

Enhancement of the Ontology of Craniofacial Development and Malformation (OCDM): Human and Mouse Craniofacial Musculoskeletal Development, Craniofacial Zebrafish Ontology (CZO) and Hub Metadata Annotation List

JLV Mejino, LT Detwiler, TC Cox, JF Brinkley

University of Washington and Seattle Children's Research Institute, Seattle, WA, USA

Data integration is one of the primary objectives of FaceBase, and such a task requires a consistent and standardized terminology for data annotation. A primary purpose of the OCDM is to provide a controlled set of terms or keywords that can be employed for data annotation, so that diverse data can be retrieved via common terms. Equally importantly, the terms are ontologically related to each other, thereby promoting integration at the level of meaning rather than simply via keywords (semantic integration).

We initially built the OCDM based on our Foundational Model of Anatomy ontology (FMA) to include components representing adult and developmental anatomy and malformations in the human and the mouse species, as well as homology mappings between these two organisms. In FaceBase 2 the focus is on cranial vault development and associated dysmorphology such as craniosynostosis, midface hypoplasia and micrognathia. Hence we augmented the spatio-structural representation of the development of the musculoskeletal system of the head (cMSK), with emphasis on the developmental properties of embryonic and fetal structures addressed in the different use cases in FaceBase.

An important addition in FaceBase 2 is the incorporation of zebrafish craniofacial research. We therefore created the Craniofacial Zebrafish Ontology (CZO) as another species-specific anatomy ontology in the OCDM. We applied the same ontological framework to the CZO as we used for the human and mouse components, and populated it with content imported from the Zebrafish Anatomy Ontology (ZAO) from ZFIN, as well as with contributions from zebrafish domain experts in the FaceBase group. We assigned the terms/concepts to their proper places in the taxonomy based on their semantic definitions, while we developed and extended, where needed, the spatio-structural relationships related to the zebrafish cMSK. We incorporated the different zebrafish developmental stages from ZFIN and published sources, as well as the preferred stage descriptions used by the zebrafish FaceBase researchers.

The HUB collects data from the different spokes of FaceBase, and a necessary step towards integration of these data is to develop metadata standards (data elements and value sets) to be used for annotating the data. These standards are necessary for keyword-based data searches and are a prerequisite for semantic searches using relations in the OCDM and other ontologies. We have provided the HUB an initial set of metadata spreadsheets based on project descriptions provided by members of the consortium. We generated 14 metadata lists that cover terminologies for the following domains: species, specimen types, human, mouse and zebrafish anatomy, human and mammalian malformation

phenotypes, human, mouse and zebrafish developmental stages/ages, mouse strains and imaging modalities/devices. Many of the values, especially those relating to human and mouse anatomy and human malformation phenotypes, were directly derived from the OCDM. Others were pulled from existing ontology sources, thereby facilitating the integration of FaceBase data in the emerging global database.