The root directory of SCL package contains three sub folders, “scripts” that contains all the programs, “data” that contains the data used, and “outputs” that contains all the output files.

1. Generate the 3D structure

The C++ program “scl.cpp” can read the Hi-C matrix and generate the 3D structure that are represented by 3D coordinates.

The following is an example:

“cd SCL1.0\_source\_code/”;

“cd scripts/”;

“g++ -o scl scl.cpp”;

“./scl -i ../data/Th1\_Cell1\_chrX.txt -o ../outputs/Th1\_cell1/chrX/50k/Th1\_Cell1\_chrX\_model -res 0.05”;

“-i” specifies the input single-cell Hi-C contact file, in which each line contains two locations for being in contact. The two locations indicate one single-cell Hi-C contact.

“-res” specifies the resolution of the inferred 3D structure. The unit of it is Mb. For example, 0.05 means the inferred 3D structure will be at 50 kb resolution.

“-o” specifies the output file that contains 3D coordinates of the inferred 3D structure.

The user can freely adjust all the parameters shown in Eq. (1)-(5) in our paper. If a parameter is not specified, a default value will be used. See below for details:

“-t” specifies the starting temperature for simulated annealing. If it is not specified, default value 10 will be used.

“-theta1” specifies theta1 in Eq. (1). If it is not specified, default value 0.7 will be used.

“-delta0” specifies delta0 in Eq. (1). If it is not specified, default value 8 will be used.

“-rho” specifies rho in Eq. (1). If it is not specified, default value 1 will be used.

“-phi” specifies phi in Eq. (1). If it is not specified, default value 0.1 will be used.

“-tau” specifies tau in Eq. (1). If it is not specified, default value 1 will be used.

“-beta” specifies beta in Eq. (1). If it is not specified, default value 1 will be used.

“-mu1” specifies mu1 in Eq. (1). If it is not specified, default value 20 will be used.

“-mu2” specifies mu2 in Eq. (3). If it is not specified, default value 2 will be used.

“-d0” specifies d0 in Eq. (3). If it is not specified, default value 1000 will be used. If the value of d0 is bigger than the length of the chromosome, i.e., the number of beads for the entire chromosome, it’s value will be changed to the total number of beads for the target chromosome.

“-d1” specifies d1 in eq. (5). If it is not specified, default value 8 will be used.

“-cost” is a parameter to control whether the cost value at each temperature will be displayed in screen. Input 1 will let the program output the cost value; 0 will not. Printing the cost value at each temperature will increase the computing time.

“-e” is the expansion index. Before adding high-resolution beads between every two consecutive low-resolution beads, we multiple the X, Y, and Z coordinates by a constant to enlarge the low-resolution structure so that there are plenty of spaces for newly-added high-resolution beads. By default, this value is set to 0.12 but can be changed by this parameter. The program may give warning message saying there are not enough spaces for some locations. This usually is fine, i.e., no need to enlarge the value of this expansion index.

“-t2” is the temperature used for the fine-tuning simulation for constructing the high-resolution structure. The default value of this parameter is 0.1.

“-times” specifies the times of tries performed for the fine-tuning simulation. A number 5 means the number of tries that will be performed will be 5 times the length of the chromosome, which is also the total number of beads. If it is not specified, default value 5 will be used.

2. Visualization of the inferred structures

You can convert the coordinate file generated by scl.cpp to “.pdb” file for visualization by the software such as Pymol or “VMD.

Example for converting a single pdb file:

“cd SCL1.0\_source\_code/scripts/visualize/”;

“perl convert\_single.pl ../../outputs/ES\_cell/chr1/500k/chr1\_500k\_model\_1.txt ../../outputs/ES\_cell/chr1/500k\_model/chr1\_500k\_model\_1.pdb 20”;

where “../../outputs/ES\_cell/chr1/500k/chr1\_500k\_model\_1.” is the coordinates file generated by scl program, ../../outputs/ES\_cell/chr1/500k\_model/chr1\_500k\_model\_1.pdb” is the output pdb file, and “20” is the maximum coordinate value. We recommend to use 20 as the maximum coordinate file based on our evaluations.

1. Select the top inferred structure from a pool of inferred structures

The user can run multiple jobs that generate multiple 3D structures and then select the top structure. We provide a program to rank the models based on Q-score.

First, use “convert\_folder.pl” to convert all the coordinate files in a folder into pdb format files, and then use “qscore.pl” to select top 1 or 3 structures among the pdb files.

Before using “qscore.pl”, it should be confirmed that “TMscore” file in the same folder is executable. If “TMscore” file is not executable (using “ls -l” to check its permission), use “chmod 755 TMscore” to give it permission for executing.

Example of executing the programs:

“cd SCL1.0\_source\_code/scripts/visualize/”;

“perl convert\_folder.pl ../../outputs/ES\_cell/chr1/500k/ ../../outputs/ES\_cell/chr1/500k\_model/ 20”;

“perl qscore.pl ../../outputs/ES\_cell/chr1/500k\_model/ ../../outputs/ES\_cell/chr1/rank\_list.txt”;

“../../outputs/ES\_cell/chr1/500k/” is the folder contains the coordinate files, “../../outputs/ES\_cell/chr1/500k\_model/” is the folder to store the output pdb files, and “20” is the maximum coordinate value. The “../../outputs/ES\_cell/chr1/rank\_list.txt” is the output ranking list of predicted 3D structures based on their Q-score.