4.3.2 - DNA Nebulization

Parameters

| Parameter | Variable | Default | Parameter | Description |
|--------------------|-------------------------------|---------|-------------------------------|---|
| Name | | Value | Range | |
| FRAG_NB_M | 📃 Unknown macro: 'mathinline' | 1 | >0 | fragmentation intensity ("rotor speed", controls the slope of the fragment size distribution) |
| FRAG_NB_LA MBDA | 📃 Unknown macro: 'mathinline' | 900 | >0 | limiting molecule size below which it is very difficult to break a |
| | | | | field |
| FRAG_NB_TH OLD | 📜 Unknown macro: 'mathinline' | 0.1 | 📃 Unknown macro: 'mathinline' | threshold on the fraction of molecules below which thermodynamics equilibrium is assumed |
| | | | | |

Algorithm

Already early reports on results from RNA-Seq experiments based on nebulization observed reads accumulating at the 5'-end of transcripts and around the center, especially of shorter transcript forms. These observations coincide well with breakpoint distributions obtained by a theoretical model of mechanical breakage that considers molecules as rigid stiffs, in which breakpoints recursively accumulate around the midpoint of iteratively broken fragments. According to this model, the average expected fragment size depends on the length of the nebulized DNA molecule: comparatively short molecules accumulate higher breaking probabilities during the time it takes to fragment the longer molecules in the transcript population.

In the light of these preliminary studies, we simulate nebulization by an iterative two-step process: first, a random orientation of the molecule in the shear

| field—i.e., the point (Unknown macro: 'mathinline') where the shearing stress is applied—is determined by random sampling un | der a | | | | |
|---|-------|--|--|--|--|
| Gaussian function centered at a molecule's midpoint. Subsequently, the breaking probability | duced | | | | |
| Unknown macro: 'mathblock' | | | | | |
| where Unknown macro: 'mathinline' is the molecule length, Unknown macro: 'mathinline' is a parameter th | nat | | | | |
| describes the limiting size below which molecules are very unlikely broken by the shearing field; is a parameter describing the force of the shear field and determines the steepness of the slope in the resulting fragment size distribution; | | | | | |
| Unknown macro: 'mathinline' finally is a constant that adjusts | size | | | | |
| exactly between Unknown macro: 'mathinline' (Unknown macro: 'mathinline') and (| | | | | |
| Unknown macro: 'mathinline'). In our model, a Bernoulli trial on Unknown macro: 'mathinline' determines | 41 | | | | |

whether a simulated break incurs at a given position. Recursive breaking continues until thermodynamics equilibrium as assumed by convergence of the fraction of breaks per time unit in the library falling below a given threshold (t=1%).