uc_Milk: An ontology for scientifically-based unambiguous characterization of mammalian milks, their composition and the biological processes giving rise to their creation

Emeline Colet
Ecole Nationale Supérieure de Chimie de Paris
Chimie ParisTech
Paris, FRANCE
emeline.colet@etu.chimie-paristech.fr

Matthew Lange, PhD
Department of Food Science and Technology
University of California Davis
Davis, CA
mclange@ucdavis.edu

***Abstract* - Recent efforts in biological ontology go to great lengths to unambiguously categorize biological entities and phenomena of the natural world, as well as their relationships with each other. This paper illustrates the importance of unambiguously characterizing mammalian milk because milk is a complex mixture of many chemical components and thus represents a key role in infant nourishment and development. In addition to the build of a computable knowledge base around mammalian milk, ontological modeling of this aspect of biology and chemistry enable increased understanding of mammalian milk composition and the biological structures and biochemical processes giving rise to their creation. Utilizing unambiguous vocabularies to compare human milk with other mammalian milks relative to the biological and behavioral survival challenges facing varied mammalian organisms and the phenotypic qualities each milk confers, is a fundamental goal of this project.**

Keywords* - *mammalian milk composition; ontological modeling; biological processes

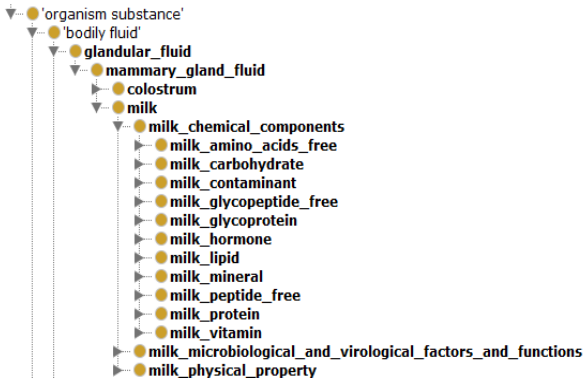
One foundational design pattern for creating uc_Milk is an ontology for unambiguous characterization of mammalian milks, their composition and the biological processes giving rise to their creation. Available online at the GitHub website *github.com/IC-FOODS/uc_Milk*, uc_Milk is part of a larger multi-ontology framework [1] currently being housed within the International Center for Food Ontology Operability, Data, and Semantics. IC-FOODS at UC Davis [2]. Milk is at once, a consumable beverage, an ingredient at the heart of the dairy industry, and often the sole nutritive source for infants--being both a source of dietary metabolites as well as a transport mechanism for xenobiotic agents. In all of its roles as a food/ingredient source, milks are known to confer specific health phenotypes as they interact with specific genotypes. Yet to date milk has received very limited attention in the world of ontological research.

We used the Basic Formal Ontology as an upper ontology because of its ubiquity in biological and biomedical ontologies, for facilitating integration by way of meta-concepts. In addition to characterizing the biological and biochemical structures and processes related to mammalian milks, since milk is a commodity as well as an essential biological fluid, uc_Milk ontology considers social, commercial, and environmental entities related to milk consumption and production (both as a food and as an ingredient in foods). These additional classifications hold promise for uc_Milk ontology usability relative to traceability and trust in commercial milk and all milk products, and provide a platform for improvement of processing and production techniques, information about milk distribution, availability and choice, as well as increased health and well-being of consumers. uc_Milk maps directly to entities from several other ontologies. Yet because milk is often at the “leaf” end of existing ontologies as in the "Environment Ontology". The "Uber Anatomy Ontology (Uberon)” and the "Gene Ontology" and "Vertebrate Taxonomy Ontology" respectively provide structures for mammary gland structures and development, the lactogenesis biological process, and formal mammalian phylogeny. uc_Milk also maps to the "Cell Ontology (CL)" for the intracellular anatomical structures and processes. uc_Milk is currently expanding feeding behaviors by mapping to the burgeoning uc_Eating ontology as well as to anatomical entities and the roles in the development of neural infant tissues in the NeuroBehavior Ontology. uc_Milk increases availability of structured knowledge relative to milk.

Lactation is the hallmark underlying biological force driving the mother-infant dyad in Mammalia. The significant energetic and metabolic costs of lactation imposed to the mother suggests that milk’s role in infant health extend well beyond simple nutrition. Millennia of selective pressures on the mother-infant dyad, have conferred in milk a dual purpose: an optimal source of nutrients, and a delivery vehicle for bioactive agents. Milk is known as nature’s most complete food because of its complex mixture of bioactive components and essential nutrients such as protein, fat, carbohydrate, minerals, vitamins, and physiologically active substances. Aside from nutritional values of milk, biologically active compounds such as casein and whey proteins have been found to be increasingly important for physiological and biochemical functions that have crucial impacts on human metabolism and health. Accordingly, the infant gut and its associated microbiota are adapted not only for the utilization and absorption of milk macronutrients, but also to respond to non-nutritive, yet bioactive molecular entities. Important non-nutritive, bioactive agents in human milk include an astonishing high number of complex oligosaccharides, and a plethora of glycoconjugated proteins and lipids. Complex prebiotic oligosaccharides are known to be present in domestic animal milks, yet are in low abundance relative to human milk, a

known source of prebiotic oligosaccharides with important effects on human health. Characterization of these molecules, known as Human Milk Oligosaccharides or HMOs started in the 1960s, yet even today, their analysis remains a challenging task due to the large number of structures and their structural complexity. No commercially viable processes for manufacturing these complex and important milk bioactives exist, making the third largest component of breast milk, conspicuously absent in today's infant formulas, weaning foods, medical foods and everyday dairy products. Understanding the extent and types of oligosaccharides present in bovine milk is an important step towards determining the feasibility of developing commercial sources. The identification, annotation and characterization of oligosaccharides and other bioactive compounds in milk in is a necessary step to evaluate the processes giving rise to their formation, and will enable future scale-up processes for recovering these ingredients for commercial applications.

The tremendous variation of milk compositions that occurs among mammalians is one of the particularly interesting aspects of lactation biology. As mammals include the largest animals of the planet, a classification has been established: mammalians have been separated into two major groups, placental and non-placental mammals. Classification systems based on molecular studies reveal three major groups or lineages of placental mammals: Afrotheria, Xenarthra and Boreoeutheria which have been divided in clades, superorders and orders [3] [4]. The aim is to compare composition of a great number mammalian milk with human milk based on fat, protein, sugar, vitamin, dry matter and energy content and to understand the influence of various factors such as consumed diet, lactation time, habitat and many others on milk composition [5]. Another important part of developing this ontology is to describe the processes giving rise to milk creation. Indeed milk is a complex mixture whose composition reflects the activities of distinct secretion and transport processes of the mammary gland and mirrors the differing nutritional requirements of mammalian mammary gland neonates [6]. To accomplish those main goals, uc_Milk covers two main points of view: on the one hand because milk is a complex biologically derived fluid, uc_Milk will serve basic science but on the other hand, since milk is involved in many dairy products, this ontology is also designed to facilitate



utilization in commercial and industrial foods.

Figure 1 : Milk components

The first part of uc_Milk has breastmilk as starting point: as mammary glandular fluid, milk supplies water and chemicals components to the neonate for his survival, proper development, and vigorous growth (Figure 1). uc_Milk provides structures in three stages of life cycle: infant stage with breastfeeding thanks to milk, pregnancy stage with biological modifications and hormones, and lactation stage which is divided in four steps : mammogenesis, lactogenesis, galactopoiesis and involution. At the same time, lactogenesis is classified as subclass of biological process due to the secretory differentiation of mammary gland and the biological production of milk. Additionally, lactation system described as subclass of exocrine system includes the biological functional anatomy of the lactating mammary gland and the biological transport and secretion pathways because solutes can enter milk through both transcellular and paracellular routes. Transcellular routes are divided in four general pathways, two for the secretion of endogenous substances and two other for the transport of exogenous substances (Figure 2) [6]. The biological factors of milk production overlap with the

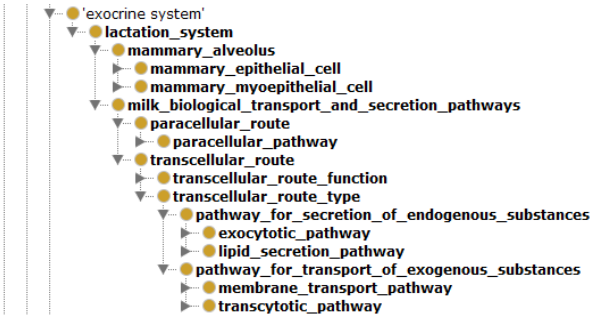


Figure 2 : Lactation system

breastfeeding behavior of uc_Eating ontology.

The second part of uc_Milk is industrial and food focused. uc_Milk provides structures in milk processing, milk process, milk producing, milk production as well as food products : milk is involved in many dairy food products such as butter, cream, or cheese. Milk attributes are subclasses of food product and the milk sensory quality attribute overlap with the uc_Sense ontology. The format of the ontology connects the subclasses of milk processing process and milk production process to the uc_Processing ontology. Furthermore, uc_Milk characterizes the composition of different mammalian milk as subclass of milk producing entity.

Using Protégé, with the streamlined interface and the program's ability to add changes and import existing ontology files improved the flow of classification and improved the reproducibility of uc_Milk. This ontology is still at an early stage so we are just beginning the process of adding axioms and we have not yet tested this initial ontology. Sustainability plans for the ontology will be developed once we receive initial feedback from the community about how paths forward for integration with related ontologies. Entity identifiers will be established after feedback is received from ICBO participants, and integration partners are identified.

REFERENCES

[1] Lange, M. C., Lemay, D. G., German, J. B. (**2007**). A multi-ontology framework to guide agriculture and food towards diet and health. *Journal of the Science of Food and Agriculture*, **87**, 1427-1434.

[2] Boulos, M. N. K., Yassine, A., Shirmohammadi, S., Namahoot, C. S, Brücker, M. (**2015**). Towards an "Internet of Food": Food Ontologies for the Internet of Things. *Future Internet*, **7**, 372-392.

[3] Tarver, J.E. et al. (**2016**). The Interrelationships of Placental Mammals and the Limits of Phylogenetic Inference. *Genome Biol Evol*, **8** (2) : 330-344.

[4] Langer, P. (**2008**). The phases of maternal investment in eutherian ammals. *Zoology*, **111** (2), 148-162.

[5] Skibiel, A.L., Downing, L.M., Orr, T.J, Hood, W.R. (**2013**). The evolution of the nutrient composition of mammalian milks. *Journal of Animal Ecology*, **82**, 1254–1264.

[6] Mc Manaman, J.L., Neuville, M.C. (**2003**). Mammary physiology and milk secretion. *Advanced Drug Delivery Reviews*, **55**, 629–641.