

Chromosome replication

Model implementation. Chromosome replication occurs through three steps that are implemented in the `ChromosomeFormation` and `ChromosomeElongation` processes. First, a round of replication is initiated at a fixed cell mass per origin of replication and generally occurs once per cell cycle (see Algorithm 1). This is in contrast to the DnaA based mechanistic model included in the *M. genitalium* model but allows for stable replication over multiple generations and in different growth conditions. Second, replication forks are elongated up to the maximal expected elongation rate, dNTP resource limitations, and template strand sequence but elongation does not take into account the action of topoisomerases or the enzymes in the replisome (see Algorithm 2). Finally, replication forks terminate once they reach the end of their template strand and the chromosome immediately decatenates forming two separate chromosome molecules (see Algorithm 3).

Algorithm 1: DNA replication initiation

Input : m_{cell} cell mass

Input : $m_{critical}$ critical initiation mass

Input : n_{origin} number of origins of replication

Input : $n_{fork,f}$ number of replication forks on forward strand

Input : $n_{fork,r}$ number of replication forks on reverse strand

Input : C length of C period

Input : D length of D period

if $\frac{m_{cell}}{n_{origin}} > m_{critical}$ **then**

$n_{fork,f} = n_{fork,f} + n_{origin}$

$n_{fork,r} = n_{fork,r} + n_{origin}$

$n_{origin} = 2 \cdot n_{origin}$

Result: When cell mass is larger than critical initiation mass m_c another round of replication is initiated with correct number of replication forks

Algorithm 2: DNA replication elongation

Input : e maximal elongation rate of replication fork
Input : p_i position of forks on chromosome where $i = 1$ to n_{fork}
Input : Δt length of current time step
Input : $c_{dNTP,j}$ counts of dNTP where $j = 1$ to 4 for dCTP, dGTP, dATP, dTTP
Input : L_k total length of each strand of chromosome from origin to terminus
 where $k = 1$ to 4 for forward/complement and reverse/complement.
for each replication fork i on sequence k **do**
 1. Based on replication fork position p_i and maximal elongation rate e
 determine “stop condition” (s_i) for replication fork assuming no dNTP
 limitation.
 $s_i = \min(p_i + e \cdot \Delta t, L_k)$
 Stop condition is either maximal elongation rate scaled by the time step or
 the full length of sequence (i.e. the fork will terminate in this time step).
 2. Derive sequence between replication fork position (p_i) and stop condition
 (s_i).
 3. Based on derived sequence calculate the number of dNTPs required to
 polymerize sequence $c_{dNTP,i}^{req}$.
 4. Elongate up to limits:
 if all($c_{dNTP,i}^{req} < c_{dNTP,j}$) **then**
 | Update the position of each replication fork to stop position
 | $p_i = s_i$
 else
 | Attempt to equally elongate each replication fork update position of each
 | fork to maximal position given the limitation of $c_{dNTP,j}$.
 5. Update counts of $c_{dNTP,j}$ to reflect polymerization usage.
Result: Each replication fork is elongated up to the limit of available sequence,
 elongation rate, or dNTP limitation

Algorithm 3: DNA replication termination

Input : p_i position of forks on chromosome where $i = 1$ **to** n_{fork}
Input : L_k total length of each strand of chromosome from origin to terminus
where $k = 1$ **to** 4 for forward/complement and reverse/complement
Input : d_{queue} a double ended queue data structure that stores time(s) cell
division should be triggered
Input : D D-period of cell cycle (time between completion of chromosome
replication and cell division)
Input : t Current simulation time
for each replication fork i on strand k do
 if $p_i == L_k$ then
 1. Delete replication fork
 2. Divide remaining replication forks and origins of replication
 appropriately across the two new chromosome molecules
 3. Calculate time cell should trigger division based on current time of
 chromosome termination and push onto queue data structure
 $d_{queue}.push(t + D)$
Result: Replication forks that have terminated are removed. A new chromosome
molecule is created separating all remaining replication forks. Timer for
D-period is started.

Associated data

Parameter	Symbol	Units	Value	Reference
Chromosome sequence	-	-	-	[1]
Replication fork elongation rate	e	nt/s	967	[2]
Mass per origin at DNA replication initiation ⁽¹⁾	$m_{critical}$	origin/fg	[600,975]	[3]
C period	C	min	40	[4]
D period	D	min	20	[4]

Table 1: Table of parameters for chromosome replication process.

⁽¹⁾600 is used for anaerobic conditions where the cell mass is lower and was fit to achieve the appropriate D period. All other growth conditions use 975.

Associated files

wcEcoli Path	File	Type
wcEcoli/models/ecoli/processes	chromosome_formation.py	process
wcEcoli/models/ecoli/processes	chromosome_replication.py	process
wcEcoli/reconstruction/ecoli/dataclasses/process	replication.py	data
wcEcoli/reconstruction/ecoli/flat	genes.tsv	raw data
wcEcoli/reconstruction/ecoli/flat	sequence.fasta	raw data

Table 2: Table of files for chromosome replication.

References

- [1] F R Blattner, G Plunkett, C A Bloch, N T Perna, V Burland, M Riley, J Collado-Vides, J D Glasner, C K Rode, G F Mayhew, J Gregor, N W Davis, H A Kirkpatrick, M A Goeden, D J Rose, B Mau, and Y Shao. The complete genome sequence of *Escherichia coli* K-12. *Science (New York, N.Y.)*, 277(5331):1453–1462, September 1997.
- [2] Hans Bremer and Patrick Dennis. Modulation of chemical composition and other parameters of the cell by growth rate. *Escherichia coli and Salmonella: cellular and molecular biology*, 2:1553–1569, 1996.
- [3] W D Donachie. Relationship between cell size and time of initiation of DNA replication. *The American Journal of Gastroenterology*, 219(5158):1077–1079, September 1968.
- [4] Frederick Carl Neidhardt, John L Ingraham, and Moselio Schaechter. *Physiology of the bacterial cell: a molecular approach*. Sinauer Associates Sunderland, MA, 1990.