Classification Rules (and Exceptions)

Name:

- Cell type followed by either a column ID (determined by the visual location of the cell) or a numeric identifier to separate out different examples of a given cell type (when the location is not associated with a single column)
- The location is described by one of the 7 column IDs:
  - home
    - Central column
  - A, B, C, D, E, F
    - Always in capital text
- Letter in lower case are used to represent subtypes of a given neuron.
  - Example:
    - T4a vs. T4b: Two subtypes of the T4 neuron.
- The T4 neuron type also has an additional descriptor related to its function.
  - T4 is known to function in motion detection, and the direction selectivity of a given T4 can be determined from its anatomical arborization pattern
    - Reference: Takemura, Bharioke, et al. *Nature* (2013)
  - In this case, one of the following 4 descriptors is added to describe the direction selectivity:
    - fb: front-to-back
    - bf: back-to-front
    - ud: up-to-down
    - du: down-to-up
- Sometimes the relative position of the neuron, within the hexagonal array, is also added, in a case where it may be functionally relevant.
  - We use ant (anterior) or post (posterio)
  - Example: Tm3-A-ant
  - This is always the final descriptor added to the name.

Type:

- Determined through comparison of cellular morphology through random silver staining of different cell types or genetic targeting of fluorescent proteins to different cell types
- Silver staining reference: Fischbach and Dittrich *Cell and Tissue Research* (1989)
- Genetic targeting reference: Unpublished work by Aljoscha Nern (Janelia Farm)

Class:

- Different cell types are grouped into classes by common morphological features (e.g. similar arborization patterns).
- All classes:
  - C: Centrifugal cells
    - Cells though to send feedback signals from the medulla back to the lamina
  - L: Lamina Monopolar cells
    - Cells sending input from the lamina
    - L1, L2, and L3 are strongly retinotopic, with each receiving a large fraction of its input from photoreceptors transducing photon inputs from a single point in visual space.
  - LaWF: Large Wide-field cells
    - Neurons with wide arborizations in the lamina
  - Dm: Distal medulla

        - Cells localized entirely to layers M1 to M6
        - Arborize tangentially and are thought to provide interconnections between different medulla columns, within a region of the medulla (circumscribed by depth)
    - Pm: Proximal medulla
        - Cells localized entirely to layers M8 - M10
        - Like Dm cells; arborize tangentially and are thought to provide interconnections between different medulla columns, within a region of the medulla (circumscribed by depth)
    - Mi: Intrinsic medulla cells
        - Cells that traverse all layers of the medulla (and are generally columnar) but do not send a neurite projection out of the medulla
    - R: Photoreceptors
        - Two color photoreceptors send inputs directly from the retina to the medulla, bypassing the lamina.
    - T1: T-type neuron 1
        - T1 type cells have arborizations within the lamina and medulla. This, together with their cell body location, at the edge of border between the lamina and the medulla provide them with a T-shaped arbor.
    - T2: T-type neuron 2
        - Includes T2 and T2a neurons
            - These neurons appear to be paired, and have very similar shapes and arborizations.
        - Both types arborize within the medulla and lobula, and together with the projection to their cell bodies, also have a T-shaped arbor.
    - T3: T-type neuron 3
        - Another T-type neuron that has arborizations in medulla and lobula
    - T4: T-type neuron 4
        - A T-type neuron that has arborizations within the medulla and lobula plate.
    - Tm: Transmedulla cells
        - Cells that traverse all layers of the medulla (and are generally columnar) but send a neurite projecting out of the medulla, generally to the lobula neuropil.
    - TmY: Transmedulla Y cells
        - Within the medulla, transmedulla Y cells appear generally similar to transmedulla cells. However, their neurite projection splits into two downstream neurites, and projects to both the lobula and lobula plate neuropils (with the skeleton thereby having the shape of an inverted Y)
    - Y: Y cells
        - Named for the same projection to both the lobula and lobula plate neuropils as the TmY cells (and hence an inverted Y arborization), the difference between a Y cell and a TmY is that the arborization within the medulla does not cross all the medulla layers, and stays within the deeper layers.

Superclass:

- Different cell classes can be grouped into a superclass, generally, by their common arborization patterns.
- Some superclass are simply the full name of the class (without including any additional cell classes)
- All Superclasses:
    - Lamina:
        - Includes C, L, LaWF, and T1 cells
        - Cells that have arborizations within the lamina
    - Photoreceptors
        - Includes R cells

- Distal Medulla
  - Includes Dm Cells
- Proximal Medulla
  - Includes Pm Cells
- Lobula T-type
  - Includes three classes of neurons: T2, T2a, T3
  - These are neurons with T shaped arbors that arborize within the medulla and the lobula
- Lobula Plate T-type
  - Includes T4 neurons
  - These are neurons with T shaped arbors that arborize within the medulla and the lobula plate
- Transmedulla Cells
  - Includes Tm neurons
- TmY-type Cells
  - Includes TmY cells
- Y Cells
  - Includes Y cells

Columnar Spread:

- This key takes a value of "Single columnar" or "Multi columnar"
  - Some fraction of the total reconstructed arbor of a cell overlaps with the 7 column ROIs
  - Of this fraction, we find the fraction that is in the column with the greatest overlap.
  - If this fraction is >0.9, then we label the cell as "Single columnar".
    - This implies that >90% of the cell's arbor within the column ROIs is within a single column
    - Note that this ***does not*** imply that the entire cell's volume is within the column ROIs.
  - We apply this analysis to an entire cell type
    - We assume that each cell type must have the same classification (since cells of a given type are defined to be morphologically similar).
    - Further, we focus only on cells within the central 7 columns. These cells are most completely reconstructed.
      - We ignore any cells that are within the outer sections of the reconstruction (no matter the degree of completeness)
    - We choose the value (i.e. either single or multi columnar) based on the fraction found over the majority of cells within the type.
- Exceptions:
  - Each of the neuron types that is an exception to this rule has a biological reason for the exception. We have attempted to detail them here.
  - Dm6:
    - The arborizations for most Dm6 reconstructions are within only a single column. Hence, the quantitative metric correctly labels them as single columnar.
    - However, Dm6 is actually composed of several bushy arbors, each within a single column, connected by short, thin processes. These processes are often missed during EM reconstruction, due to their small size. Nevertheless, the identification of a given (single columnar) arbor as part of a Dm6 cell is unambiguous, due to the agreement with silver stained exemplars.
    - The silver staining results, and genetic labelling of individual neurons clearly show that all Dm4s encompass arborizations is several columns. If such a complete Dm6 was reconstructed, then it would be quantitatively classified as Multi columnar.
    - Therefore, despite the reconstruction of only single column arbors in most cases, we have used an exception to classify Dm6 as Multi columnar.

- Dm1:
  - The case of Dm1 is analogous to Dm6.
  - The EM reconstructions contain only single columnar fragments of the cells. However, by comparison with light microscopy, the cell type is clearly Multi columnar, and we have used an exception to classify it as such.
- C2
  - There is a single C2 neuron arborizing within each retinotopic structure within the lamina.
    - This does not necessarily imply that it should be single columnar.
  - However, C3 neurons have an arborization very similar in structure to its sister neuron, C3.
    - Further, C3 is a Single columnar neuron.
    - The main difference between the two is that C2 has slightly bushier arborizations. Hence, a slightly greater fraction of its arbor is within other column ROIs.
    - Since the 90% threshold is simply an arbitrarily defined cut-off (chosen to select as many classifications correctly as possible), it is possible that C2 simply sits just short of the threshold.
    - Therefore, by observation of its arborization pattern, we decided that it was reasonable to utilize an exception and label it as Single columnar.
- L4
  - This lamina monopolar cell is located in the exterior of the column ROIs (see next section on Columnar location). Hence, its arborizations spread into three surrounding columns.
  - However, the size of the arbor of L4 is similar to that of L1, L2, and L3 - three other lamina monopolar cells, which are clearly single columnar (indeed, the extent of their arbors is used as a working definition of the extent of the column).
    - Therefore, it is reasonable to label this cell as Single columnar (and exterior) - by analogy to the other lamina monopolar cells.
- T2
  - This cell type is quantitatively identified as single columnar. However, it is physically located adjacent (and often intertwined) to two other analogous Lobula T-type cells (T2a, and T3). Both of these cells are Multi columnar.
  - Hence, we assume that this cell must simply sit just over the arbitrary 90% threshold, and - by analogy to the other neurons of the same type - we label this cell as Multi columnar.
- Tm20
  - This cell appears - visually - to be Single columnar. Hence, we have labeled it as such.
  - We believe that the reason for the mis-classification of this cell type may be the small number of examples within the central 7 columns. Since this cell is not found within all columns, we may simply have a small sample error.

Columnar Location:

- This key value takes a value of "Interior" or "Exterior".
  - The metric used to define the value of this key is the fraction of the total reconstructed volume of the cell (within the seven column region) that is within the column ROIs.
  - If this fraction is <0.3, then the cell is "Exterior". If the fraction is >=0.3, then the cell is "Interior".
    - This threshold was chosen because 30% of the total volume within the reconstruction is within the column ROIs. Therefore, a completely random cell which covers the entire region would necessarily have 30% of its volume within the column ROIs. We would reasonably term such a cell as external (since it is not internal to any column, and does not have a concentration of its volume within any of the columns).
  - This metric is not identical to the Columnar spread metric, though it is dependent on it. It is designed to measure the concentration of the volume within the columns vs. outside the columns, whereas

the Columnar spread metric is designed to measure the concentration of volume within a single column vs. the rest.
    - The metric is dependent on the Columnar spread metric in the following way:
        - If the Columnar spread is Multi columnar, the Columnar location is assumed to be "Exterior". This is not due to any deeper biological understanding. It simply allows the smallest set of exception cases.
- Exceptions:
    - Each of the neuron types that is an exception to this rule has a biological reason for the exception. We have attempted to detail them here.
    - L4:
        - L4 is a single columnar neuron (see above for the exception) that - through visual inspection - appears displaced from the single columnar interior neurons associated with the surrounding columns.
        - Hence it appears reasonable to define it as exterior.
        - The reason it is not automatically classified as exterior may have to do with its small total arbor size. Any small differences in the extent of its arbor, on a specific side, would greatly affect the fraction of its volume that is within a specific ROI (when the main body extent is between three different columns).
            - This means that - often - a significant fraction of its arbor is within one specific column, although visually, it appears to have its main body axis located between columns.
            - This randomness is supported by the fact that the identity of the column where it has its maximum extent appears random, and uncorrelated with the retinotopic point in space from which it receives most of its inputs (in the lamina). (This is determined by identifying the L1, L2, and L3 neurons corresponding to the same point in space, and identifying the column to which these single columnar interior neurons send their projections.)
    - Dm8, Tm9, and Tm4:
        - These three cell types are the only examples of multi-columnar interior neurons.
        - Although they send processes to many columns (and hence have a large fraction of their volume outside the columns), they still have a majority of their primary arborization oriented along a single column. Hence, anatomists have classified them as interior (and we include them as an exception to the rule).
    - Tm23/24:
        - We define Tm23/24 as an exterior cell.
        - This neuron has a particularly unique morphology - with only a thin neurite that traverses the entire medulla without having any branching arborizations.
        - Given its structure, it can correspond to two different identified neurons within the light microscopy datasets - either Tm23 or Tm24. However, these cells cannot be distinguished in the limited reconstruction within the medulla (hence the naming convention of the cell type).
        - These thin neurites appear to pass through the medulla at random. This does mean that, in some cases, they travel through a column ROI.
            - Indeed, numerically more axons within the 7 column region traverse column ROIs than not
            - However, the general anatomical impression - including axons outside of the 7 column region suggests that this is simply a sampling bias by the random choice of which 7 columns to reconstruct.
        - Therefore, it is reasonable to say that Tm23/24 is not organized with respect to the medulla columns and, hence, is not interior to the columns (i.e. in this two-valued key, it is labeled as exterior).

- One important caveat is that this population of neurons could actually be composed of both Tm23 and Tm24 neurons, and hence one might be interior, while the other might be exterior (with the combination appearing to be random). However, at present, we cannot distinguish this.

Column ID:

- The column ID is the column in which a specific single-columnar, interior neuron is located in.
  - For multi columnar or exterior neurons, by definition, it is not possible to define such a column, and hence we have left this value empty.
  - We also do not provide an ID for any neurons which are not in the central 7 columns. This is because there are only column ROIs for the central columns. Hence the quantitative overlaps for the outer columns are not available, in our dataset.

Column Volume Fraction:

- The fraction of the total reconstructed volume within the 7 central columns: Home, A, B, C, D, E,F.
- The normalization is made of the total volume within the 7 column reconstructed region (in the DVID database, these volumes can be found by querying the graph_roi_full labelgraph object).
  - Indeed, all the volumes within "column I" can be found by querying the appropriate labelgraph object in the DVID database (i.e. graph_roi_i).
- To compute the fractions used to obtain the Columnar spread, it is necessary to renormalize these values to include only the fraction of the volume within the ROIs.
- However, to compute Columnar location, it is only necessary to sum over all the fractions (and compare against the total that is not included, i.e. that is exterior to the columns).

Synaptic Fractions:

- There are four sets of synaptic fractions that have been precomputed. Each can be computed from the synapse.json, by querying all the columnar / layer ROIs to identify which one (if any) contains each synapse. The synapses are then associated by body, so as to identify the fraction of synapses in a specific ROI, for a given body.
- The location of each synapse can be divided into a presynaptic specialization, a Tbar, and a postsynaptic specialization, a PSD.
  - Column Tbar Fraction:
    - Fraction of Tbars for a given cell, in each of the 7 columns.
    - Details where the cell provides output to other cells.
  - Layer Tbar Fraction:
    - Fraction of Tbars for a given cell, in each of the 10 layers.
    - Details where the cell provides output to other cells.
  - Column PSD Fraction:
    - Fraction of PSDs for a given cell, in each of the 7 columns.
    - Details where the cell receives input from other cells.
  - Layer PSD Fraction:
    - Fraction of Tbars for a given cell, in each of the 10 layers.
    - Details where the cell receives input from other cells.
- In all these vectors, the total may not sum to 1, because of the presence of synapses outside of the ROIs. However, they should always be <= 1.