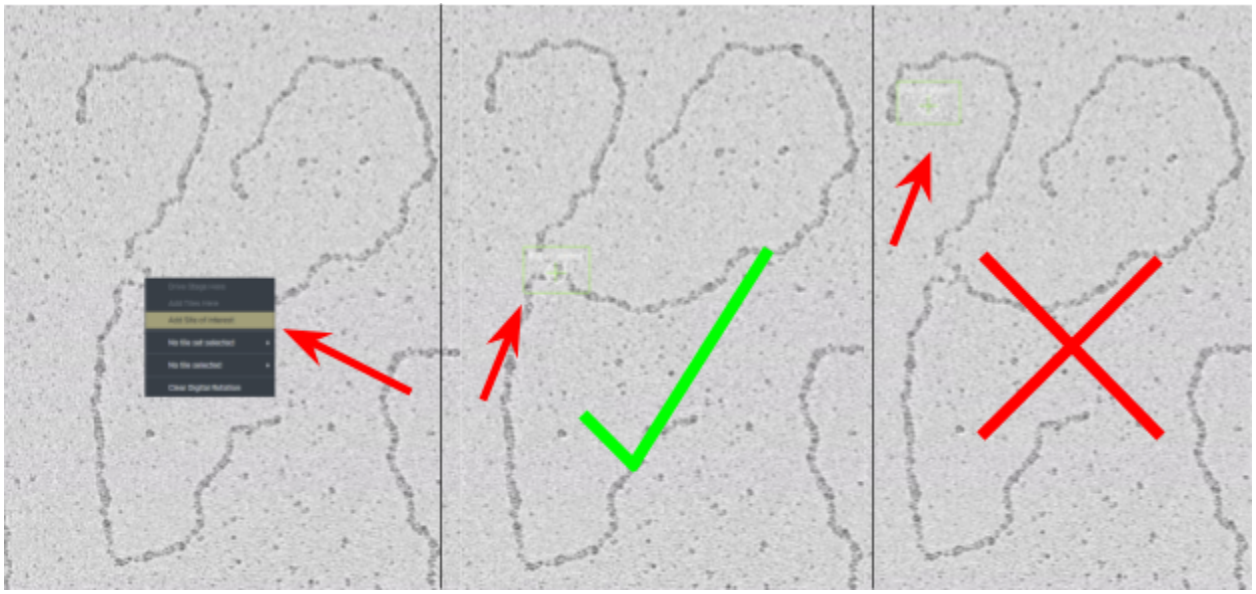
Talos Manual Annotation Workflow

This documentation describes the workflow to detect DNA replication forks from Talos EM images, stitch them and perform the measurements on them.

Annotation instructions

We use the Thermo Fisher MAPS software to annotate potential forks in Talos images.
After the acquisition of the images, the data is copied to a share and then the MAPS viewer is accessible via a ZMB virtual machine.

Within the MAPS software, you can open an experiment, search for forks and right click on a potential fork and select “Add Site of Interest”. Add this site to the center of the fork, not somewhere along the molecule. You can label the potential forks in any way you want, e.g. just accepting the default names or giving them names like “Unsure”, “Reversed Fork”, “Symmetric Fork” etc. The names you give the Sites of Interest will be the names of the stitched images afterwards. Allowed characters are letters, numbers and parentheses (because these names will be filenames of the stitched images).



Documentation by Joel Lüthi, 18/06/2019

Stitching images

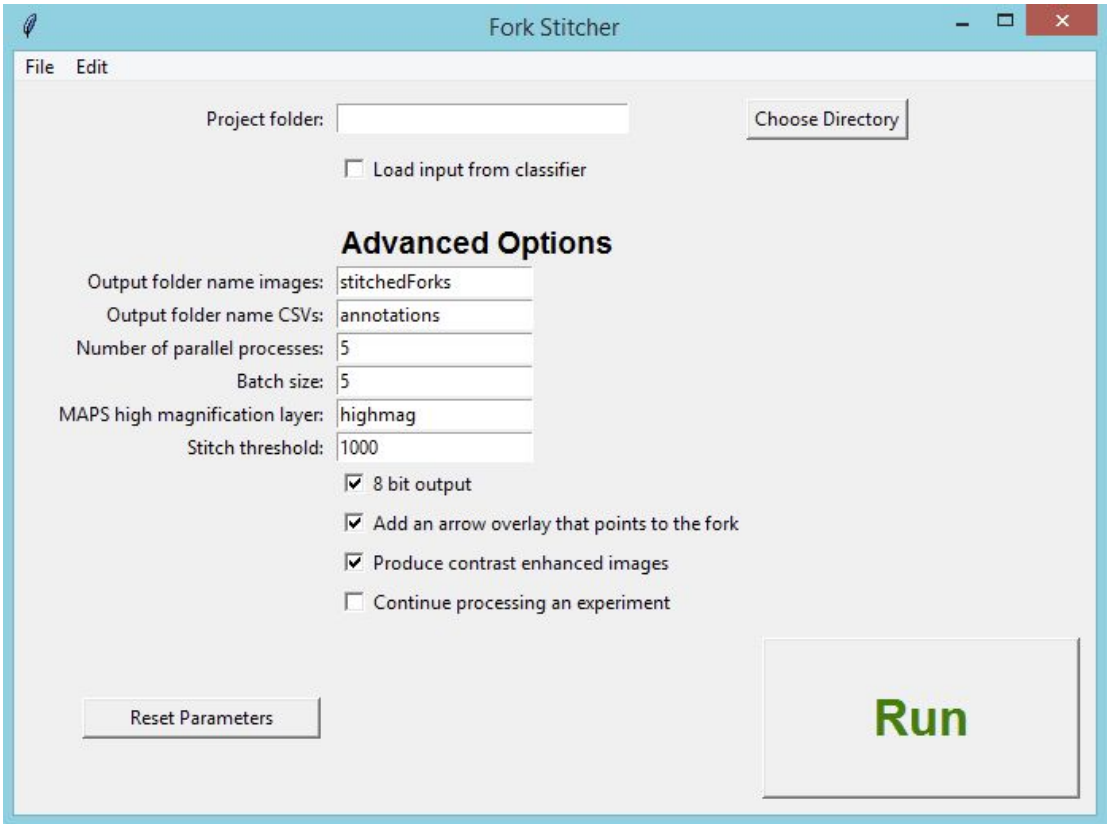
Starting the Stitcher

Make a Shortcut to the Stitcher.exe program and place the shortcut on your desktop. Then double click the shortcut.
Do not move the Stitcher.exe program itself. It can be found e.g. on the ZMB shares in GroupLopes/Stitcher.

If you want to change something in the program or test changes, you can run the program directly from the python source code. Follow the instructions here on how to install and run it: https://github.com/jluethi/ForkStitcher

Starting the Stitcher

Select the Project Folder: The folder that contains the MapsProject.xml file and the LayerData folder.
Change any of the options you want to adapt (see Working with the Stitcher below for details)
Click “Run”.



Documentation by Joel Lüthi, 18/06/2019

Working with the Stitcher

Project folder: The folder that contains the MapsProject.xml file and the LayerData folder.

Load input from classifier: Option to load the annotations from a classifier output csv file instead of from the MAPS project directly. By default, it's not active and the annotations from MAPS are loaded.

Classifier CSV file: If Load input from classifier is selected, this option appears. Input the full path to the csv file that the classifier created as output using the file picker button. You need to select the csv file, not just the folder where the file is in.

Advanced Options:

Most of the time, you will not have to change these options. Here is what they do:

Output folder name images: The name of the folder where the stitched images should be placed. The Stitcher creates a folder with this name in the *Project folder*.

Output folder name CSVs: The Stitcher creates a series of csv files in the *"Output folder name CSVs"* folder. It then goes through those files in batches (e.g. 5 at a time) and processed the annotations saved in them. After finishing a batch, it deletes the csv file and creates a new one that ends in _stitched.csv with additional information. In the end, the Excel Sheet containing information about all stitched images will be placed in this folder.

Number of parallel processes: How many stitching processed should be run in parallel. Depends on the CPU power and RAM available on your machine. On Image Processing VMs, 3 should work stable. 4 may work as well. If you tried 20, you would probably crash the Stitcher or making it incredibly slow because it runs out of memory (RAM).

Batch size: How many annotations are processed in each batch. Keeping this number low is useful if the Stitcher would crash or if you want to modify the batches being processed (see Continue Processing an Experiment).

High magnification layer: The name of the Layer within MAPS that contains the high magnification images (e.g. not the overview images). The Stitcher will stitch the images from that layer. Defaults to 'highmag' in the current Talos settings of the Lopes lab.

Stitch threshold: Sometimes the Stitcher cannot find reasonable stitching parameters. This threshold allows the user to define a limit. If an image would be moved more than this threshold, the Stitcher will not perform a stitching but instead just provide the original center image. If this happens too often, increase this threshold. If you have weirdly stitched images and would prefer to just get the center image, decrease this threshold. 1000 - 2000 is a reasonable default.

Documentation by Joel Lüthi, 18/06/2019

8-bit output: Whether the image should be saved as an 8-bit image (or a 16-bit image). Given the low dynamic range in the image, saving as 8-bit is a good way to save 50% of the image size.

Add an arrow overlay that points to the fork: Whether the Stitcher should add an arrow overlay pointing to the annotation.

Produce contrast enhanced images: Whether the Stitcher should apply Normalize Local Contrast ImageJ filter to the images to enhance the contrast.

Continue processing an experiment: You can continue processing an experiment that has been stopped by clicking the *"Continue Processing an Experiment"* folder. If you click that button, the Stitcher does not parse the MAPS XML file and does not create new csv files in the *"CSV folder name"* folder. Instead, it processes the remaining csv files in the folder that have not been processed yet.
Continuing to process an experiment has 3 use-cases:
1. You need to stop processing, because your VM time has run out or you want to shut down your computer.
2. The Stitcher crashed in the middle of processing and you want it to continue without redoing everything it already did before.
3. You actually don't want to process all annotations in the MAPS XML file. In this case, you can start the Stitcher and shut it down when it starts to stitch the first images. Then, you delete/edit the csv files that contain annotations you are not interested in and then rerun the Stitcher using *"Continue Processing an Experiment".*

Stopping the Stitcher: You can stop the Stitcher by closing the interface and confirming the shut down.

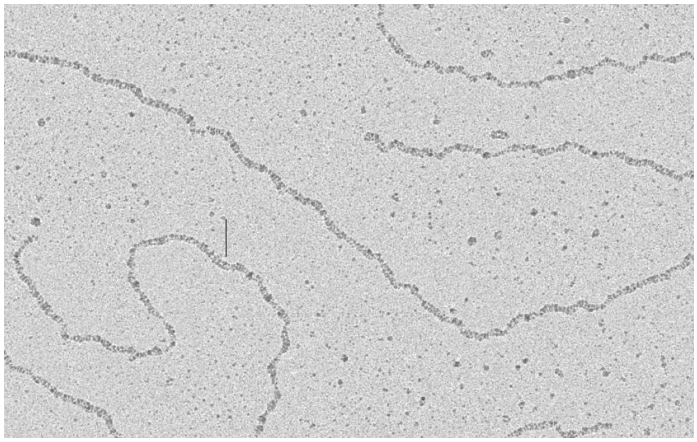
Analyzing Forks

The stitching tool puts all the stitched forks in the *Stitched images folder name* folder, which is in the project folder (e.g. Group Lopes\Sebastian\Projects\8384_1_siRad51C_CPT_3rd). It also puts an excel file that is prepared for the analysis of the forks in the *CSV folder name* folder. Look for the file that ends in "_fused.xlsx".
Open the forks in ImageJ. They are typically 130 Mb, so best do this from a VM where the connection is very fast. Check whether something is actually a fork or some kind of false positive (see scoring instructions below) and enter it accordingly into the spreadsheet. Do not modify the prefilled columns, they are needed to find the forks as an input to the Fork classifier. There is a green arrow pointing towards the potential fork. You can turn this arrow off by going to Image;Overlay;Hide Overlay and use Show Overlay to turn it back on again.
Also, perform any measurements you want to do on the forks and enter them in the excel sheet as well.

Scoring instructions

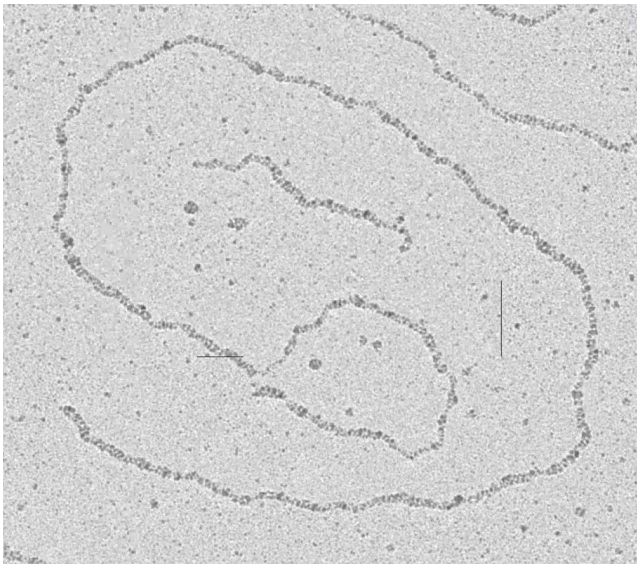
In order to incorporate these images into the training data of the neural network, we require the following classifications. The neural network only looks at the center of the fork, thus we can train it not to detect linear DNA but likely can't find a way to avoid detecting loops or crossings. Therefore, it is important to know whether something is not a fork and just linear DNA or whether something is not a fork, but on a very high zoom level may reasonably look like a fork (e.g. a loop).
When you analyze the stitched images of potential forks, put those classifications into the spreadsheet (csv file) that is automatically generated and put in the project folder.

Linear DNA

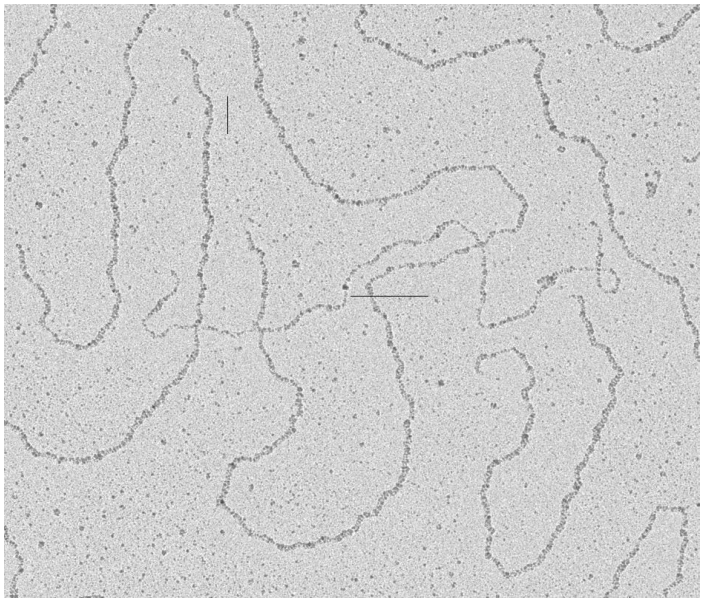


Documentation by Joel Lüthi, 18/06/2019

Loop



Crossing



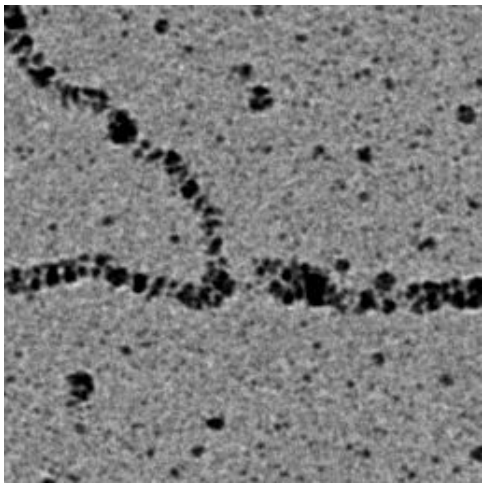
Other False Positive

Anything that is not a fork, but doesn't fit in any of the categories above

Documentation by Joel Lüthi, 18/06/2019

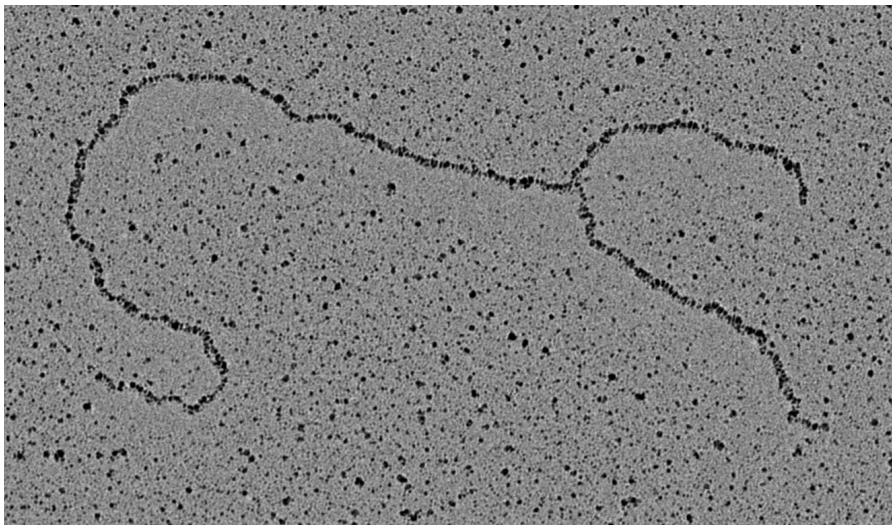
Missing Link Forks

Something that looks like a fork but is not fully connected



Fork

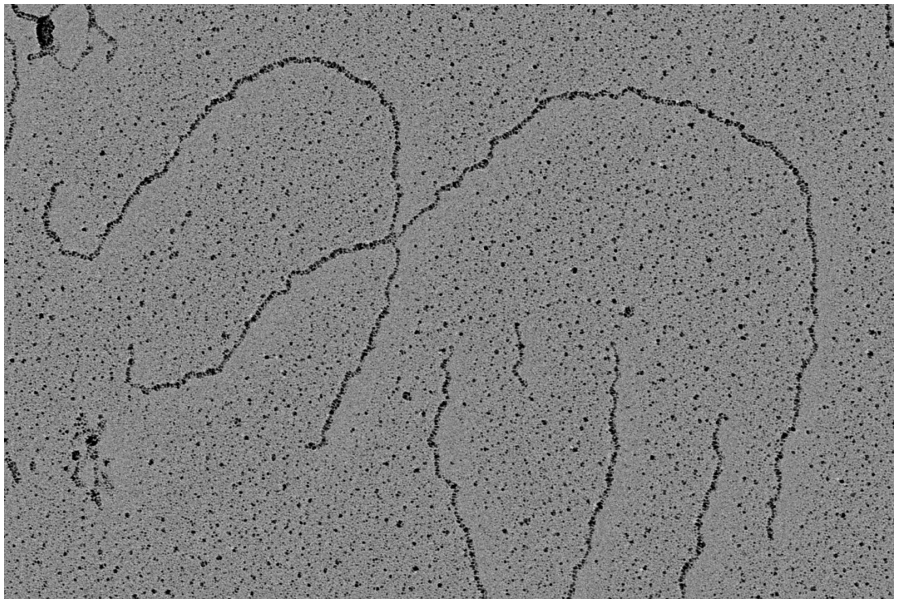
Symmetric or Asymmetric



Documentation by Joel Lüthi, 18/06/2019

Reversed Fork

Symmetric or Asymmetric



Documentation by Joel Lüthi, 18/06/2019