Supplementary Information

An integrated one step system to extract, analyze and annotate all relevant information from image-based cell screening of chemical libraries

Obdulia Rabal, Wolfgang Link, Beatriz G. Serelde, James R. Bischoff and Julen Oyarzabal[*]

Experimental Therapeutics Programme, Centro Nacional de Investigaciones Oncológicas (CNIO), Melchor Fernández Almagro 3, 28029 Madrid, Spain.

Table of Contents

Supplementary Table S1- Cellular features collected by AttoVision software within BD Pathway Bioimager.

#	Cellular Feature	Definition
1	Nucleus_Pixels	Number of pixels of the nuclear region (Nuclear Area)
2	Nucleus_Verticies	Number of pixels on the inner and outer boundaries of the nuclei.
3	DAPI_Intensity	Average pixel intensity within the nuclear region on Ch1 (DAPI)
4	Cytoplasm_Pixels	Number of pixels of the cytoplasmic region
5	Cytoplasm_Verticies	Number of pixels on the inner and outer boundaries of the cytoplasm
6	GFP cyt_Intensity	Average pixel intensity within the cytoplasmic region on Ch2 (GFP)
7	GFP nucl_Intensity	Average pixel intensity within the nuclear region on Ch2 (GFP)
8	Nucleus_W	Bounding Box Width of the nuclear region
9	Nucleus_H	Bounding Box Height of the nuclear region
10	Cytoplasm_W	Bounding Box Width of the cytoplasmic region
11	Cytoplasm_H	Bounding Box Height of the cytoplasmic region

Supplementary Table S2- Cellular features collected by HCS tool for automated image analysis.

#	Cellular Feature	Definition
1	nuclear_plus_gfp_SumOfPixelIntensities	Sum of pixel intensities within the nucleus and cytoplasm on Ch2 (GFP)
2	nuclear_plus_gfp_SumOfSquaredPixelIntensities	Sum of squared pixel intensities within the nucleus and cytoplasm on Ch2 (GFP)
3	nuclear_plus_gfp_SumOf3rdPowerPixelIntensities	Sum of 3rd Power Pixel Intensities within the nucleus and cytoplasm on Ch2 (GFP)
4	nuclear_plus_gfp_SumOf4thPowerPixelIntensities	Sum of 4th Power Pixel Intensities within the nucleus and cytoplasm on Ch2 (GFP)
5	nuclear_plus_gfp_PixelIntensityMean	Average Pixel Intensity within the nucleus and cytoplasm on Ch2 (GFP)
6	nuclear_plus_gfp_PixelIntensityVariance	Variance in pixel intensity within the nucleus and cytoplasm on Ch2 (GFP)
7	nuclear_gfp_SumOfPixelIntensities	Sum of pixel intensities within the nucleus on Ch2 (GFP)
8	nuclear_gfp_SumOfSquaredPixelIntensities	Sum of squared pixel intensities within the nucleus on Ch2 (GFP)
9	nuclear_gfp_SumOf3rdPowerPixelIntensities	Sum of 3rd Power Pixel Intensities within the nucleus on Ch2 (GFP)
10	nuclear_gfp_SumOf4thPowerPixelIntensities	Sum of 4th Power Pixel Intensities within the nucleus on Ch2 (GFP)
11	nuclear_gfp_PixelIntensityMean	Average Pixel Intensity within the nucleus on Ch2 (GFP)
12	nuclear_gfp_PixelIntensityVariance	Variance in pixel intensity within the nucleus on Ch2 (GFP)
13	nuclear_plus_dapi_SumOfPixelIntensities	Sum of pixel intensities within the nucleus and cytoplasm on Ch1 (DAPI)
14	nuclear_plus_dapi_SumOfSquaredPixelIntensities	Sum of squared pixel intensities within the nucleus and cytoplasm on Ch1 (DAPI)
15	nuclear_plus_dapi_SumOf3rdPowerPixelIntensities	Sum of 3rd Power Pixel Intensities within the nucleus and cytoplasm on Ch1 (DAPI)
16	nuclear_plus_dapi_SumOf4thPowerPixelIntensities	Sum of 4th Power Pixel Intensities within the nucleus and cytoplasm on Ch1 (DAPI)
17	nuclear_plus_dapi_PixelIntensityMean	Average Pixel Intensity within the nucleus and cytoplasm on Ch1 (DAPI)
18	nuclear_plus_dapi_PixelIntensityVariance	Variance in pixel intensity within the nucleus and cytoplasm on Ch1 (DAPI)
19	nuclear_dapi_SumOfPixelIntensities	Sum of pixel intensities within the nucleus on Ch1 (DAPI)
20	nuclear_dapi_SumOfSquaredPixelIntensities	Sum of squared pixel intensities within the nucleus on Ch1 (DAPI)
21	nuclear_dapi_SumOf3rdPowerPixelIntensities	Sum of 3rd Power Pixel Intensities within the nucleus on Ch1 (DAPI)
22	nuclear_dapi_SumOf4thPowerPixelIntensities	Sum of 4th Power Pixel Intensities within the nucleus on Ch1 (DAPI)
23	nuclear_dapi_PixelIntensityMean	Average Pixel Intensity within the nucleus on Ch1 (DAPI)
24	nuclear_dapi_PixelIntensityVariance	Variance in pixel intensity within the nucleus on Ch1 (DAPI)
25	nuclear_BlobArea	Nuclear Area (pixels)
26	nuclear_BlobPerimeter	Nuclear Perimeter
27	nuclear_BlobCircularity	Nuclear circularity
28	nuclear_BoundingBoxWidth	Nuclear Bounding Box Width
29	nuclear_BoundingBoxHeight	Nuclear Bounding Box Height
30	nuclear_EquivalentDiameter	Diameter of the circle of the same area as nucleus

31	nuclear_ConvexArea	Area of the convex hull (area of the smallest convex set of pixels containing the entire nuclear object)
32	nuclear_Solidity	Nuclear area divided by the area of the convex hull
33	nuclear_FormFactor	Measure of nuclear circularity. (4*pi*area/perimeter squared)
34	nuclear_DistanceToNearestNeighborCentroid	Distance to nearest nucleus (Centroid - Centroid)
35	nuclear_DistanceToNearestNeighborBoundary	Distance to nearest nucleus (Boundary - Boundary)
36	nuclear_dapi_SumOfPixelIntensities_T	Sum of pixel intensities within the nucleus on Ch1 (DAPI) with Autothresholding segmentation
37	nuclear_dapi_SumOfSquaredPixelIntensities_T	Sum of squared pixel intensities within the nucleus on Ch1 (DAPI) with Autothresholding
38	nuclear_dapi_SumOf3rdPowerPixelIntensities_T	Sum of 3rd Power Pixel Intensities within the nucleus on Ch1 (DAPI) with Autothresholding
39	nuclear_dapi_SumOf4thPowerPixelIntensities_T	Sum of 4th Power Pixel Intensities within the nucleus on Ch1 (DAPI) with Autothresholding
40	nuclear_dapi_PixelIntensityMean_T	Average Pixel Intensity within the nucleus on Ch1 (DAPI) with Autothresholding segmentation
41	nuclear_dapi_PixelIntensityVariance_T	Variance in pixel intensity within the nucleus on Ch1 (DAPI) with Autothresholding
42	nuclear_BlobArea_T	Nuclear Area (pixels) with Autothresholding segmentation
43	nuclear_BlobPerimeter_T	Nuclear Perimeter with Autothresholding segmentation
44	nuclear_BlobCircularity_T	Nuclear circularity with Autothresholding segmentation
45	nuclear_EquivalentDiameter_T	Diameter of the circle of the same area as nucleus with Autothresholding segmentation
46	nuclear_ConvexArea_T	Area of the convex hull with Autothresholding segmentation
47	nuclear_Solidity_T	Nuclear area divided by the area of the convex hull with Autothresholding segmentation
48	nuclear_FormFactor_T	Measure of nuclear circularity. (4*pi*area/perimeter squared) with Autothresholding

Supplementary Table S3 – Results from BD Pathway Bioimager

	Total Compounds
Total Compounds in HCS	33,992
Primary Hits identified with BD Pathway	448
Hits from VS analysis complementing primary Hits	91
Confirmed secondary Hits	228
PI3K Inhibitors	54

Supplementary Table S3 Footnote
Of the 33,992 compounds tested, 448 (1.32%) met the active criterion defined using the BD Pathway Bioimager and produced 60% inhibition of Akt-mediated FOXO-GFP translocation. To maximize, in terms of quality and quantity, the information obtained from this screening; once hits from primary screening were selected an additional step was performed: identification of false positives (FP) and false negatives (FN). Thus, all hits from primary screening together with those initially non-hits rescued by virtual screening (VS) techniques were further confirmed in a second HCS round. Virtual screening, and visual inspection of a number of wells, complemented the initial hit pool with 91 additional compounds. For those 228 confirmed positive hits, PI3K inhibitory activity was experimentally estimated through its corresponding biochemical assay[14]

Supplementary Table S4– *p*-Values obtained for each of the 13 cytological features initially considered for cytotoxicity assessment for the 168-membered library. Different Student's *t*-tests were conducted for different plates, comparing reference cytotoxic compound *vs* untreated (controls) wells and non-cytotoxic compounds *vs* untreated (controls). Bold values show significance level of 1% or better (*p*-value < 0.01). 7 cytological features (*in bold*) were finally selected.

	Cytotoxic *vs* untreated			Non-cytotoxic *vs* untreated		
	Plate I (n=14)	Plate II (n=9)	Plate III (n=13)	Plate I (n=92)	Plate II (n=19)	Plate III (n=83)
Cytological Feature (All nuclear)	*P*	*P*	*P*	*P*	*P*	*P*
nuclear_dapi_PixelIntensityVariance_T_Mean	**4.E-15**	**1.E-03**	**2.E-06**	0.92	0.03	0.38
nuclear_EquivalentDiameter_T_Mean	**9.E-10**	**6.E-03**	**3.E-08**	0.94	0.11	0.31
nuclear_BlobPerimeter_T_Mean	**3.E-09**	**4.E-03**	**5.E-09**	0.94	0.20	0.17
nuclear_BlobArea_T_Mean	**3.E-09**	**4.E-03**	**4.E-08**	0.77	0.10	0.30
nuclear_ConvexArea_T_Mean	**5.E-09**	**4.E-03**	**2.E-08**	0.75	0.12	0.27
nuclear_dapi_PixelIntensityMean_T_Mean	**1.E-07**	**3.E-04**	**4.E-08**	0.99	0.43	0.14
Cell_Density_Mean	**3.E-05**	**1.E-03**	**4.E-04**	0.10	0.90	0.29
nuclear_dapi_SumOf4thPowerPixelIntensities_T_Mean	**5.E-06**	0.04	**2.E-07**	0.81	0.81	0.15
nuclear_BlobCircularity_T_Mean	**2.E-05**	0.02	**3.E-05**	0.86	0.45	0.45
nuclear_FormFactor_T_Mean	**2.E-05**	0.02	**3.E-05**	0.86	0.45	0.45
nuclear_dapi_SumOf3rdPowerPixelIntensities_T_Mean	**4.E-05**	0.15	**4.E-07**	0.86	0.59	0.16
nuclear_dapi_SumOfSquaredPixelIntensities_T_Mean	**4.E-03**	0.98	**3.E-06**	0.92	0.34	0.19
nuclear_dapi_SumOfPixelIntensities_T_Mean	0.09	0.07	0.01	0.95	0.15	0.35
nuclear_Solidity_T_Mean	0.77	0.73	0.04	0.84	0.32	0.89

Supplementary Table S5 - Confusion matrix, kappa values, sensitivity and specificity obtained for each of the different definitions established to classify compounds as cytotoxic or not from the 168-membered reference library. Each of the seven significant cytological features found with Student's *t*-test, was defined as outlier if the mean value for a well was 1, 1.5, 2 or 3 standard deviations (SD) away from the mean values of the negative controls. In different models, a minimum number of outlier features was required to classify compounds as cytotoxic or not and the corresponding confusion matrix resulting from classification was calculated.

|	Number of outlier features required	TP	TN	FP	FN	*Kappa*	Sensitivity	Specificity
1 SD	7	12	155	1	0	0.96	100	99
	6	12	141	15	0	0.57	100	90
	5	12	137	19	0	0.51	100	88
	4	12	119	37	0	0.31	100	76
	3	12	116	40	0	0.29	100	74
	2	12	102	54	0	0.21	100	65
	1	12	67	89	0	0.10	100	43
1.5 SD	7	12	156	0	0	1.00	100	100
	6	12	144	12	0	0.72	100	92
	5	12	142	14	0	0.70	100	91
	4	12	134	22	0	0.63	100	86
	3	12	127	29	0	0.58	100	81
	2	12	117	39	0	0.53	100	75
	1	12	89	67	0	0.42	100	57
2 SD	7	11	156	0	1	0.95	92	100
	6	**12**	**152**	**4**	**0**	**0.86**	**100**	**97**
	5	12	147	9	0	0.76	100	94
	4	12	141	15	0	0.69	100	90
	3	12	140	16	0	0.68	100	90
	2	12	126	30	0	0.57	100	81
	1	12	103	53	0	0.47	100	66
3 SD	7	9	156	3	3	0.82	75	98
	6	12	152	2	0	0.86	100	99
	5	12	154	2	0	0.92	100	99
	4	12	151	5	0	0.84	100	97
	3	12	150	6	0	0.82	100	96
	2	12	141	15	0	0.69	100	90
	1	12	123	33	0	0.56	100	79

II. Supplementary Figures
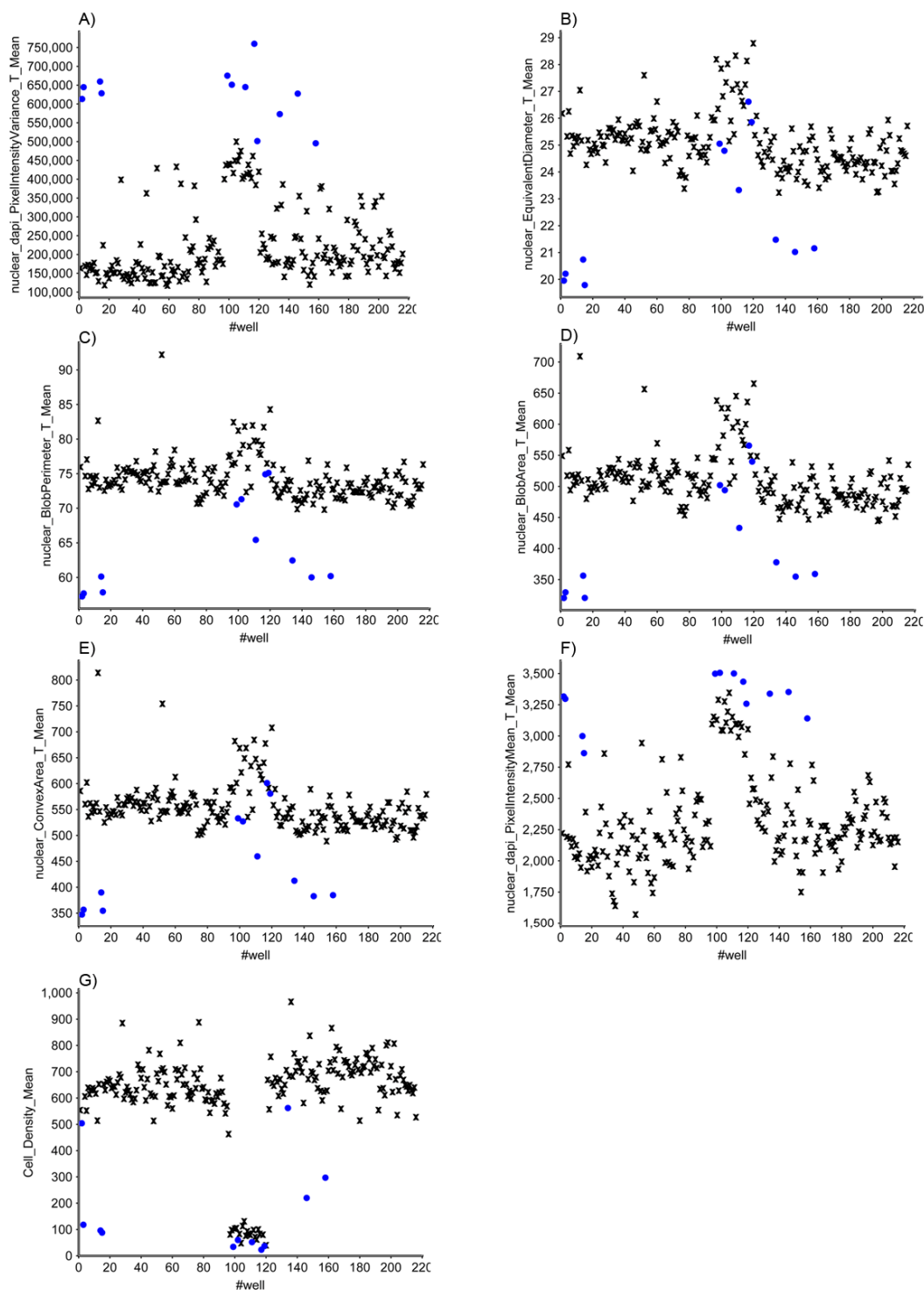
Supplementary Fig. S1. Key cytological features plot for phenotypic profiling as cytotoxic compound plot for the 12 reference cytotoxic compounds (blue dots) and for the 156 non-cytotoxic compounds and negative controls (black crosses); clearly this is playing a role.

Actual Value

		Positive	Negative	Total
	Positive	TP	FP	p'
Prediction Outcome	Negative	FN	TN	n'
	Total	p	n	N

Supplementary Fig. S2. Confusion Matrix, with the number of true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN).
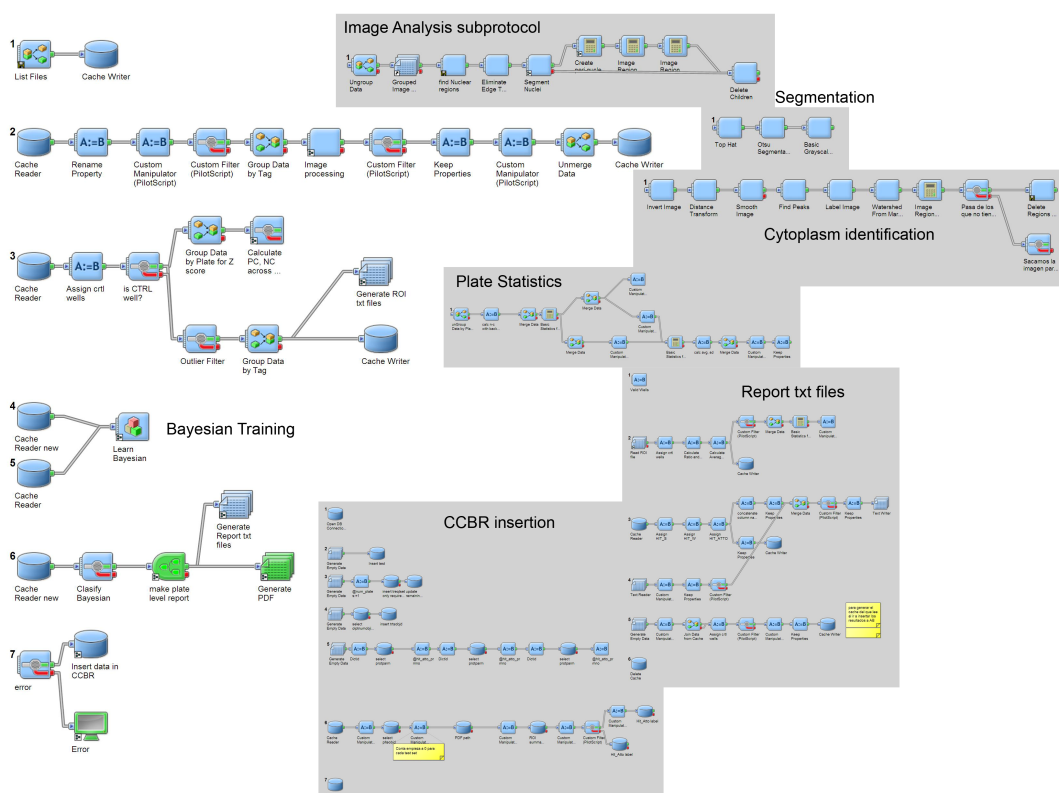
Supplementary Fig. S3. Screenshot of the Pipeline Pilot protocol. Many steps are collapsed into different subprotocols, some of them are exemplified in those gray-background images.