

**Christophe Lefèvre**

**Comparative genome bioinformatics and bioinformatics of the lactation program: Mammosapiens**

Recently the genome sequences of a rapidly growing number of organisms have become available. The comparisons of these genomes yield tremendous insights into the genes that are essential for life and those that define the species, revealing the mechanisms of evolution and the hidden mechanisms of gene regulation.

Mammals are characterized by the total dependency of the new born on milk produced by the maternal mammary gland and lactation is one of the most remarkable products of evolution. The rich mammalian diversity found in Australia provides a unique resource to study the evolution of lactation. Mammalian species have evolved a variety of lactation strategies. For example, marsupials give birth to a relatively immature embryo after a short gestation period. By contrast, in Eutherians most of the development occurs in Utero. Thus, the marsupial young depends on milk for a significant period of time of its development. As a result, the study of lactation in marsupial is not only interesting for the exploration of the evolution of the lactation system in mammals, but also provides a unique model to explore the role of milk factors on the control of mammalian development.

The Mammosapiens project is employing high throughput technology platform, including genomics, transcriptomics, proteomics, metabolomics and bioactivity screens, for the study of lactation in mammals with extreme lactation strategies. A bioinformatics resource is used to support storage and analysis of the data generated.

**1) e-Research environment for Lactation Bioinformatics**

The project is to develop an online internet based platform for the bioinformatics analysis of sequence and gene expression analysis of data related to the study of the lactation system in a number of mammalian species generated by high throughput genomics technologies (sequencing, microarray, proteomics, metabolomics).

Development will be done in PHP using MySQL, and scripting for integration with bioinformatics software.

Reference: MammoSapiens: eResearch of the lactation program. Building online facilities for collaborative molecular and evolutionary analysis of lactation and other biological systems from gene sequences and gene expression data. Kevin R Nicholas, Amit Kumar, Yvan Strahm, David Powell, Torsten Seemann, Kerry A. Daly, Amelia Brennan, Karensa Menzies, Julie Sharp, Matthew Digby and Christophe Lefèvre (2008). In: eResearch Australasia 2008, Melbourne,

Australia, 28 September - 3 October 2008.

<http://espace.library.uq.edu.au/view/UQ:155351>

**2) Analysis of gene expression data microarrays related to lactation in a number of mammalian species ( platypus, echidna, wallaby, seal, cow, human...).**

A large library of gene expression data is being collected for a number of mammalian species. The project is to pursue global analysis, including comparative analysis of different species combining gene expression and genome sequence, and develop online data mining interface of this data (in association with other projects).

References:

- "[Fur Seal Adaptations to Lactation – Insights into Mammary Gland Function](#)". Julie A. Sharp, Kylie N. Cane, Christophe Lefevre, John P.Y. Arnould and Kevin R. Nicholas. *Curr Top Dev Biol*. Chapter 6, 2006;72:275-308. Review.
- "[Isolation and characterization of two whey protein genes in the Australian dasyurid marsupial the stripe-faced dunnart \(\*Sminthopsis macroura\*\)](#)". A. A. De Leo, C. Lefevre, D. Topcic, E. Pharo, J.F. Cheng, P. Frappell, M. Westerman, J. A. M. Graves, and K. R. Nicholas. *Cytogenet Genome Res*. 2006;115(1):62-9.
- "[The tammar wallaby and fur seal; models to examine Local control of lactation](#)". Amelia J. Brennan, Julie A. Sharp, Christophe Lefevre, Denijal Topcic, Aurelie Auguste and Kevin R. Nicholas. *J Dairy Sci*. 2007 Jun;90 Suppl 1:E66-75. Review. PMID: 17517753.
- "[The Fur Seal – a Model Lactation Phenotype to Explore Molecular Factors Involved in the Initiation of Apoptosis at Involution](#)". Julie A. Sharp, Christophe Lefevre Amelia J. Brennan and Kevin R. Nicholas. *J mam gland biol & neo Review*, 2007, 12(1):47-58.
- "[Molecular evolution of Monotreme and marsupial whey acidic protein genes](#)". Julie Anne Sharp, Christophe Lefevre, Kevin R Nicholas. *Evol Dev*. 2007 Jul-Aug;9(4):378-92.
- "[Analysis of the expression of immunoglobulins throughout lactation suggests two periods of immune transfer in the tammar wallaby \(\*Macropus eugenii\*\)](#)". Daly KA, Digby M, Lefevre C, Mailer S, Thomson P, Nicholas K, Williamson P. *Vet Immunol Immunopathol*. 2007 Jul 20. [Epub ahead of print] PMID: 17727962.
- "Significance, origin and function of bovine milk proteins: the biological implications of manipulation or modification. In; *Milk Proteins: From Expression to Food*". Sheehy P. A., Williamson, P., Sharp, J., Menzies, K., Lefevre, C., Digby, M., Nicholas, K.R. and Wynn, P.C. Book chapter (Editors: Abby Thompson, Mike Boland, Harjinder Singh) Elsevier. (in press)
- "[Characterization and expression of Peroxiredoxin 1 in the neonatal tammar wallaby \(\*Macropus eugenii\*\)](#)". Daly KA, Lefevre C, Nicholas K, Deane E, Williamson P.

Comp Biochem Physiol B Biochem Mol Biol. 2007 Sep 4; [Epub ahead of print] PMID: 17913528.

["Lactation transcriptomics in the Australian marsupial, \*Macropus eugenii\*: transcript sequencing and quantification"](#). Christophe M. Lefèvre, Matthew R. Digby, Jane C. Whitley, Yvan Strahm, Kevin R. Nicholas. BMC Genomic, 2007 Nov 13;8:417. PMID: 17997866.

["Identification and transcript analysis of a novel wallaby \(\*Macropus eugenii\*\) basal-like breast cancer cell line"](#). Sharp JA, Mailer SL, Thomson PC, Lefevre C, Nicholas KR.

Mol Cancer. 2008 Jan 7;7(1):1.

["A population of mammary epithelial cells do not require hormones or growth factors to survive"](#). Brennan AJ, Sharp JA, Khalil E, Digby MR, Mailer SL, Lefèvre CM, Nicholas KR. J Endocrinol. 2008 Mar; 196(3):483-96.

["Identification, characterization and expression of cathelicidin in the pouch young of tammar wallaby \(\*Macropus eugenii\*\)"](#). Daly KA, Digby MR, Lefèvre C, Nicholas KR, Deane EM, Williamson P. Comp Biochem Physiol B Biochem Mol Biol. 2008 Mar;149(3):524-33.

["CD14 and TLR4 are expressed early in tammar \(\*Macropus eugenii\*\) neonate development"](#). Daly KA, Lefèvre C, Nicholas K, Deane E, Williamson P. J Exp Biol. 2008 Apr;211(Pt 8):1344-51.

["Genome analysis of the platypus reveals unique signatures of evolution"](#).

Wesley C. Warren et al. Nature. 2008 May 8;453(7192):175-83.

["Uncoupling the mechanisms that facilitate cell survival in hormone-derived bovine mammary explants"](#). Brennan A, Sharp J, Lefevre C, Nicholas K. J Mol Endocrinol. In Press.

### **3) Combinatorial algorithm for comparative genome sequence analysis**

The project is to develop combinatorial analysis algorithm and sequence analysis applications for the identification of sequence conservation in many species and its visualisation in the context of sequence annotation.

This will be based on previous work describing combinatorial algorithm for sequence analysis. Applications will be implemented and used to analyse mammalian genomes evolution and function, in particular genomic regions encoding milk protein genes.

#### References:

" [The position end-set tree: A small automaton for word recognition in biological sequences.](#)" Christophe Lefèvre and Joh-E Ikeda. CABIOS (1993), 9, 3, 343-348.

" [Pattern recognition in DNA sequences and its application to consensus footprinting.](#)" Christophe Lefèvre and Joh-E Ikeda. CABIOS (1993), 9, 3, 349-354.

"[A small automaton for word recognition in DNA sequences and its application to consensus analysis of regulatory elements in DNA regions controlling gene expression.](#)" Christophe Lefèvre and Joh-E Ikeda. Proceedings of the First International Conference on Intelligent Systems for Molecular Biology (eds. L.

Hunter, D. Searls, and J. Shavlik) AAI/MIT Press, Menlo Park CA, 1993, p 243-250.

["A fast word search algorithm for the representation of sequence similarity in genomic DNA"](#). Christophe Lefèvre and Joh-E Ikeda. Nucleic Acid Research, (1994) Vol 3,404-411.

#### **4) Network Biology of the lactation system**

The accumulation of information about biological networks in the last years has enable an approach where biological network can be integrated and mined to infer functional relationships.

Such approaches will be developed and used to analyse functional relationships in the lactation system.

Networks may be obtained from available data (protein interaction, metabolic pathways) or constructed from literature or other analysis.

#### **5) The visible mammary epithelial cell**

This project will construct a virtual model of the mammary epithelial cell and attempt to simulate its regulation during lactation.

## **6) PhD projects on offer at the Australian Animal Health Laboratory, CSIRO Livestock Industries, Geelong**

**Project number: G9 Mining the chicken genome. See <http://www.csiro.gov.au/files/files/p9uu.pdf>**

This project aims to extract valuable information from the chicken genome. The goal will be to develop databases and search and analysis tools to address new research areas. For example automated gene feature extraction tools will be developed to allow the identification and cataloguing of 5' and 3' untranslated regions as well as upstream promoter regions. These databases will be used in the meta-analysis of various messenger RNA (mRNA) and microRNA (miRNA) expression data sets to identify possible miRNA/mRNA regulatory interactions and co-ordinately expressed gene sets. Such analysis may allow the refinement of miRNA target prediction models. The overall goal is to contribute to our understanding of the regulatory mechanisms involved in controlling gene expression in the chicken.

*For more information contact:* Dr Robert Moore (Rob.Moore@csiro.au ), phone + 61 3 5227 5760 or Dr Christophe Lefevre (Christophe.Lefevre@deakin.edu.au), phone + 61 3 5227 2037.