OncoCL: A Cancer Cell Ontology

Karen E. Rasmussen and Mary E. Dolan*
The Jackson Laboratory, Bar Harbor, ME 04609

ABSTRACT

Pathologists and molecular biologists don't speak the same language – even when they are talking about the same thing. Basic surgical pathology for tumor biopsies and resections has evolved separately from basic cancer biology, such that surgical pathology and molecular biology are only just starting to merge. Our goal in this work is that OncoCL provide a shared language.

We have developed an ontology, OncoCL, to describe cancer cells and provide a framework for consistent annotation of cancer-associated genomic and phenotypic data. OncoCL captures the spatiotemporal, histopathologic, and molecular properties of tumorigenesis at recognized stages in cancer initiation and at critical stages of cancer progression.

The conceptual basis of our approach is that cancer cell phenotypes derive from the succession of alterations based on acquisition of the cancer hallmarks described by Hanahan and Weinberg. We have implemented a default logic representation based on linking to the existing Cell Ontology (CL), a representation of normal cell types, as the 'cell of origin' -- that is, a normal/canonical cell type that maintains all its normal characteristics until it undergoes oncogenic change.

1 ONCOCL DESIGN PRINCIPLES

Hanahan and Weinberg proposed an organizing principle for cancer (2000), six cancer hallmarks. More recently, they included two emerging hallmarks: reprogramming of energy metabolism and evading immune destruction; and two enabling characteristics for the acquisition hallmark capabilities: genome instability and inflammation (2011). Our model represents oncogenic change in a cell as the acquisition of these hallmarks.

As we developed OncoCL, we found that, along with purely 'cellular' characteristics, the tissue and tumor context of the cell was an important part of characterizing the cell as part of cancer progression.

The Open Biomedical Ontology community (Smith), after years of ontology development, has formulated a set of principles for good ontology development, among them: embedding in the Basic Formal Ontology (BFO), using only relations from the Relation Ontology (RO), and reusing reference ontologies as much as possible. OncoCL reuses CL, UBERON (Anatomy Ontology), BTO (BRENDA Tissue Ontology), Pathway Ontology (Rat Genome Database pathway and Reactome). In addition, we reuse selected terms from several ontologies: PATO (Quality Ontology: 'cellular_phenotype' and 'morphology'), CPO (Cellular Phenotype Ontology: 'cellular_phenotype' and 'morphology'); SO (Sequence Ontology: 'variant' and 'alteration').

We have introduced four additional classes: CanonicalCell, CancerCell, CancerHallmark, CancerRelatedGene.

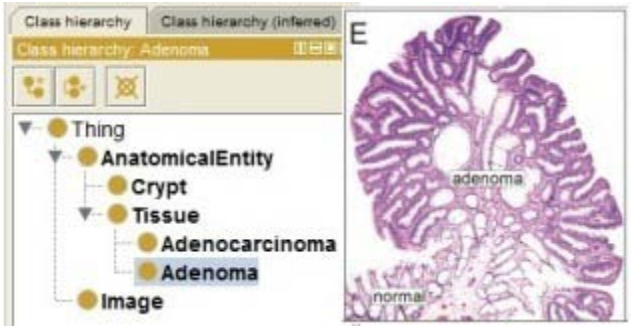
* To whom correspondence should be addressed: Mary.Dolan@jax.org

2 IMPLEMENTATION OF ONCOCL

As our first use case, we have modeled human colorectal cancer (CRC) where the progression of cell transformations has been well studied. Our initial model is a comprehensive description of CRC progression, from normal epithelium through metastasis with the goal of integrating data from molecular biology and histopathology as identifying characteristics.

Vogelstein (1990) proposed a model of colorectal cancer development as a result of genetic mutations. The 'Vogelgram' correlates molecular biology of progression with morphology as a polyp transitions to cancer.

As a tool for developing OncoCL we are using the freely-available, open-source Protégé Ontology Editor. Protégé is based on OWL, a language with very expressive formal semantics. In addition, we incorporate a user-developed Protégé plug-in (Balhoff) that enables integration of images of cells and tissues with appropriate classes to facilitate use by pathologists.



ACKNOWLEDGEMENTS

This work is supported by NIH with funding through NCI CA155825.

REFERENCES

Balhoff J. (2011) https://github.com/balhoff/image-depictions-view

Bard J, Rhee SY, Ashburner M. (2005) An ontology for cell types. *Genome Biology*, 6, R21.

Fearon ER, Vogelstein B. (1990) A genetic model for colorectal tumorigenesis. *Cell*, 61(5):759-67.

Hanahan D, Weinberg RA. (2000) The hallmarks of cancer. *Cell*, 100, 57-70; (2011) The hallmarks of cancer: The Next Generation. *Cell*, 144, 646-674.

Smith B et al. (2007) The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol*. 25(11):1251-5.