10-3-12 Master database file and folder layout

This document describes locations, file types, naming, and sample name keys for the L1, L2, and L3 data

Folder: SAIC GBM/

*10-3-12-Xeno-Aliases (JCA)* (This is the key used to identify barcode array combination for each tumor sample, and it’s biological replicates). Note all Barcodes beginning with a 2 should have a leading zero.

*STK 90028 Array Layout.txt* Array layout describing phosphorylatable spot location on the grid, along with sequence of the peptides and other data specific to the STK arrays (applies to all images starting with 02\*\*\*\*\*\*\*) or 2\*\*\*\*\*\*\* barcodes).

*PTK 86312 Array Layout.txt* Array layout describing phosphorylatable spot location by X, Y coordinate on the grid, along with sequence of the peptides, and other data specific to the PTK arrays (applies to all image files starting with 63\*\*\*\*\*\*\*)

Folder: SAIC GBM/OrthotopicRawData(L1)/

Folder: SAIC GBM/OrthotopicRawData(L1)/PTK SAIC Ortho-GBMs/

Barcode specific folders (ex. Name: (generally 3 barcodes per folder) with subfolders(per each batched run) including image folders (*example name 631008610\_631008611\_631008612-on 1200PTKlys50v01-run 110419084153*) containing images taken over time (*in folder example name.* *631008610\_631008611\_631008612-ImageResults).*

Folder SAIC GBM/OrthotopicRawData(L1)/STK SAIC Ortho-GBMs/

Barcode specific folders (3 barcodes per folder, with 4 arrays each) with sub folders(per each batched run) including image folders (*example name 028017506\_028017507\_028017508-on STKlys50disp12-run 110420124824*) containing images taken over time (*in folder example name.* *028017506\_028017507\_028017508-ImageResults).*

Folder SAIC GBM/L2 Median Signal minus background’

Barcode and Array labeled Median Signal minus background values, per sample, per cycle number are included. Each value (Median Signal minus background is essentially ‘brightness’ of each spot, per sample (see row (give specific row numbers for barcode and array, and cycle number) labels that designate values. Cycles increase with time sequentially (where exposure time is 50ms), except for the final cycle, usually 94 for PTK, and 124 for STK (note exceptions listed specifically for barcodes). In these final cycles (94, 124, or ‘0’ as designated) the pictures vary with exposure time (see row l# label exposure time).

Folder SAIC GBM/L3data/9-7-12KinPTKSTKvINIs100.xlsx

*9-7-12KinPTKSTKvINIs100.xlsx* (L3 data*) includes the following sheets:*

*A. PTK vINI*

*B. PTK s100*

*C. STK s100*

*The PTK vINI, PTK s100 and STK s100 values are listed per tumor sample name, and tumor biological replicate name per column in this file, not per barcode and array as the SAIC GBM\L2 Median Signal minus background’ values are.  Peptides spots are listed by PamGene peptide ID per row (these ID's can be cross referenced for further spot data via the 'PTK 86312 Array Layout.txt' or 'STK 90028 ArrayLayout.txt' files to determine peptide sequence, name, and other information). STK chips do not have a calculated  vINI due to experimental constraints, thus there is no STK vINI sheet):*

*Descriptions of the data in the 9-7-12KinPTKSTKvINIs100.xlsx  are as follows:*

*A. PTK vINI(initial velocity) is best described as a kinetic read . It integrates the earliest (cycle number) spot brightness values and calculates a curvilinear fit per peptide spot per sample.*

*B. PTK s100(postwash steady state for the Tyrosine chip) uses later exposures, after the lysate has been rinsed from the array, and integrates, 10, 20,50,100, and 200ms exposure into one value (approximately slope \*100) per peptide spot per sample*

*C. STK s100(postwash steady state, for the STK chip) uses later exposures, after the lysate has been rinsed from the array, and integrates, 10, 20,50,100, and 200ms exposure into one value (approximately slope \*100) per peptide spot per sample.*

Image files (L1) naming:

Each image file is 710-712kb tif image file (697x520 pixels), although when attaching (as seen in the attached file) they can be as large as 960kb

1. Each image is of the entire array (all spots) at a specific cycle number, and exposure time.

File Naming: The files are named as such:

028017507\_W1\_F1\_T100\_P94\_I204\_A30.tif

All files that start with 02\*\*\*\*\*\*\* (or 2\*\*\*\*\*\* (leading zero dropped) in the *9-6-12 Kinomics UIDs-Barcodes-key Orthopic Raw Data (EXCEL).xlsx* lookup/key file are STK (serine/threonine kinome) files and their layout (X, Y grid coordinates per peptide spot) is referenced by the *90028 Array Layout.txt* file.

All files that start with **63**\*\*\*\*\*\*\* are PTK files and their layout (X, Y grid coordinates per peptide spot) is referenced by the *86312 Array Layout.txt* file.

“028017507\_W1” is the chip serial/barcode number (which contains 4 arrays, also can be extrapolated as array ‘x’ location within the pamstation). 4 independent samples are run on each chip in one of the 4 arrays (W1, W2, W3, W4, in the *9-6-12 Kinomics UIDs-Barcodes-key Orthopic Raw Data (EXCEL).xlsx* key this is referred to as A1, A2, A3, A4 respectively. These are notations Well (W) and Array(A) are interchangeable). The barcode and well combination is all that is required to determine which sample you are looking at. For example in this case 028017507\_W1 in the the *9-6-12 Kinomics UIDs-Barcodes-key Orthopic Raw Data (EXCEL).xlsx* can be found on row 113 as:

Barcode Array(well) sample name biological rep UID

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 28017507 | A1 | GBM ortho 6 | GBM ortho 6-B | kinomicsGil00038 |  |

* *Note the leading zero is dropped on in this file 028017507 becomes 28017507*

“F1” is not used

“T100” is the image exposure time (T100 in this case is 100ms, where T10,T20,T50,T100, and T200 are all valid millisecond exposure times).

“P94” denotes the protocol step in the Evolve program. (in PTK experiments these will normally range from P32 to P94\* in STK these will normally range from P62 to P124\*). The final protocol step takes multiple exposure times in increasing exposure times at the final cycle. (i.e. @ P94 (final cycle)in a PTK capture, or P124(final cycle) in the STK the 10ms exposure is taken before the 20ms, then the 50, then 100, then 200) even though the P94 or P124(or P0 in the 3 STK barcodes listed in red below )is the same. For earlier image captures (all 50ms) the P\* determines the sequence (P32 then P37, the P42, etc).

“I204” describes the sequence of the image was captured (in PTK experiments these will range from I1 to I216\* in STK these will range from I1 to I144\*)

\*Sometimes an error can cause a step to be re-run/restarted using a modified protocol, and in these rare cases the P\*\* and I\* numbers may not be comparable across different barcode well combinations (although the 3 barcodes in affected experiments will all have comparable P\*\* and I\*\*\* labelings).

For example a software restart affected the 3 barcodes and associated image file listed immediately below where the P\*\* number is ‘shifted’ by -124 prior to the STK postwash image capture (note P\*\* before this are not used in STK analysis, and were not captured).

This only effects the 3 following barcodes (and all wells(W1-W4) in them):

028017114, 028017115, 028017116 where P0 in the file name is equivalent to P124 in all the other unaffected STK files (02\*\*\*\*\*\*\*)

Any other data not referenced via barcode\_well with the UID key is not used or may include test wells. Also the 3 barcodes below are included in the L2 data (but not referenced via the UID/sample name key , and were not used for further analysis due to a mechanical failure during the run):

028017221, 028017222, 028017223

Joshua Anderson

[janders7@uab.edu](mailto:janders7@uab.edu)

[janders211@gmail.com](mailto:janders211@gmail.com)