

## **WP17: Informatics - the EMPReSS SOP and EuroPhenome databases**

To fully realise the potential of model organisms to bridge the gap between phenotype and genotype it is essential to provide structured descriptions of phenotypes that can be interpreted in a consistent fashion. We have recently suggested ways of using combinations of ontologies to describe mouse phenotypes<sup>1,2</sup> and provided tools that allow the storage, active updating and visualisation of multiple ontologies<sup>3</sup>. Central to this approach is the establishment of validated SOPs for the measurement of phenotypic attributes. We believe the screen or assay used to ascertain phenotype governs the phenotype that will be detected and therefore must play a central role in phenotype representation. We are in the process of constructing the controlled vocabularies of assays required for implementation of this schema. From this perspective the EMPReSS resource represents a collection of well-defined assays that can be linked directly to phenotype data.

To facilitate the storage and processing of the SOPs we have created SOPML, an XML language that allows the description of SOPs. The SOPs generated by different institutes are automatically annotated using a SOP template and stored in SOPdb, the underlying database. The documents produced by this process are automatically validated against the XML schema and manually curated by a domain expert before being committed in the database. The combination of XML-based markup, RDF document metadata and XSLT transforms facilitates filtering, advanced indexing, searching and rendering of the information<sup>4</sup>. The EMPReSS SOP database is accessible at <http://empress.har.mrc.ac.uk><sup>5</sup>.

We have also developed a phenome database for Eumorphia, EuroPhenome, which holds the baseline data from our standardisation and validation experiments on inbred strains. The EuroPhenome database links phenome data to the EMPReSS SOPs and is available through [www.europhenome.eu](http://www.europhenome.eu).

---

<sup>1</sup> Gkoutos et al (2004). Building mouse phenotype ontologies. *Pac. Symp. Biocomput.* 178-189

<sup>2</sup> Gkoutos GV, Green ECJ, Mallon AM, Hancock JM, Davidson D. (2005) Using ontologies to describe mouse phenotypes. *Genome Biol.* 6: R8

<sup>3</sup> Gkoutos et al (2005). CRAVE: a database, middleware and visualisation system for phenotype ontologies. *Bioinformatics.* 21, 1257-1262

<sup>4</sup> Gkoutos et al (2001). Chemical markup, XML, and the world-wide web. III. Toward a signed semantic chemical web of trust. *J. Chem. Inf. Comp. Sci.* 41: 1124-1130

<sup>5</sup> Green ECJ, Gkoutos GV, Lad HV, Blake A, Weekes J, Hancock JM (2005) EMPReSS: European mouse phenotyping resource for standardized screens. *Bioinformatics.* 21:2930-2931