nature portfolio

https://doi.org/10.1038/s41596-020-00478-x

Supplementary information

Tracing DNA paths and RNA profiles in cultured cells and tissues with ORCA

In the format provided by the
authors and unedited

Supplementary Data Legends

Supplementary Data 1 | Primary probe synthesis costs. An EXCEL file that contains the information of all the components - including vendor, product number, unit, etc - needed for primary probe synthesis. Importantly, we have included the total cost to make a single primary probe.

Supplementary Data 2 | Cost of readout probe hybridization. An EXCEL file listing the reagents needed for each hybridization round. The user can edit this file to calculate the total cost per each experiment. The cost depends upon the number of readout probes, fluorescent oligos, and buffer used. The user can choose which fluorescent oligos to use; however, we advise that these oligos should be imaged in a different channel from the fiducial.
As an example, we have chosen to do a three-color experiment that has a total of 60 unique barcodes. This means that 30 readout probes (e.g. Readout_001: GGCCTCGAACGAACGATAGCcgaatgctct**GATCCGATTGGAACCGTCCCA**) will bind to a Cy5 fluorescent oligo (Cy5/**TGGGACGGTTCCAATCGGATC**)- due to its complementary sequences (shown in bold) - and a separate set of 30 readout probes (e.g. Readout_031: AGTGCGCACGAGTTGAACTGtttgctcgca**CGGTTGGAGGTGCAGGTGAG**) will bind to an Alexa 750 fluorescent oligo (Alex750/**CTCACCTGCACCTCCAACCG**). The fiducial sequences will bind to an Alexa 488 fluorescent oligo. All these sequences are provided under **Data Availability**. Thus, there should be no fluorescent cross-talk between these three channels. Also, doing a three-color experiment saves time as the length of the experiment is reduced (see **Supplementary Data 3**) and other fluidics components. However, if the user wants to do a two-color experiment, then they should edit the top of the table under the fluorescent oligo options (you can change the 'bulk-Alexa 750' from 1 to 0, if you want to image readout data only in Cy5). Additionally, bulk-Cy5 and bulk-Alexa750 simply means that the user decided to buy the fluorescent oligo(s) as a bulk instead of a single order. Buying in bulk saves money.

Supplementary Data 3 | Length of imaging experiment and ChrTracer3 analysis. An editable EXCEL file that is useful to estimate the amount of time it will take to complete an imaging experiment and analyze the raw data in ChrTracer3. We have listed all the parameters that would influence the length of time for both of these. Please note that in 'Imaging Time' we have listed the 'Number of readouts per hybridization' to be '2', which means we are showing an example of a three-color experiment Details of a three-color experiment can be found in the legend of **Supplementary Data 2**.

Supplementary Data 4 | Raw data storage. An editable EXCEL file to help the user calculate the amount of storage space needed for an imaging experiment. This will also calculate the cost to store the data.

Supplementary Data 5 | Components for our homebuilt microscope. An EXCEL file containing detailed information - including the supplier(s), part number(s), quantity, units, and total costs - of each component used for our microscope. The total cost of the homebuilt microscope is also included in this table.

Supplementary Data 6 | RNA Table. An example of a RNAtable.xlsx file that is used for downstream analysis, such as in BuildMosaicsGUI.

The first column ‘FolderName’ should match the name of the folder within the RNA_Expt folder. The ‘Readout’ column lists the readout numbers used in the experiment. The ‘DataType’ column has either ‘H’ or ‘R’, which stands for a readout Hybridization or a readout Repeat (in our example, there is no ‘R’). The ‘HybNum’ column represents the order in which the readouts were imaged during the experiment. The ‘channels’ column lists the channels that were used to image the readout and the fiducial data. In this case, we are using the 647-nm and 561-nm laser to record the readout and fiducial data, respectively. Specify which channel was used for the fiducial imaging under the ‘fiducialChannel’ column. The ‘channelLayout’ column is designated as ‘alternate’ to indicate that the channels were sequentially imaged. The column ‘bufferFrames’ is ‘10’, which means that there were 10 frames imaged in the beginning and end of the *z*-scan for a FOV that were still and not moving in either *z* direction. ‘totalFrames’ is ‘220’ and this indicates that there were 220 frames imaged in total for both channels. The last column, ‘rnaNames’ lists out the name of each RNA species that was imaged.

Supplementary Data 7 | DNA Table. An example of ExperimentalLayout.xlsx that is used for ChrTracer3 and downstream GUIs. The columns are the same as that of RNAtable.xlsx (see legend for **Supplementary Data 6**), with two exceptions. 1) For DNA labeling, we typically relabel some of the barcodes at the experiment for quantification purposes. Thus, you will see that at the bottom of the ‘DataType’ column, the last 5 rows are labelled as ‘R’, indicating that these are the barcodes that were relabelled. 2) There is no ‘rnaNames’ column as this is a DNA table.

Supplementary Data 8 | ChrTracer3 output file. An example of a fov001_AllFits.csv table that is generated at the end of ChrTracer3 analysis.
The first three columns ‘x’, ‘y’, and ‘z’ indicate the 3D position of a barcode in nm relative to the current crop box (add the value in locusX and locusY to the ‘x’ ‘y’ values to translate into stage coordinates in nm); ‘h’ is the brightness of the molecule, ‘wx’, ‘wy’, and ‘wz’ are the standard-deviations in each dimension associated with the Gaussian fit (in nm), a is the magnitude of the fit and b is the background; ‘xL’ ‘xU’, ‘yL’, ‘yU’, ‘zL’, ‘zU’ are the lower and upper uncertainty bounds on the x,y,z positions computed; ‘readout’ is readout name, ‘dataType’ is either ‘R’, ‘A’, or ‘H’ indicating if the spot detected was a repeat-control, an alignment-control for chromatic correction, or a normal data hybridization respectively; ‘chn’ records the laser used to image the spot, ‘fov’ the field of view in which it was recorded, ‘locusX’ and ‘locusY’ record the x,y coordinates of the spot in field of view, ‘tableOrder’ records the position the data set was loaded and is used for troubleshooting, ‘hybe’ records which labeling round (hybridization) the spot was collected in; ‘chnNum’ records for multicolor acquisition, which order the spot was acquired in; ‘idx’ is an index that records the number of peaks found in the spot image (typically restricted to be only the brightest); ‘panel’ records the order in which the spot was displayed in the corresponding troubleshooting panels, ‘s’ is the the index of the current spot with respect to the current FOV, ‘fs’ is a Cantor-Pair unique number per chromosome created from ‘s’ and ‘fov’; ‘fsr’ is a Cantor-Pair of readout and fs, a generating a unique number for every spot in the experiment; ‘xshift’, ‘yshift’, ‘zshift’ are the translations, in nm, applied to correct drift for the spot; ‘shiftScore’ is a binary value of whether the drift correction was considered reliable or uncertain, fid_x, fid_y, and fid_z are the centroid coordinates of the fiducial relative to the crop box, fid_h is the brightness of the fiducial signal.