nature portfolio

https://doi.org/10.1038/s41596-021-00653-8

Supplementary information

Pooled genetic perturbation screens with image-based phenotypes

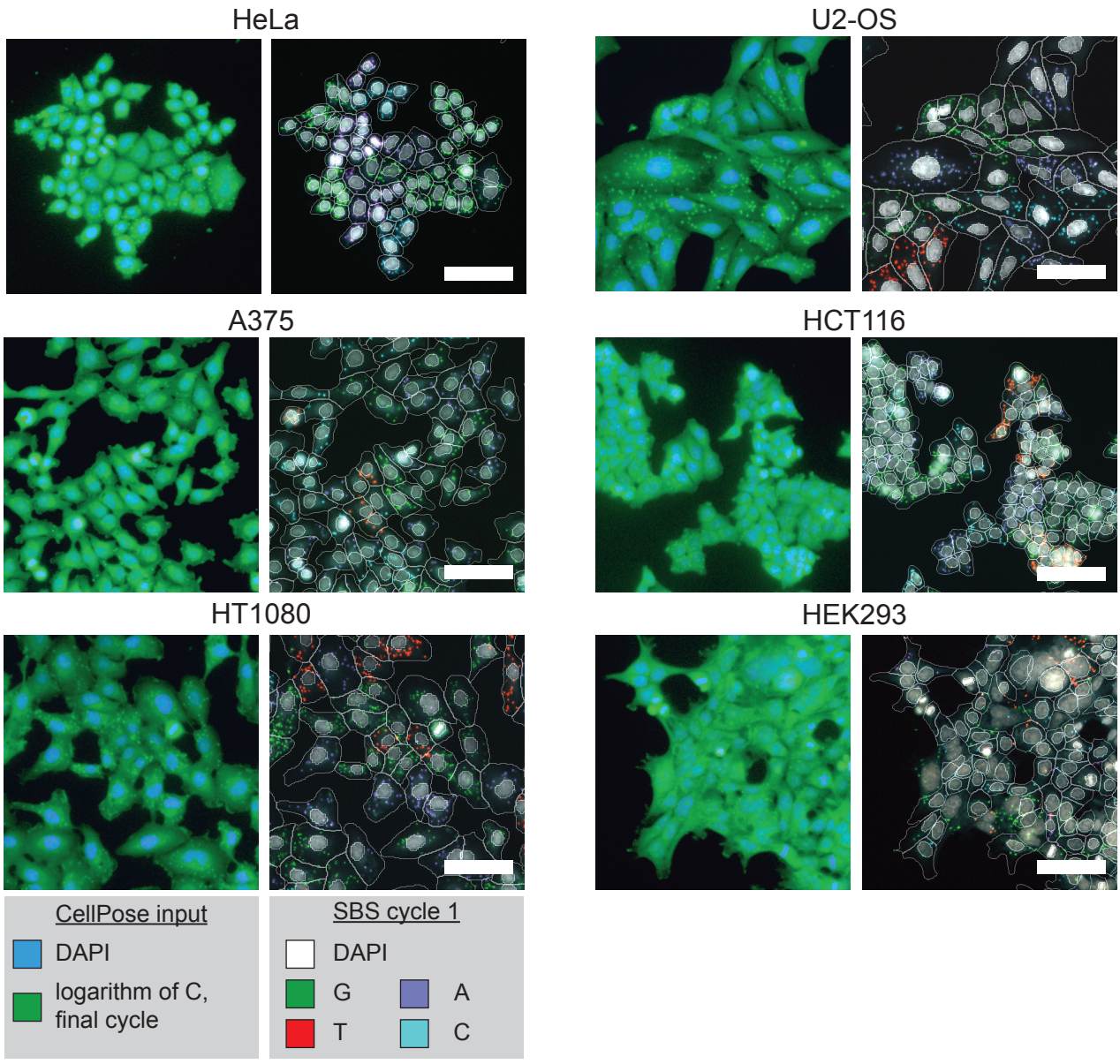
In the format provided by the authors and unedited

a

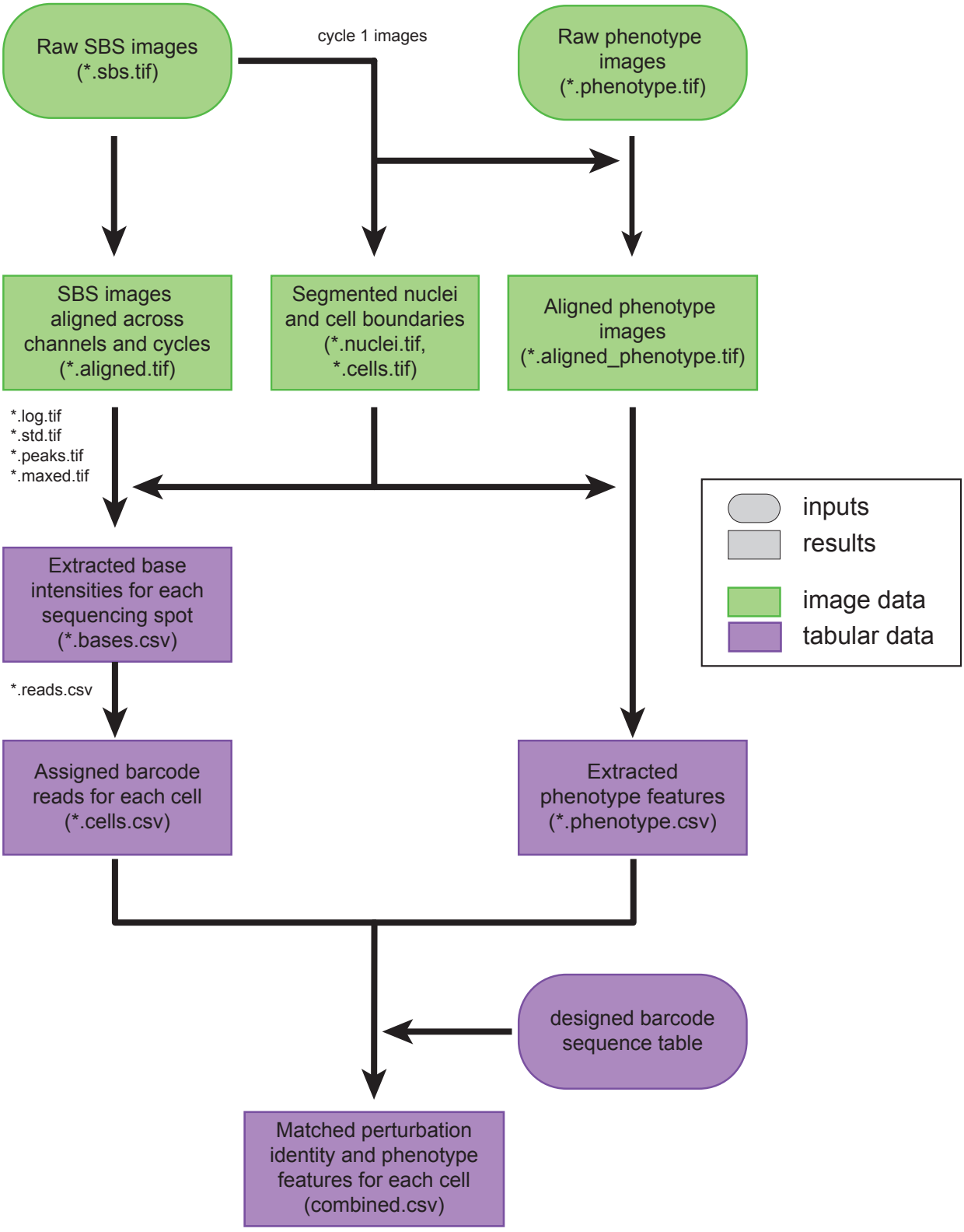
Cell line	Cells with ≥1 counts for most abundant mapped read*	Cells with ≥2 counts for most abundant mapped read*	NFkB screen published	Notes	Recommended
HeLa	94%	89.0%	yes		yes
A549	-	-	yes		yes
A375	92.8%	86.3%			yes
HT1080	91.8%	91.0%			yes
U2-OS	91.8%	84.7%		large surface area reduces cell throughput	yes
HCT116	95.5%	91.6%	yes	cells tend to clump	only if required by phenotyping assay
HEK293	93.2%	79.3%		high background staining, low sequencing spot count	only if required by phenotyping assay
THP-1	-	-		low sequencing spot count	no

*cell lines compared in the same experiment on a single plate,
mean of 2 replicate wells with 4 cycles of sequencing, 10 sgRNAs in the CROPseq vector

b



Supplementary Figure 1. Cell line comparison. (a) Suitability of example cell lines for optical pooled screening. NFkB screens were published in ref.[43] (b) Example cell and nuclear segmentation using CellPose, a generalist algorithm based on a deep learning model trained on diverse cell and image types. For each cell line, the left panel shows the exact RGB input provided to CellPose, and the right panel shows segmentations using the “cyto” and “nuclei” models in single-model mode (scale bar, 100 μm). All cell lines were analyzed identically, varying only in a single “diameter” parameter automatically generated by the CellPose “cyto” calibration model. Segmentation of a 1.2mm x 1.2mm field of view took ~4 min on 1 CPU.



Supplementary Figure 2. Image analysis pipeline overview. Flowchart demonstrating the major steps of the screen processing pipeline. Input data are represented by rounded shapes; image and tabular data are in green and purple, respectively.