



Hematopoietic Cell Types: Prototype for a Revised Cell Ontology



Alexander D. Diehl
The Jackson Laboratory



July 24, 2009



The Cell Ontology

- An ontology of cell types built by biologists for the needs of data annotation and analysis.
- The main branch organizes and enumerates in vivo cell types.
- The Cell Ontology covers cell types from all of biology.
- The Cell Ontology is not
 - A list of specific cell lines, immortal or otherwise, although the Cell Ontology may be used to describe such cells if they correspond to an in vivo cell type.
 - A list of in vitro methods for preparing cell cultures, although the Cell Ontology may be used to describe the resulting cells if they correspond to an in vivo cell type.

The Cell Ontology

- Initially developed in 2003 by Jonathan Bard, David States, Michael Ashburner, and Seung Rhee.
- First described in 2005, Genome Biology, 6:R21.
- Current version relies solely on is_a and develops_from relationships, with extensive multiple inheritance.
- Available as the file cell.obo at <http://obofoundry.org> and <http://bioportal.bioontology.org/>.
- Lindsay Cowell is the Cell Ontology liaison to the OBO Foundry to review the ontology for conformity to the OBO Foundry principles.

Use Cases for the Cell Ontology

- Data Annotation (ongoing for GO, future use for IEDB, IDO, and VO).
- Cross product term formation with GO, MP, and other ontologies (ongoing).
- Representation of flow cytometry results.
- Immune system modeling.
- Nervous system modeling.
- Others?

NIAID Cell Ontology Workshop

May 2008

- Gathering of NIAID experts on immune system cells with an interest in ontology.
- Two day meeting focusing on:
 - Ontology best practices
 - T cells
 - B cells
 - Dendritic cells
 - Macrophages

NIAID Cell Ontology Workshop

May 2008

- Many additional cell types needed to represent immune cell types more fully.
- Need to restructure ontology to remove multiple inheritance.
- Desire to rebuild ontology via logical definitions for terms based on cross-products.

Additional cell types added to represent immune cell types

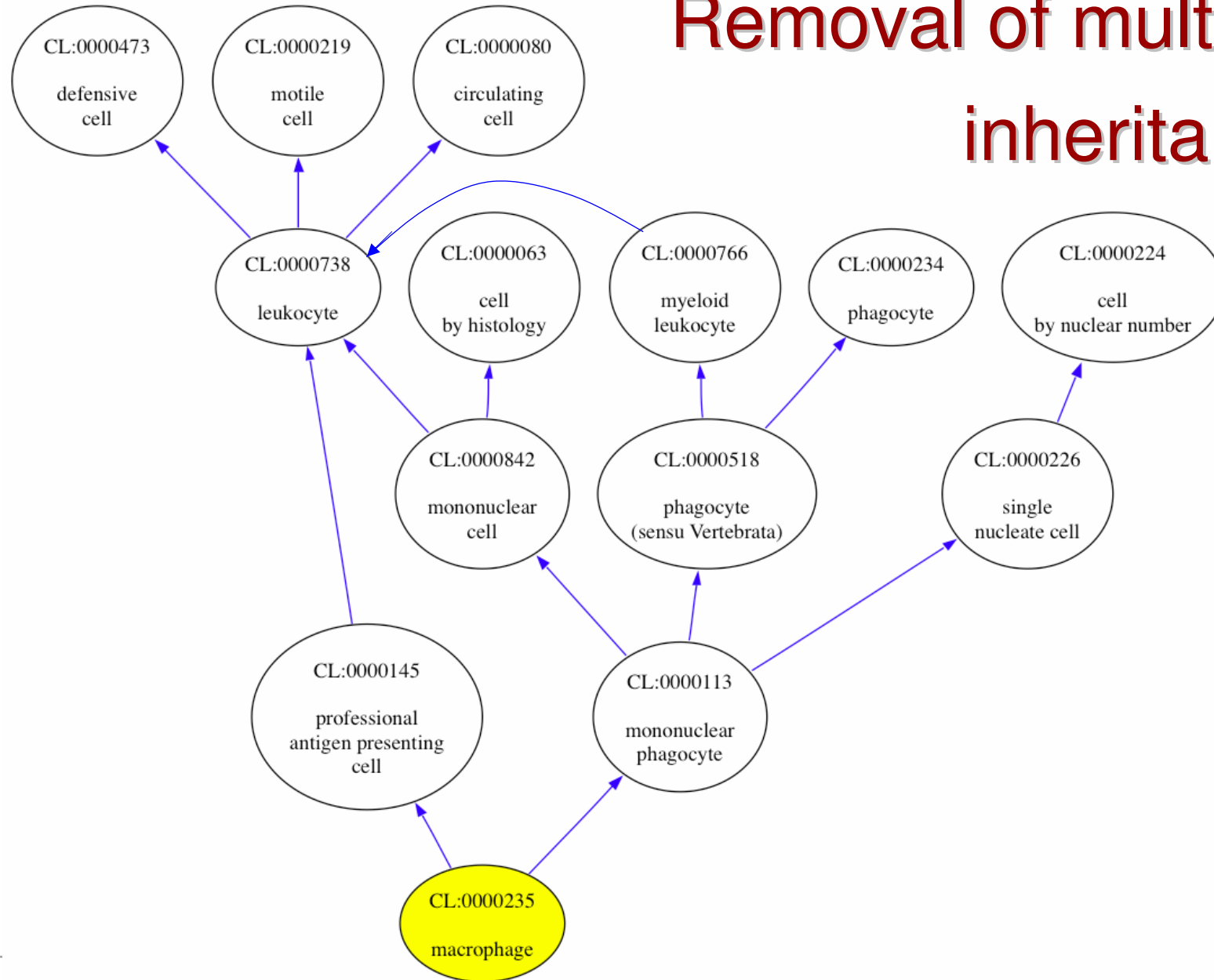
Missing cell types:

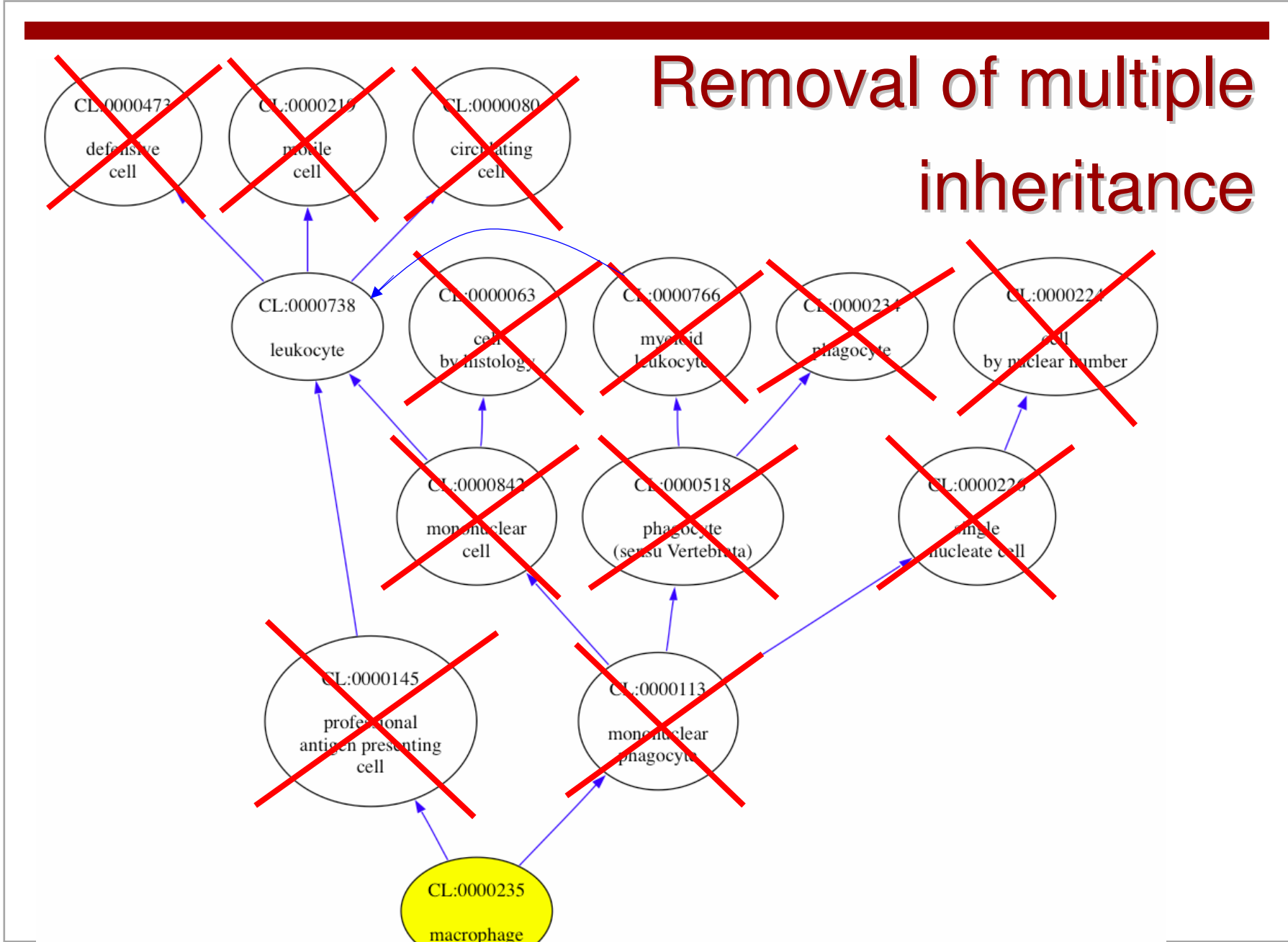
- Th17 cell
- Tr1 cell
- Plasmablast
- inflammatory macrophage
- etc.

Precisely specified cell types :

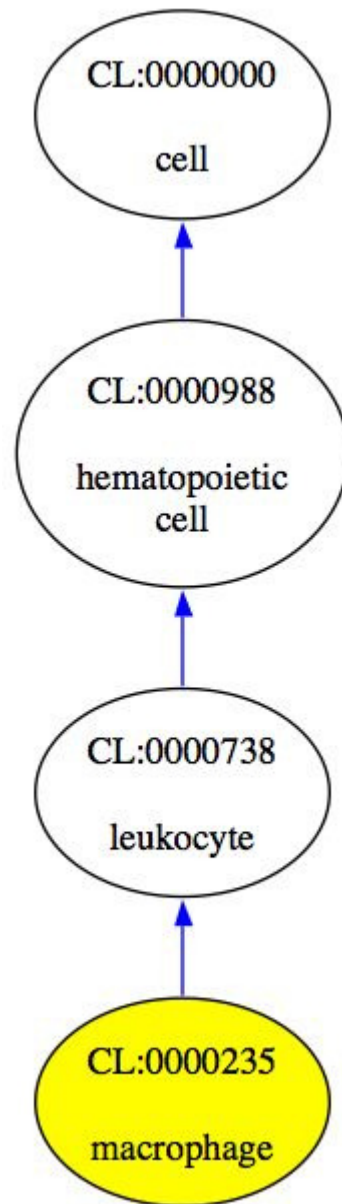
- CD4-negative, CD8-negative type I NK T cell secreting IL-4
- mature CD8alpha-positive, CD11b-negative dendritic cell
- IgA Memory B Cell
- etc.

Removal of multiple inheritance





Removal of multiple inheritance



How do we describe cells?

- Morphology
- Surface markers
- Transcription factors
- Location
- Role or process involvement
- Lineage

How do we describe cells?

Link to external ontology

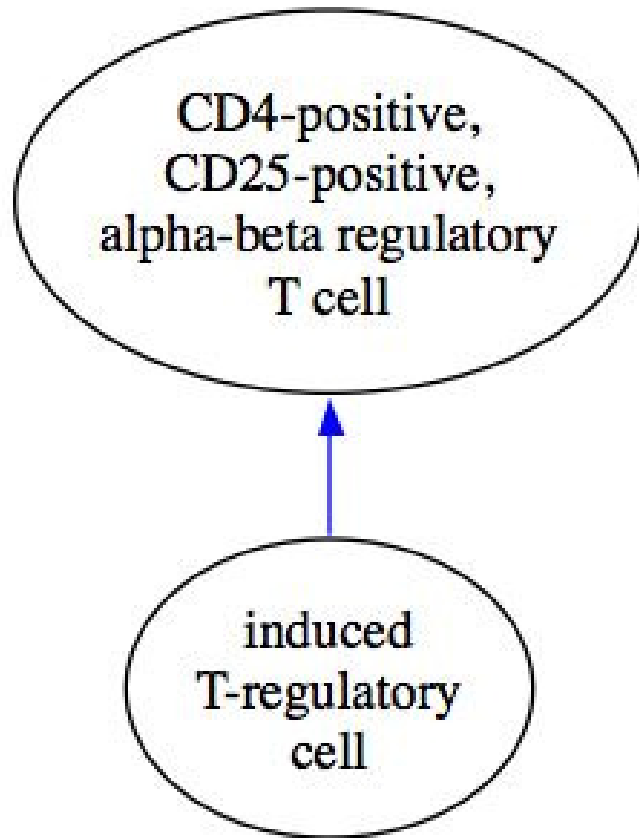
- Morphology
 - Surface markers
 - Transcription factors
 - Location
 - Role or process involvement
 - Lineage
- Protein Ontology
- Anatomy Ontology (FMA)
- GO Biological Process
- Cell Ontology
-
- The diagram illustrates the mapping of cell description categories to external ontologies. Arrows indicate the following connections: 'Surface markers' and 'Transcription factors' both point to 'Protein Ontology'; 'Location' points to 'Anatomy Ontology (FMA)'; 'Role or process involvement' points to 'GO Biological Process'; and 'Lineage' points to 'Cell Ontology'.

How do we describe cells?

Link via specific relationships

- Morphology
 - Surface markers
 - Transcription factors
 - Location → □□located_in
 - Role or process involvement → □□participates_in
 - Lineage → □□develops_from/has_lineage
- has_part
-

Rebuild ontology via logical definitions based on cross-products



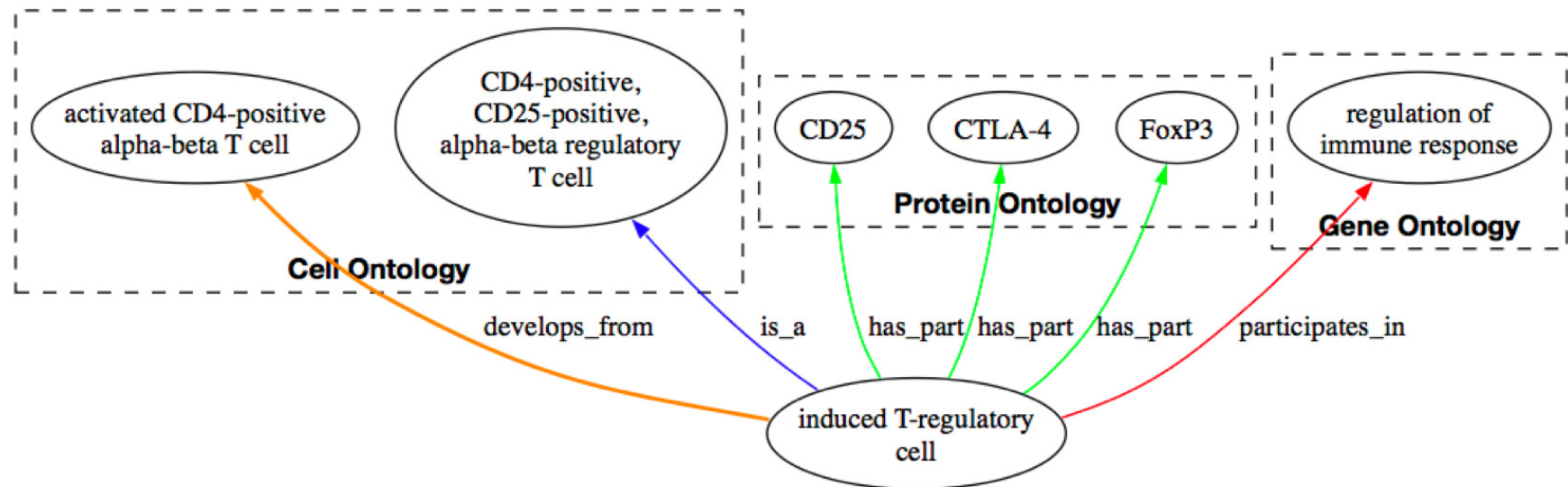
name: CD4-positive, CD25-positive, alpha-beta regulatory T cell

def: "A CD4-positive, CD25-positive, alpha-beta T cell that regulates overall immune responses as well as the responses of other T cell subsets through direct cell-cell contact and cytokine release."

name: induced T-regulatory cell

def: "CD4-positive alpha-beta T cell with the phenotype CD25-positive, CTLA-4-positive, and FoxP3-positive with regulatory function."

Rebuild ontology via logical definitions based on cross-products



Rebuild ontology via logical definitions based on cross-products

```
id: CL:0000902
name: induced T-regulatory cell
def: "CD4-positive alpha-beta T cell with the phenotype CD25-
positive, CTLA-4-positive, and FoxP3-positive with regulatory
function."
intersection_of: CL:0000792 ! CD4-positive, CD25-positive,
alpha-beta regulatory T cell
intersection_of: has_part PRO:000001380 ! CD25
intersection_of: has_part PRO:000001852 ! CTLA-4
intersection_of: has_part PRO:000001350 ! FoxP3
intersection_of: participates_in GO:0050776 ! regulation of
immune response
relationship: develops_from CL:0000896 ! activated CD4-
positive, alpha-beta T cell
```


A two step process: CL 1.5 and CL 2.0

- “CL 1.5” is an intermediate step incorporating the results of the NIAID workshop on immune cell types within a modified CL and continuing to use only `is_a` and `develops_from` relationships.
- The term definitions in CL 1.5 contain most of the information required to build logical definitions.
- In CL 1.5 multiple inheritance for the immune cell types has largely been removed.
- The CL 1.5 improvements to the hematopoietic cells are implemented in the publicly available `cell.obo` file.

A two step process: CL 1.5 and CL 2.0

- “CL 2.0” represents the Cell Ontology rebuilt with logical definitions based cross-products with external ontologies. We are currently seeking dedicated funding for this work.



B cells

- Working Group: Martin Zand, Richard Scheuermann
- Approximately 50 new cell types.
- Better coverage of
 - precursor B cells
 - immature B cells
 - germinal center T cells
 - memory B cells
 - effector B cells
 - plasma cells
- Focused on definition by cell surface marker expression and transcription factor expression, with some cells also described by participation in particular processes.

T cells and NK cells

- Working Group: Penelope Morel, Alexander Diehl
- Approximately 50 new cell types.
- Better coverage of
 - Memory T cells
 - Regulatory T cells
 - Effector T cells
 - NK T cells
 - NK cells
- Focused on definition by cell surface marker expression and transcription factor expression, with some cells also described by participation in particular processes.

Monocytes and Macrophages

- Working Group: Elizabeth Gold and Anastasia Nijnik
- Approximately 30 new cell types.
- Better coverage of
 - tissue macrophages
 - activated macrophages
- Focused on definition by anatomical location with certain types having additional cell expression details.



Dendritic cells

- Working Group: Anna Maria Masci, Lindsay Cowell
- Approximately 30 new cell types.
- Better coverage of
 - conventional dendritic cells
 - plasmacytoid dendritic cells
 - species-specific dendritic cell subtypes
- Focused on definition by cell surface marker expression.
- Paradigm for CL2.0 – DC-CL.

DC-CL

Dendritic cells

[Term]

id: DC_CL:0000010

name: immature CD8_alpha-negative CD11b-positive dendritic cell

def: "immature CD8_alpha-negative CD11b-positive dendritic cell is_a CD8_alpha-negative CD11b-positive dendritic cell that

has_low_plasma_membrane_amount CD80, CD86, and MHCII." [AMM:amm]

comment: Current Opinion In Immunology 2008 20: 61-67

intersection_of: DC_CL:0000012 ! CD8_alpha-negative CD11b-positive dendritic cell

intersection_of: has_low_plasma_membrane_amount DC_CL:0000094 ! CD80

intersection_of: has_low_plasma_membrane_amount DC_CL:0000096 ! CD86

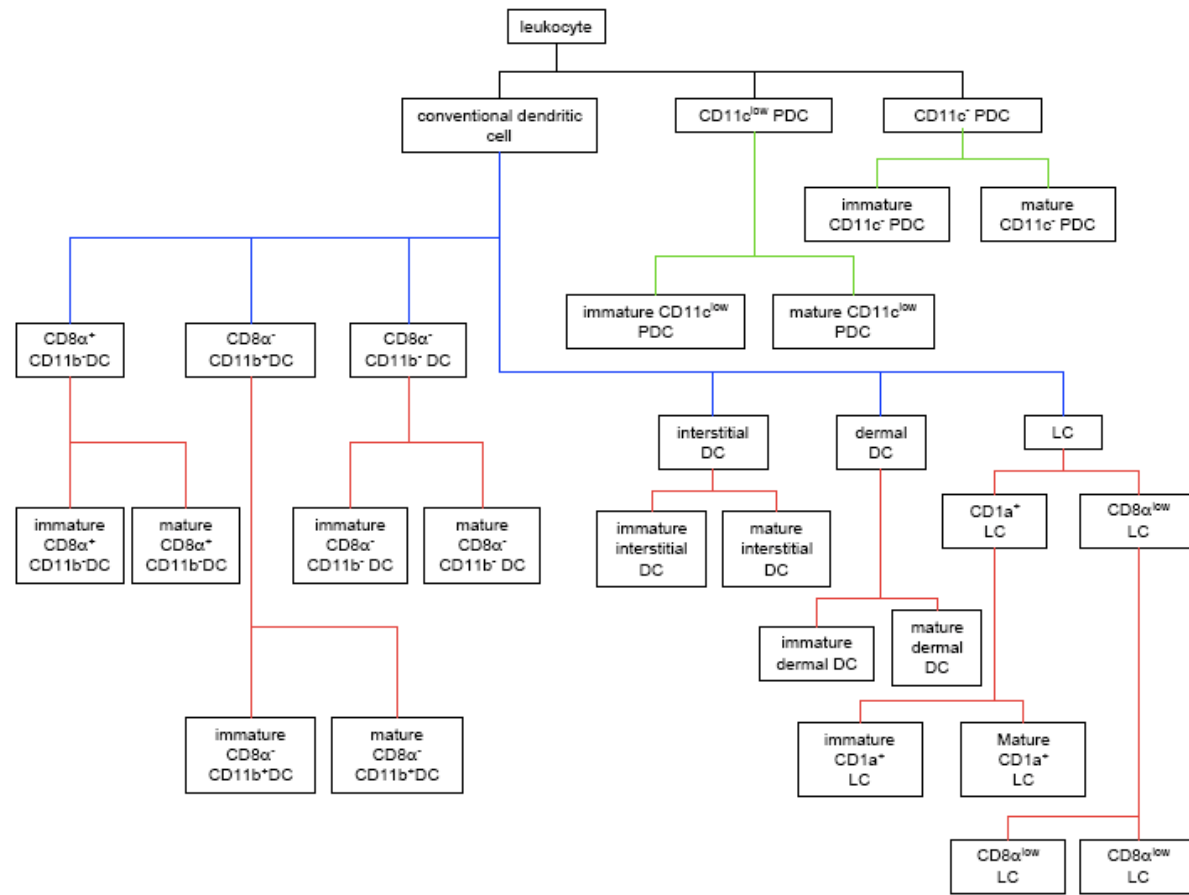
intersection_of: has_low_plasma_membrane_amount DC_CL:0000103 ! MHCII

relationship: develops_from DC_CL:0000041 ! common dendritic precursor

relationship: has_function DC_CL:0000135 ! antigen processing activity

DC-CL

Dendritic cells





Issues

- Some multiple inheritance remains for certain immune cell types in CL 1.5.
- Not all cell types can be easily defined by surface marker phenotype.
- The four subontologies represent four different styles to revise the particular cell types in question.
- Some new cell types are highly specified and species-specific. Data annotators, however, also need general cell types.
 - Example: plasmacytoid dendritic cell

Conclusions & Future Work

- The CL 1.5 improvements to the hematopoietic cells are available in the public cell.obo file.
- The full logical definition/cross-product approach will be applied to all hematopoietic cells: “Hemo-CL”, and eventually the entire Cell Ontology for all cell types.
- The workshop approach with domain experts worked well for this revision.
- Cell definition is complicated by species differences as well as developmental and activation states.

Acknowledgments

Participants in NIAID Cell Ontology Workshop

Martin Zand	Lindsay Cowell
R.E.W Hancock	Jingming Ma
Terry Meehan	Chris Mungall
Bjoen Peters	Christopher Cavnor
Elizabeth Gold	Thomas Kepler
Jeremy Seto	Penelope Morel
Elena Naumova	Anastasia Nijnik
Bali Pulendran	Richard Scheuermann
Alex Sette	Alison Deckhut
Alison Yao	Lynda Chiodetti
Ashley Xia	Timothy Gondre-Lewis
Alexander Diehl	

The Jackson Laboratory

Judith Blake

The Protein Ontology

Cecilia N. Arighi

The Gene Ontology Consortium

Funding

NIAID

NHGRI HG000330 (MGD)

NHGRI HG002273 (GO)



End