RNA degradation

Model Implementation. The *E. coli* model provides a more detailed, mechanistic representation of RNA degradation compared to the *M. genitalium* model. Unlike the previous model, the gene functionality of endoRNase and exoRNase is mechanistically integrated to evaluate: (1) rates of RNA degradation due to endo-nucleolytic cleavage, and (2) rates of nucleotides digested by exoRNases. These mechanisms are implemented in the `RnaDegradation` process (detailed in Algorithm 1).

Endo-nucelolytic Cleavage. RNAs are cleaved by nine different endoRNases, each of which are assumed to have the same rate of cleavage but can have a different specificity for cleavage of mRNAs, tRNAs or rRNAs. The rate of cleavage for each RNA is determined with a Michaelis-Menten kinetic equation:

$$r_i = k_{cat,endo} \cdot c_{endo,i} \cdot f_i \tag{1}$$

where i indicates the RNA from each gene, r_i is the rate of cleavage of each RNA species, $k_{cat,endo}$ is the rate of cleavage for a single endoRNase, $c_{endo,i}$ is the count of endoRNases specific to each RNA species, and f_i is the saturation of endoRNases for each RNA and is defined:

$$f_i = \frac{\frac{c_{RNA,i}}{K_{M,i}}}{1 + \sum_j \frac{c_{RNA,j}}{K_{M,j}}} \tag{2}$$

where $c_{RNA,i}$ is the count of each RNA and $K_{M,i}$ is the Michaelis constant for each RNA, i. The saturation fraction accounts for competitive binding of each RNA species for the available endoRNases. The Michaelis constant for each RNA is determined by setting Eq. 1 equal to the first order approximation based on measured (or assumed, if measurement

data is not available) RNA half lives:

$$k_{cat,endo} \cdot c_{endo,i} \cdot \frac{\frac{c_{RNA,i}}{K_{M,i}}}{1 + \sum_j \frac{c_{RNA,j}}{K_{M,j}}} = \frac{\ln 2}{\tau_{RNA,i}} \cdot c_{RNA,i} \qquad (3)$$

where $\tau_{RNA,i}$ is the half life for each RNA. Expected counts of endoRNases and RNA transcripts for an average cell are used to solve the non-linear equation for each $K_{M,i}$.

During simulations, individual RNA counts are low so r_i from Eq. 1 would be very low ($\ll 1$) for most RNA. To get integer counts to degrade within a timestep, the total number of RNAs expected to be degraded is first determined. Then, samples are drawn from a multinomial distribution of available RNAs to degrade until this total is reached. This is done separately for each RNA group (mRNA, tRNA and rRNA) based on the known endoRNase affinity for each group:

$$R_{endo,group} = \sum_{i \in group} r_i \cdot \Delta t \qquad (4)$$

where r_i is as defined in Eq. 1, Δt is the length of the simulation timestep, and $i \in$ $group$ means that the RNA is in a specific RNA group (mRNA, tRNA or rRNA). The multinomial distribution for each RNA to degrade is sampled with a probability for each RNA species based on the saturated fraction of that species:

$$p_{deg,group,i} = \frac{f_i \cdot \mathbf{1}_{i \in group}}{\sum_{j \in group} f_j} \qquad (5)$$

where $\mathbf{1}_{i \in group}$ is an indicator function that is 1 if i in in $group$ and 0 otherwise.

Exo-nucleolytic Digestion. Endo-nucleolytic cleavage produces non-functional RNA fragments, which are then degraded to individual nucleotides via nine different exoRNases. Once degraded, these nucleotides can be recycled by the `Metabolism` process. ExoRNase capacity is determined as the following:

$$R_{exo} = k_{cat,exo} \cdot c_{exo} \cdot \Delta t \qquad (6)$$

2

where R_{exo} is the exoRNase digestion capacity, $k_{cat,exo}$ is the rate of cleavage for a single exoRNase, c_{exo} is the total count of exoRNases, and Δt is the length of the simulation timestep. Since non-functional RNA fragments are modeled in aggregate and not individually like functional RNA molecules, we do not include a K_M term for the digestion capacity and assume full saturation.

Algorithm 1: RNA degradation: endo-nucelolytic cleavage and exo-nucleolytic digestion

---

Input : $K_{M,i}$ Michaelis constants of each mRNA transcript binding to endoRNases where $i = 1$ **to** n_{RNA}

Input : $k_{cat,endo}$, $k_{cat,exo}$ catalytic rate of endoRNase and exoRNase

Input : c_{endo}, c_{exo} count of endoRNases and exoRNases

Input : $c_{frag,i}$ count of nucleotides in non-functional RNA fragments where $i = 1$ **to** 4 for AMP, CMP, GMP, UMP

Input : $c_{nt,i}$ count of free nucleotides where $i = 1$ **to** 4 for AMP, CMP, GMP, UMP

Input : c_{mRNA}, c_{tRNA}, c_{rRNA} count of each mRNA, tRNA and rRNA

Input : c_{H_2O}, c_{PPi}, c_{proton} count of small molecules

Input : `multinomial()` function that draws samples from a multinomial distribution

Input : `countNTs()` function that returns counts of AMP, CMP, GMP, and UMP for a given non-functional RNA fragment

Input : `lengthFragments()` function that returns the total number of bases of all RNA fragments

```
/* Endo-nucleolytic cleveage */
```

1. Calculate fraction of active endoRNases (f_i) that target each RNA where $i = 1$ **to** n_{gene}

$$f_i = \frac{\frac{c_{RNA,i}}{K_{M,i}}}{1 + \sum_j \frac{c_{RNA,j}}{K_{M,j}}}$$

2. Calculate total counts of RNAs to be degraded ($R_{endo,group}$)

$$R_{endo,mRNA} = \sum_{i \in mRNA} k_{cat,endo} \cdot c_{endo,i} \cdot f_i \cdot \Delta t$$
$$R_{endo,tRNA} = \sum_{i \in tRNA} k_{cat,endo} \cdot c_{endo,i} \cdot f_i \cdot \Delta t$$
$$R_{endo,rRNA} = \sum_{i \in rRNA} k_{cat,endo} \cdot c_{endo,i} \cdot f_i \cdot \Delta t$$

3. Determine probabilities for multinomial distributions of RNAs to degrade for each RNA group ($p_{deg,group,i}$)

$$p_{deg,mRNA,i} = \frac{f_i \cdot \mathbf{1}_{i \in mRNA}}{\sum_{j \in mRNA} f_j}$$
$$p_{deg,tRNA,i} = \frac{f_i \cdot \mathbf{1}_{i \in tRNA}}{\sum_{j \in tRNA} f_j}$$
$$p_{deg,rRNA,i} = \frac{f_i \cdot \mathbf{1}_{i \in rRNA}}{\sum_{j \in rRNA} f_j}$$

4. Sample multinomial distributions for each group $R_{endo,group}$ times, with probability determined by relative endoRNase saturation, to determine counts of RNAs that are converted into non-functional RNAs (d_i)

$$d_i = \texttt{multinomial}(R_{endo,mRNA}, p_{deg,mRNA,i})$$
$$+ \texttt{multinomial}(R_{endo,tRNA}, p_{deg,tRNA,i})$$
$$+ \texttt{multinomial}(R_{endo,rRNA}, p_{deg,rRNA,i})$$

---

5. Increase number of RNA fragments. Decrease RNA and water counts by amount required for RNA hydrolysis and increase pyrophosphate counts for the removal of the 5' pyrophosphate

$$c_{frag} = c_{frag} + \text{countNTs}(d_i)$$
$$c_{RNA,i} = c_{RNA,i} - d_i$$
$$c_{H_2O} = c_{H_2O} - \sum_i d_i$$
$$c_{PPi} = c_{PPi} + \sum_i d_i$$

```
/* Exo-nucleolytic digestion */
```
6. Compute exoRNase capacity (R_{exo})

$$R_{exo} = k_{cat,exo} \cdot c_{exo} \cdot \Delta t$$

if $R_{exo} > \sum c_{frag,i}$ **then**

Update nucleotide, water and proton counts

$$c_{nt,i} = c_{nt,i} + c_{frag,i}$$
$$c_{H_2O} = c_{H_2O} - \text{lengthFragments}(c_{frag})$$
$$c_{proton} = c_{proton} + \text{lengthFragments}(c_{frag})$$

Set counts of RNA fragments equal to zero $(c_{frag,i} = 0)$

else

Sample multinomial distribution c_{frag} with equal probability to determine which fragments are exo-digested $(c_{frag,dig})$ and recycled

$$c_{frag,dig,i} = \text{multinomial}(R_{exo}, \frac{c_{frag,i}}{\sum_i c_{frag,i}})$$

Update nucleotide, water, proton counts, and RNA fragments

$$c_{nt,i} = c_{nt,i} + c_{frag,dig,i}$$
$$c_{H_2O} = c_{H_2O} - \text{lengthFragments}(c_{frag,dig})$$
$$c_{proton} = c_{proton} + \text{lengthFragments}(c_{frag,dig})$$
$$c_{frag,i} = c_{frag,i} - c_{frag,dig,i}$$

Result: RNAs are selected and degraded by endoRNases, and non-functional RNA fragments are digested through exoRNases. During the process water is consumed, and nucleotides, pyrophosphate and protons are released.

Associated data

Parameter	Symbol	Units	Value	Reference
EndoRNase catalytic rate	$k_{cat,endo}$	RNA counts/s	0.10	See GitHub
ExoRNase catalytic rate	$k_{cat,exo}$	nt digested/s	50	See GitHub
mRNA half-lives[1]	τ_{mRNA}	min	[1.30, 31.40]	[1]
tRNA, rRNA half-lives	τ_{tRNA}, τ_{rRNA}	hour	48	[1]
Michaelis constant	K_M	RNA counts	-	See GitHub
RNAse mechanism of action	-	-	endo-/exo-RNAse	See GitHub
EndoRNase specificity[2]	-	-	Specificity for mRNA, tRNA, rRNA	See GitHub

Table 1: Table of parameters for RNA degradation process.
[1]Non-measured mRNA half-lives were estimated as the mean mRNA half-life (5.75 min).
[2]Matrix relating each endoRNase to the type(s) of RNA that it targets.

Associated files

wcEcoli Path	File	Type
wcEcoli/models/ecoli/processes	rna_degradation.py	process
wcEcoli/reconstruction/ecoli/dataclasses/process	rna_decay.py	data
wcEcoli/reconstruction/ecoli/flat	rnas.tsv	raw data
wcEcoli/reconstruction/ecoli/flat	endoRnases.tsv	raw data

Table 2: Table of files for RNA degradation.

References

[1] Jonathan A Bernstein, Arkady B Khodursky, Pei-Hsun Lin, Sue Lin-Chao, and Stanley N Cohen. Global analysis of mrna decay and abundance in escherichia coli at single-gene resolution using two-color fluorescent dna microarrays. *Proceedings of the National Academy of Sciences*, 99(15):9697–9702, 2002.