Supplemental Materials

Molecular Biology of the Cell

Seal *et al.*

Supplementary Information

From Pixels to Phenotypes: Integrating Image-Based Profiling with Cell Health Data Improves Interpretability

Srijit Seal[1,2], Jordi Carreras-Puigvert[3], Shantanu Singh[1], Anne E Carpenter[1], Ola Spjuth*[3], Andreas Bender*[2]

[1]Imaging Platform, Broad Institute of MIT and Harvard, Cambridge MA, USA

[2]Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge, United Kingdom

[3]Department of Pharmaceutical Biosciences and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

Cell Painting; Cell Health; BioMorph, Toxicity Prediction; Cell Morphology; Interpretability; Image-based profiling

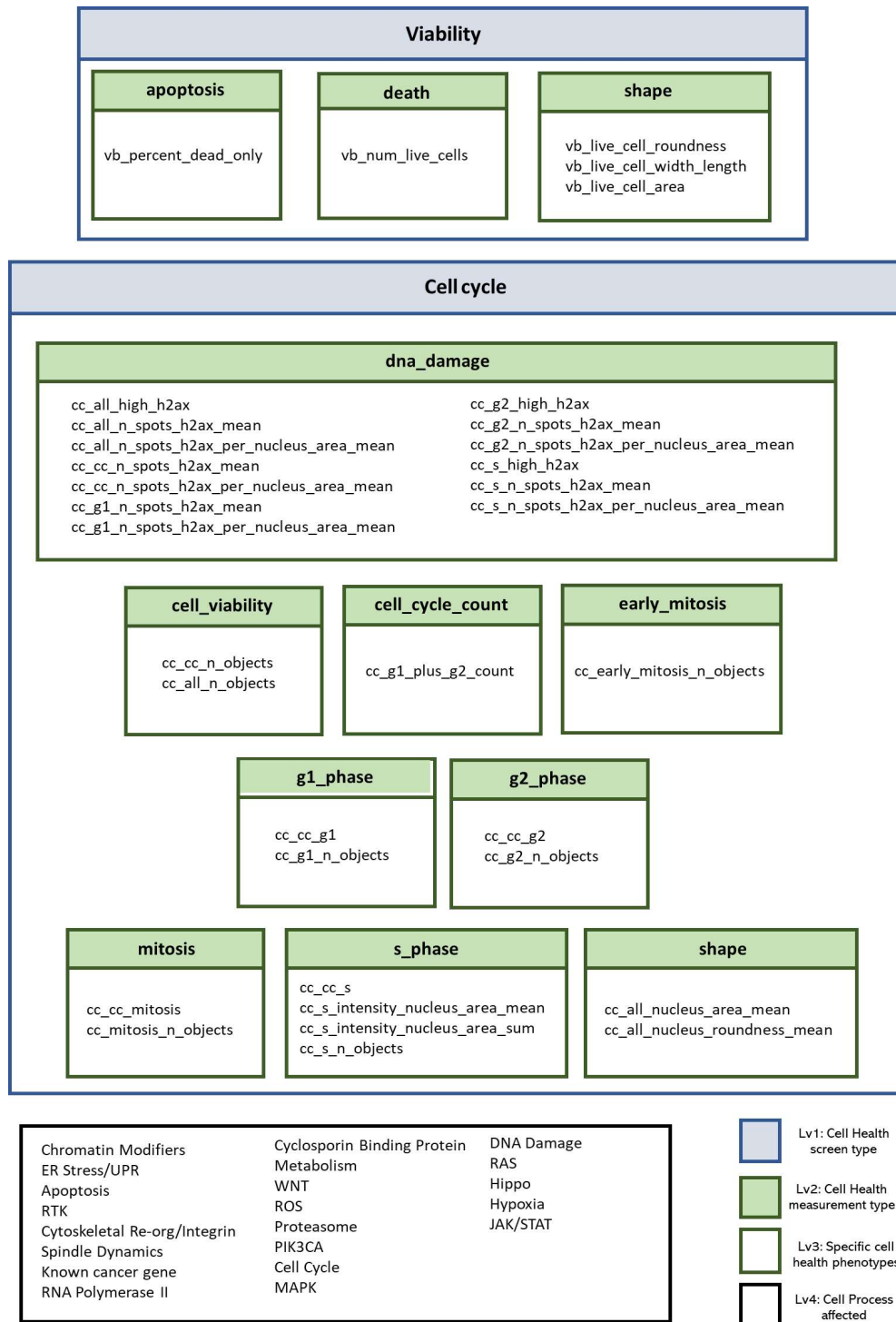* Email: ola.spjuth@farmbio.uu.se, ab454@cam.ac.uk

Viability

apoptosis

vb_percent_dead_only

death

vb_num_live_cells

shape

vb_live_cell_roundness
vb_live_cell_width_length
vb_live_cell_area

Cell cycle

dna_damage

cc_all_high_h2ax
cc_all_n_spots_h2ax_mean
cc_all_n_spots_h2ax_per_nucleus_area_mean
cc_cc_n_spots_h2ax_mean
cc_cc_n_spots_h2ax_per_nucleus_area_mean
cc_g1_n_spots_h2ax_mean
cc_g1_n_spots_h2ax_per_nucleus_area_mean

cc_g2_high_h2ax
cc_g2_n_spots_h2ax_mean
cc_g2_n_spots_h2ax_per_nucleus_area_mean
cc_s_high_h2ax
cc_s_n_spots_h2ax_mean
cc_s_n_spots_h2ax_per_nucleus_area_mean

cell_viability

cc_cc_n_objects
cc_all_n_objects

cell_cycle_count

cc_g1_plus_g2_count

early_mitosis

cc_early_mitosis_n_objects

g1_phase

cc_cc_g1
cc_g1_n_objects

g2_phase

cc_cc_g2
cc_g2_n_objects

mitosis

cc_cc_mitosis
cc_mitosis_n_objects

s_phase

cc_cc_s
cc_s_intensity_nucleus_area_mean
cc_s_intensity_nucleus_area_sum
cc_s_n_objects

shape

cc_all_nucleus_area_mean
cc_all_nucleus_roundness_mean

Chromatin Modifiers
ER Stress/UPR
Apoptosis
RTK
Cytoskeletal Re-org/Integrin
Spindle Dynamics
Known cancer gene
RNA Polymerase II

Cyclosporin Binding Protein
Metabolism
WNT
ROS
Proteasome
PIK3CA
Cell Cycle
MAPK

DNA Damage
RAS
Hippo
Hypoxia
JAK/STAT

Lv1: Cell Health screen type

Lv2: Cell Health measurement type

Lv3: Specific cell health phenotypes

Lv4: Cell Process affected

Figure S1: BioMorph terms' hierarchy with relationships for (a) Level 1: Cell Health assay type, (b) Level 2: Cell Health measurement type, (c) Level 3: Specific Cell Health phenotypes, (d) Level 4: Cell process affected (representative of morphological characteristics)
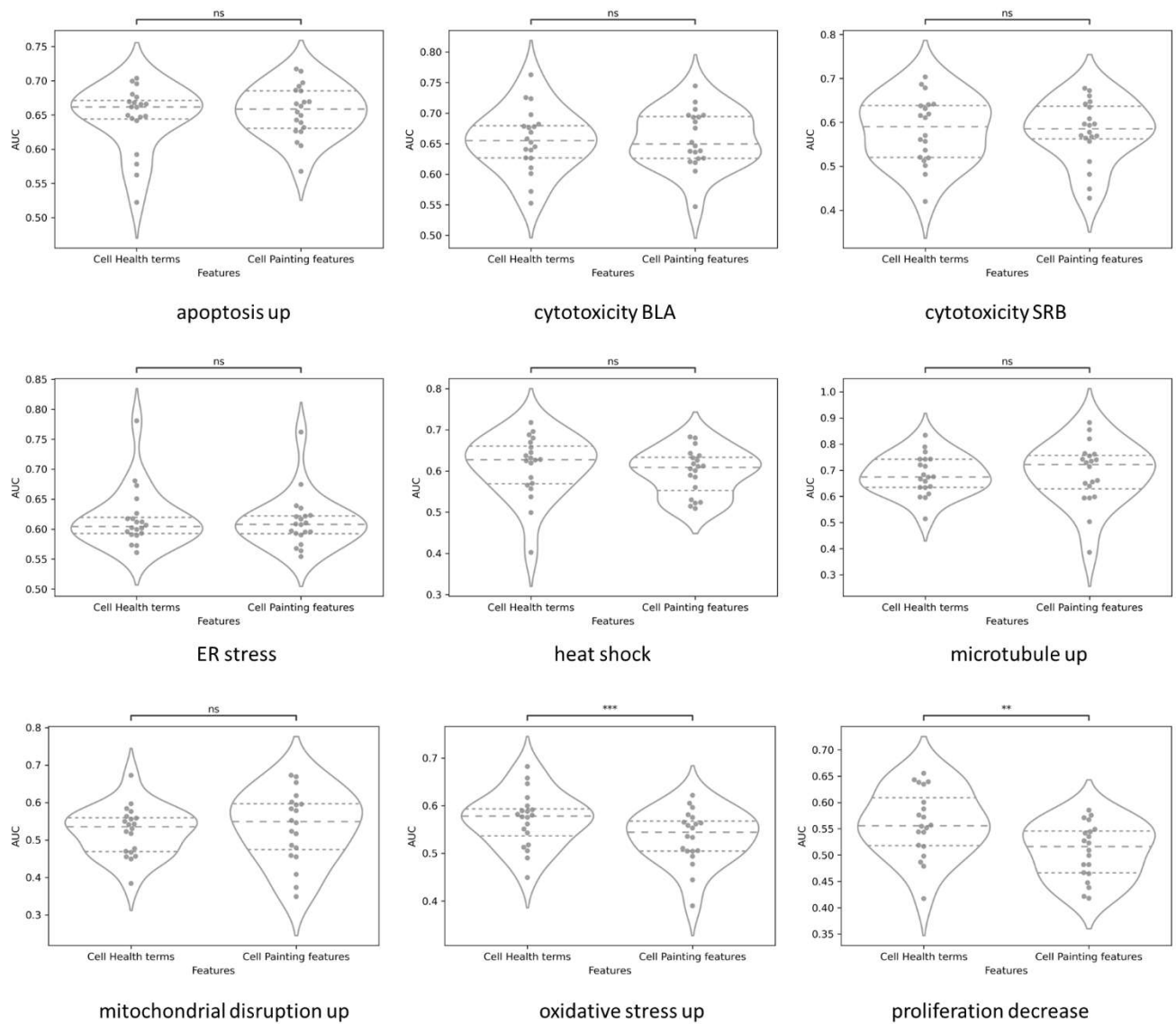
p-value annotation legend:
ns: p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

Figure S2: Distribution AUC for both models, namely, Cell Painting features and BioMorph terms as features when predicting 9 different broad biological activities in 5 times repeated 4-fold nested cross-validation.