

Notes for Tumor Extraction Part

Main idea

Think of the procedure as split into two separate parts 1. Tumor Extraction - Find separate tumors (90% done) 2. Tumor Pairing, within samples (0% done)

Things that Remain to be done for Tumor Extraction

1. Flip y-axis in graphical output, is erroneously oriented. [X]
2. Make sure that multiple arguments can be passed (nargs and "*" usage) [X]
3. Change name of graph function
4. Returns error when having no input file, must fix this. Left "TODO" in code for marking
5. Coloring is weird when no tumor spots are present
6. Make sure that original tumor annotation is left in output for those regions that are "masked" out, incase of invasive spread out we still want to be able to use these for analysis
7. Include replicate and patient in input files. [X]
8. Do not make dependent on patient name, remove regexp and replace with simple input handling

Things to be done for Tumor Pairing

1. Decide if this is to be separate or included in analysis
2. Need a way of aligning and pairing up tumors
 - i. One idea is to minimize the following objective function $\min \min \left\{ v_i A u_j \right\} \quad \forall \quad i, j$
where v_i are the vectors from "sample COM" to "tumor foci COM" i for one section and u_i for another section. A would be rotational matrix and once "aligned" it's easy to assign tumors based on their angular value. A cost function for correctly aligning the templates (especially if equidispersed number of loci are present) must be added
 2. Other idea is to convert to polar coordinates and then align
 3. If one could first align sections, just looking at angular position would be sufficient. Perhaps two-step procedure in the sense
 - i. Align sections
 - ii. Find best overlapping tumors.

This is also tractable since one then could assign tumors that are present in one replicate but not other to the same tumor in another, approximate to nearest foci.

4. KABSCH Algorithm basically do this LINK One could probably subsample the set of points, in order to reduce the covariance matrix computation (and also get equal set of points)