

Proposed updated HIPC standard for specifying cell populations

Two fields are currently utilized in ImmPort to specify cell populations in FACS and CyTOF experiments. We here propose more rules for how they should be filled out, in order to enable validation upon time of submission, and to improve the ability to compare datasets from different centers.

Field 1: “Population definition”

This field contains the marker pattern (or ‘gating strategy’) upon which cells are identified as belonging to the target population.

- Individual markers (or gates) should be separated by commas. For example, a possible definition of interferon secreting T cells is: “singlet, viable, CD14-, CD3+, IFNG+, TNFA-”.
- For gates based on protein markers:
 - the protein names should be drawn from (1) PIR (Protein Informatics Resource) : <http://pir.georgetown.edu/>, utilizing “PRO-short labels” or “exact synonyms”, or a value from Table S1 of ‘approved exceptions’ (which should ultimately be integrated into PRO)
 - the protein names should be followed by an indication of the detection level of the marker using the preferred labels as follows:

Marker state	Preferred label	Alternative labels
Negative	-	neg
Positive	+	pos
Low	+-	dim, lo
Intermediate	+~	int, medium, med
High	++	bright, hi

- For gates that are not based on cell protein markers, we have identified the following common cases, which should be identified by the preferred label:

Preferred label	Based on	Alternative labels
lymphocyte	size (FSC vs SSC)	ly, lymp, lymph, lymphocyte, Lymph, Lymphs, Lymp, Lymphocytes
monocyte	size (FSC vs SSC)	mo, mono, monos, MNC, Monocytes, Mono

granulocyte	size (FSC vs SSC)	gran
intact	size (FSC vs SSC)	Intact_cells, Intact_cells_population
singlet	relative dimensions (SSC or FSC A vs H, H vs W, A vs W)	sing, singlets, Singlet, Singlets, doublet_excluded, sing-F, intact_singlet
viable	dye	live, Annexin-, live/dead stain
proliferated	dye	CFSE-, TracerViolet-
infected	Infection marker	
MHC:epitope specific	MHC:epitope staining	

- Additional specifications for a given gate/marker, that is not possible to be captured by any of the above should be included in square brackets. For example: “viable [Annexin negative], CD3+, CD4+, [adherent]”, clarifies in the first bracket how viability is assessed (with Annexin) , and the second set of square brackets allow a user to specify a type of gating falling outside of the current specifications completely (here: adherence properties of cells).

Field 2: “Population name”

This field specifies a name for the cell population being measured, which ideally is the name of an existing cell type. It should be completed as follows

- A. Names should be chosen from: (1) the Cell Ontology (<http://obofoundry.org/ontology/cl.html>) or (2) the list of HIPC population names in the ImmPort spreadsheet (which ultimately should go into the cell ontology or be obsolete).
- B. If no appropriate name for the cell population exist in the lists from (A), it should be reported as:

‘C’ & ‘X’

Where C is the closest parent cell population name from (A) and X is a list of modifiers, specified as described in “cell population” definition above. For example:

‘T cell’ & IFNg+,TNFA-

Specifies T cells producing IFNg but not TNFa

Examples of cell population definitions and names:

<u>Population definition</u>	<u>Population name</u>
CD45+, CD3+, CD8+, CD4-, CD38+	'CD8-positive, alpha-beta T cell' & 'CD38+'
singlet, monocyte, CD3-, CD19-, CD14+, CD16-	'classical monocyte'
viable, singlet, IFNG+, TNF-	'cell' & 'IFNG+, TNF-'
intact, infected[Dengue Virus]	'epithelial cell' & 'infected[Dengue Virus]'