

**Figure 1: The SMA CellProfiler pipeline workflow.** *A.* The original image (left) is masked and normalized using ExpressHist and slideNormalize in the slideToolKit workflow and used as input by CellProfiler; the graph (right) shows the tonal distribution in the digital whole-slide image on a RGB scale. *B.* The input image is converted to a gray scaled image (left); the graph (right) shows the tonal distribution in the gray scaled image. *C.* A Gaussian filter is applied to smoothen the image and reduce image artefacts (artifact size 20 pixels) and noise (left); the graph (right) shows the tonal distribution after smoothening. *D.* The gray scaled image is inverted, *i.e.* non-tissue will become black (left); the graph (right) shows the tonal distribution after inverting. *E.* The colors, *i.e.* stains, are unmixed using the original image (left): 3, 3'-diaminobenzidine (DAB, middle), and Hematoxylin (HE, right). *F.* The tissue area is identified, as demarcated by the green line in the left image, the DAB-stained nuclei are identified (middle), and the HE-stained nuclei are identified (right). Areas that are excluded due to size (minimal size 8 pixels) are demarcated (not shown); the middle and right image shows all the identified DAB- or HE-positive objects in random colors. *G.* Finally the data for each tile are saved in a comma-separated table, including meta-data such as tile positions, image location, object counts (there could be multiple patches of stained areas or tissue). The original image (left) is used to overlay the DAB- and HE-positive objects (nuclei are demarcated with yellow and red) and the tissue area (demarcated with green). Sample used: AE4780.SMA.ndpi.