BREC

Inputs
genetic and physical maps from built-in dataset or imported files

Step 0 - Data pre-processing

- DQC Test data quality → Low quality / Good quality
- Apply data cleaning? → No / Yes → Clean data

Main process
From a Marey map
to identified heterochromatin regions,
centromere and telomere(s),
to extrapolated local recombination rate estimates

Outputs
- Recombination rate estimates
- Chromatin boundaries
- Statistics
- Interactive plots
- Visualize and download

Step 1 - Estimate Marey Map-based local recombination rates

- Correlate physical and genetic distances to build the Marey map
- Generate two regression models: third degree polynomial and Loess
- Compute the prime derivative for both models and return them as preliminary recombination maps

Step 2 - Identify the chromosome type

- Retrieve the smallest value of recombination rate from both maps
- Test whether the two corresponding markers have the same physical position or not
- Deduce the type (and subtype if applicable) of the chromosome
- Return the result if the decision is possible, otherwise, ask for the user's input

Step 3 - Prepare the heterochromatin boundaries identification

- Apply R-squared to test the goodness-of-fit and accumulate its values twice: forwards and backwards
- Compute chromosome-specific sliding window whose size is the largest inter-marker physical distance
- Apply the sliding window to test the slope variation on each accumulative R-squared curve
- If **telocentric**, deduce the position of the centromere on the arm (left or right).
- If **atelocentric (metacentric)**, test whether a centromeric gap exists or not

Step 4 - Identify the centromeric boundaries

- Set the smallest value of polynomial-based recombination rate as the starting point for the sliding window
- **Telocentric**: move the sliding window far from the centromere side while keeping the starting point as the beginning of the centromeric region
- **Atelocentric**: use two sliding windows on both R-squared curves. Move them in opposite directions to expand the starting point into a centromeric region
- When each sliding window stops, get the physical position of its ending edge and return it as the left (or right) centromeric boundary

Step 5 - Identify the telomeric boundaries

- Detect telomeric boundaries identified by the most significant depletion on the R-squared curve
- **Telocentric**: return one physical position of the corresponding depletion on the telomere side
- **Atelocentric**: return two physical positions of both corresponding depletions of the two telomeres

Step 6 - Extrapolate the local recombination rates and display the outputs

- Retrieve physical ranges of the identified heterochromatin regions
- Reset to zero the values of the local recombination rate estimates along these regions
- Return the Loess-based recombination map as BREC's result
- Generate the interactive plot that includes all BREC outputs