Additional Information for CASIMIR meeting in Corfu

EUCOMM DATABASE

(i) BioMart

BioMart is a generic query optimised data management system developed at the EBI, originally as part of the data mining requirements of Ensembl (http://www.biomart.org). I have used this system to prepare prototypes for several mouse resources including a high level view of the **KOMP/EUCOMM** pipeline status of genes at the Sanger, **Europhenome** phenotypic data, **Eurexpress** in situ expression data and EMMA mouse mutant strain archive. The advantage of BioMart is that all these resources can be hosted at separate institution but a single web portal used to query them in an integrated manner. Integrated queries are possible where these data repositories share common identifiers such as Ensembl or MGI stable IDs. I have set up a test server at http://demo.biomart.org/biomart/martview to show these resources integrated with the existing **Ensembl mouse gene and variation** marts.

As part of an FP7 grant application I have also helped Vivek Iyer (Bill Skarnes group) to prepare a more extensive mart interface to the **GENETRAP** production pipeline of the ES cell lines including those produced for EUCOMM. This is also shown on the http://demo.biomart.org/biomart/martview interface.

I will be presenting this work on the *thursday afternoon* at the Casimir Corfu meeting as part of the **WP**5 section

(ii) DAS

BioMart also acts as an easily configurable DAS server. This allows integrated graphical presentation of mouse resource data on DAS clients such as the Ensembl web browser http://www.ensembl.org/Mus_musculus/contigview?c=1:75247177.5;w=200000. From contigView you can select **Manage Sources** in the **DAS Sources** menu and use the server at http://demo.biomart.org/biomart/das to select source such as **EurExpress, EUCOMM or EMMA** to display as a track in **Ensembl contigView**.
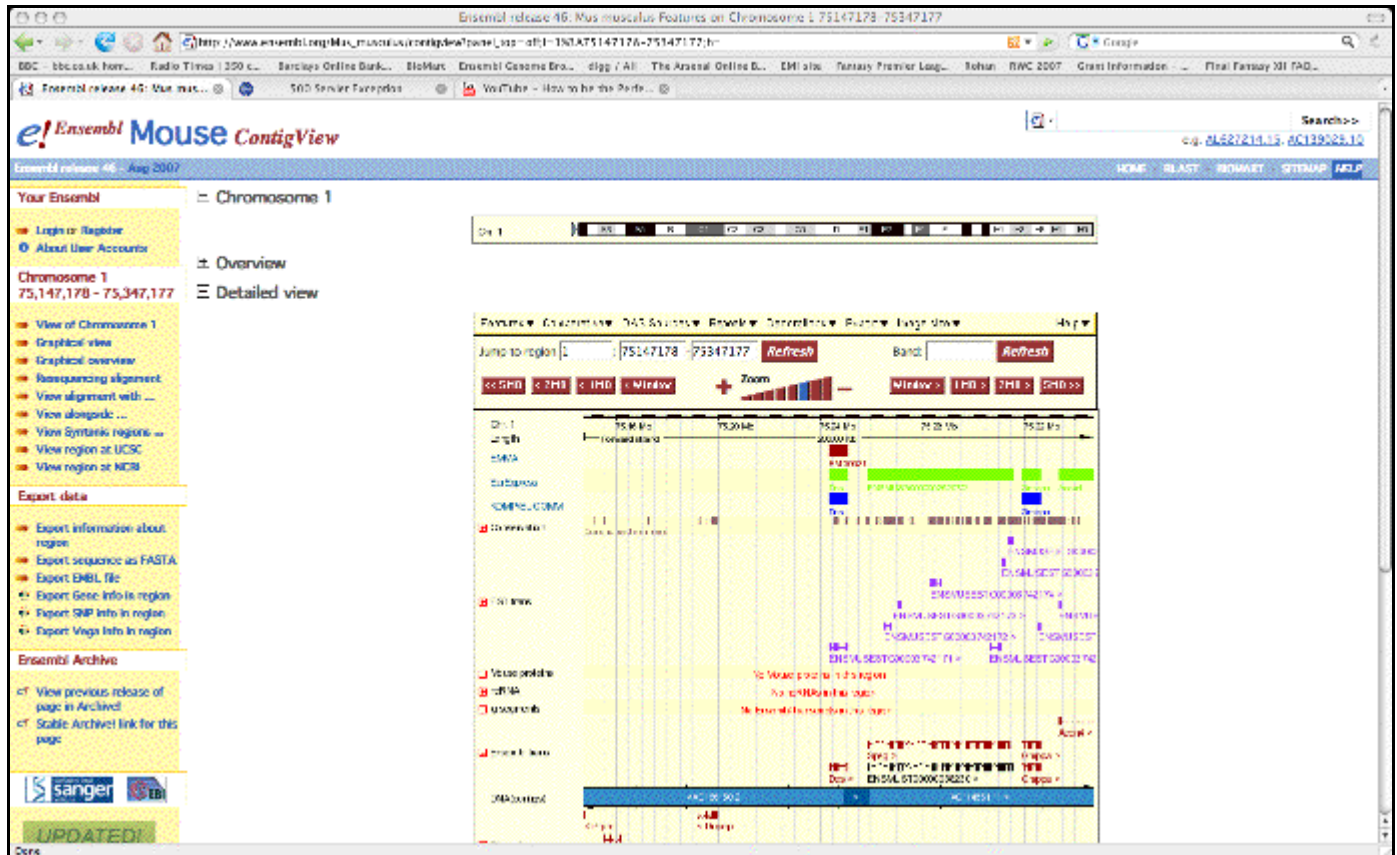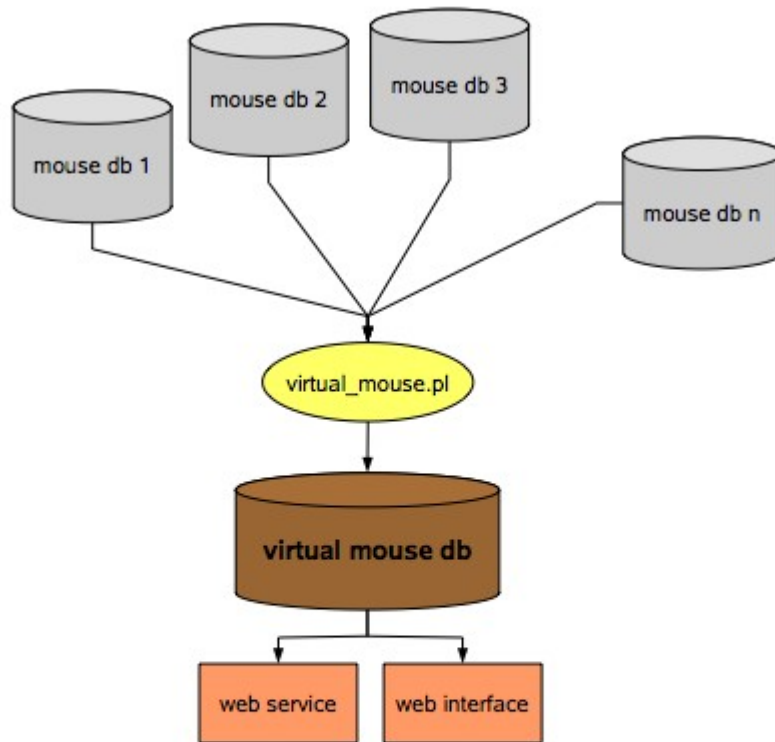

Damian Smedley (*Ensembl mouse informatics project leader*)

VIRTUAL MOUSE

(i) Virtual Mouse

The virtual mouse displays the mean value of all values of all mice ever determined on the world and whose values were added to databases accessible via webservices. The first picture shows a first draft of the virtual mouse web page.

Concept of the Virtual Mouse
Arrows mark the flow of data.
The script 'virtual_mouse.pl' has the following features:
- SOAP client: collecting mouse data from external databases via web service
- statistical analysis (calculation of mean value, standard deviation etc)
- feeding the database 'virtual mouse db' with results of statistical analysis.

The initial draft of the simple database structure of the virtual mouse is shown in the next picture.

Virtual Mouse DB

property (p_id, m_id, p_age, p_name, p_meanValue, p_standardDeviation, p_numberOfValues);
mouse (m_id, m_sex, m_strain, m_mutant);
update (u_date, p_id);
source (s_name, s_URL);
link_p_s (s_name, p_id);
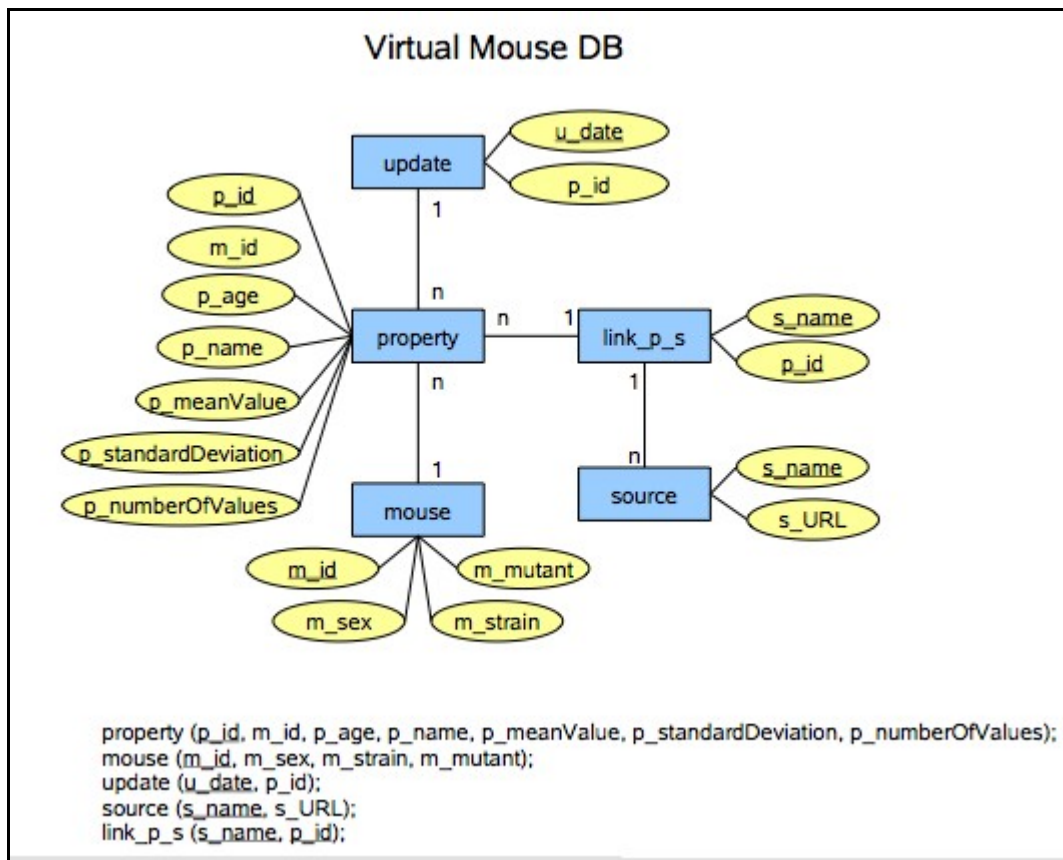
The virtual mouse will collect data for a particular parameter from all databases accessible and then calculates the mean value of it and displays this value. This value can then be used as a reference to a particular inbred mouse line or a particular mouse mutant. As the virtual mouse value of a given parameter is not so much depending of a single measurement, it is most likely a very robust value. It will allow for example the comparisons of values of a given inbred mouse strain measured at different laboratories in the world and could thereby reveal the robustness of a given parameter for mouse phenotyping.

The virtual mouse project is also a prototype that will tackle the basic problems of database interactions and tools like GO terms, Procedures and Values and so forth have to be transmitted in a standardized form in order to be able to combine the values from various phenotyping databases.

1 Ida Retter and 2 Werner Müller (*1 University of Braunschweig and 2 University of Manchester*)