. Figure 3 shows screen captures of the internal web portal and the Java based dissemination database.

Future Developments:

While the various stages of work are still in their infancy in terms of implementation, there is now a series of useable databases and visualisation software which can be built upon at the users request. It is up to the laboratory staff and collaborators to fully test these tools and report back on desirable changes. Newer technology will also be included in the development of these tools such as a move to DB2 for the dissemination database, inclusion of OGSA-DAI for Grid access and further developments of visualisation and mapping tools.

Acknowledgements

The project (DGEMap) is supported by the European Community - Research Infrastructure Action under the FP6 "Structuring the European Research Area" Programme (Contract number 011993).

Developmental material was obtained from the MRC-Wellcome Human Developmental Biology Resource (HDBR). The HDBR provides human embryonic and fetal material to the international scientific community and is held at the Institute of Human Genetics (Newcastle University) and the Institute of Child Health (University College, London) (*http://www.hdbr.org* and email hdbr@ncl.ac.uk or hdbr@ich.ucl.ac.uk).

www.dgemap.org

For further information, contact :
dgemap@ncl.ac.uk
DGEMap office +44(0)191 241 8646