README file for the master equation solution for the methylation profile and associated analysis and plotting code

Sarah Sandholtz and Quinn MacPherson

4/17/20

This folder contains the MATLAB code used in [1] for calculating the master equation (ME) solution for methyl speading along a chromosome based on the chromatin structure. A complementary Monte Carlo (MC) simulation (in Fortran) predicts the chromatin configuration from a fixed methylation sequence. Also included in this folder is code used in [1] for identifying neighboring nucleosomes, calculating the correlation between the initial and final methylation sequences, and calculating and plotting the dynamics of methylation reestablishment. Basic descriptions of these scripts are below. Raw data for the chromatin configurations from the MC simulations in [1] can be found in the subfolder MCsimdata, and raw data for the methylation sequences in [1] can be found in the subfolder MEdata.

1. S. H. Sandholtz, Q. J. MacPherson, and A. J. Spakowitz. "Physical Modeling of the Heritability and Maintenance of Epigenetic Modifications," *PNAS*, submitted (2020).

In order to run this set of simulation, calculation, and analysis code, you will need the ability to run code written in Fortran and MATLAB. Code and documentation for the MC simulations can be found in the SpakowitzLab GitHub repository in the MBS2018_PNAS branch (https://github.com/SpakowitzLab/wlcsim/tree/MBS2018_PNAS).
Below is a brief description of the analysis and plotting scripts used in [1].

saveNeighbors.f90

This code is used to find neighboring nucleosomes. It takes as input the configuration file from an MC simulation and gives as output a list of pairs of proximal nucleosomes. Specify "NT," the number of nucleosomes in the MC simulation. Specify "radius," the distance within which two nucleosomes will be considered neighbors. This number is in units of the discretization of the MC simulations, in which 1 unit corresponds to 28.7 nm.

compileSaveNei.sh

This code compiles the Fortran code "saveNeighbors.f90."

binning.f90

This code is called in the "saveNeighbors.f90" script. It establishes the bins used in the MC simulation, which are relevant for identifying neighboring nucleosomes.

"masteqn_avghp1.m"

This MATLAB code calculates the average number of neighboring nucleosomes and solves the master equation for the methylation profile.

"calc.m"

This MATLAB code calculates the overall fraction of methylated nucleosomes, the window-averaged methylation at each nucleosome, and the correlation between the initial and current methylation sequences.

"plot_dynamics.m"

This MATLAB code calculates and plots the average fraction reamplification (grouped by nucleosomes in euchromatin, on the boundary, and in heterochromatin) as a function of time.