

## Variables in Murugan and Kreiman – A theoretical framework of transcription-splicing coupling

Variable	Units	Definition	Description	Where defined	Where used	Comment
$x$	bases		snRNP position	<a href="#">Page 4</a>	<a href="#">F1, Eq1-4</a>	$x(t=0)=x_0$
$y$	bases		RNAPII position	<a href="#">Page 4</a>	<a href="#">F1, Eq1-4</a>	$y(t=0)=y_0$
$t$	sec		Time	<a href="#">Page 4</a>	<a href="#">F1, Eq1-4</a>	
$x_d$	bases <sup>2</sup> /sec		1D diffusion constant for snRNP	<a href="#">Eq1</a>	<a href="#">F1, Eq1-12</a>	$0.092 \mu\text{m}^2/\text{sec} \sim 8 \times 10^5 \text{ bases}^2/\text{s}$
$k_E$	bases/sec		RNAPII elongation rate	<a href="#">Eq1</a>	<a href="#">F1, F2, Eq3-13</a>	72 bases/sec
$\xi_{x,t}$			Gaussian white noise	<a href="#">Eq1</a>	<a href="#">Eq1</a>	
$n$	bases		position of IEJ	<a href="#">Page 4</a>	<a href="#">F1, F2, Eq4-7</a>	
$P_{x,y,t}$			Joint probability	<a href="#">Eq2</a>	<a href="#">Eq2</a>	
$G_{x_0, y_0, t}$		$G_{x_0, y_0, t} = \int_0^n \int_0^n P_{x, y, t   x_0, y_0, 0} dx dy$	Probability that $x$ and $y$ are between 0 and $n$	<a href="#">Page 5 (Eq3)</a>	<a href="#">Eq3</a>	
$T_{x_0, y_0}$	Sec	$T_{x_0, y_0} = - \int_0^\infty t \partial_t G_{x_0, y_0, t} dt = \int_0^\infty G_{x_0, y_0, t} dt$	Mean first passage time	<a href="#">Page 5 (Eq4)</a>	<a href="#">F2A,B Eq4</a>	
$\tau_{S,1D3D}$	Sec	$\tau_{S,1D3D} = n/k_E + n^2/x_d + \tau_t/n$	Average overall search time (including 1D and 3D)	<a href="#">Eq5</a>	<a href="#">F2C Eq5</a>	
$\tau_t$	bases sec		3D diffusion-controlled time constant	<a href="#">Eq5</a>	<a href="#">Eq5-11</a>	$10^9 \text{ bases s}$
$n_{opt}$	bases	$n_{opt} = (\Phi^{1/3} + x_d^2 \Phi^{-1/3} - x_d) / 6k_E$	Solution when minimizing <a href="#">Eq5</a>	<a href="#">Eq6</a>	<a href="#">F2, F4 Eq56</a>	
$\Phi$		$\Phi = x_d (54\tau_t k_E^3 - x_d^2 + 6\sqrt{3\tau_t k_E^3 (27\tau_t k_E^3 - x_d^2)})$		<a href="#">Eq6</a>	<a href="#">Eq6</a>	
$\min \tau_{S,1D3D}$	Sec		Using $n_{opt}$ in <a href="#">eq5</a>	<a href="#">Page 6</a>	<a href="#">F2C</a>	
$\tau_{S,3D}$	Sec	$\tau_{S,3D} = n/k_E + \tau_t$	Average overall search time (3D only)	<a href="#">Page 6 (Eq7)</a>	<a href="#">F2C Eq7</a>	
$n_c$	Bases	$n_c = (\Omega^{1/3}/6 + 2\tau_t x_d \Omega^{-1/3})$	Position where $\tau_{S,1D} = \tau_{S,1D3D}$	<a href="#">Eq8</a>	<a href="#">F2C</a>	
$\Omega$		$\Omega = (-108\tau_t x_d + 12\sqrt{-12\tau_t^3 x_d^3 + 81\tau_t^2 x_d^2})$		<a href="#">Eq8</a>	<a href="#">Eq8</a>	
$L$	Bases		Average dissociation length	<a href="#">Before Eq9</a>	<a href="#">Eq9</a>	
$\tau_{S,d}$	Sec	$\tau_{S,d} = n/k_E + n(L^2 / (6x_d) + \tau_t/n)/L$	Search time with snRNP dissociation	<a href="#">Eq9</a>	<a href="#">F2C Eq9</a>	
$L_{opt}$	Bases	$L_{opt} = \sqrt{6x_d \tau_t/n}$	Minimizes $\tau_{S,d}$	<a href="#">After Eq9</a>		
$\min \tau_{S,d}$	Sec	$\min \tau_{S,d} = n/k_E + \sqrt{2\tau_t n / 3x_d}$	Using $L_{opt}$ in <a href="#">Eq9</a>	<a href="#">Eq10</a>	<a href="#">Eq10</a>	
$k_{on,n}$	bases <sup>-1</sup> sec <sup>-1</sup>	$k_{on,n} = 1/\tau_{S,1D3D}$	Bimolecular forward on-rate constant	Scheme I		
$k_{off,n}$	sec <sup>-1</sup>		Dissociation constant	Scheme I	<a href="#">Eq11</a>	$10 \text{ s}^{-1}$

## Variables in Murugan and Kreiman – A theoretical framework of transcription-splicing coupling

$p_{n,1D3D}$		$p_{n,1D3D} = N_0 / \left( N_0 + k_{off,n} \left( n/k_E + n^2/x_d + \tau_t/n \right) \right)$	Probability of snRNP at IEJn	<b>Eq12</b>	<b>F2D, Eq11</b>	
$N_0$	Molecules		Number of freely diffusing snRNP	<b>Eq11</b>	<b>Eq11-12</b>	
$p_{n,3D}$		$p_{n,3D} = N_0 / \left( N_0 + k_{off,n} \left( n/k_E + \tau_t \right) \right)$	Probability of snRNP at IEJn considering only 3D diffusion	<b>Eq12</b>	<b>F2D Eq12</b>	
$S_{s,n}$		$S_{s,n} = 100 \int_0^n p_{m,1D3D} dm / n$	Splicing efficiency	<b>Eq13</b>	<b>F5 Eq13</b>	
$\mu$	Bases		Value that maximizes $S_{s,n}$	<b>After Eq13</b>	<b>F6</b>	
$_{c,k}v_i$			Signal from exon at base $i$ , transcript $c$ , tissue $k$	<b>After Eq13</b>		Note: this is different from $s_{e,c,k}$
$g_{c,k}$	Molar	$g_{c,k} = \int_0^n v_i di / n$	Transcript signal	<b>After Eq13</b>		
$c$	index		Transcript index	<b>Eq 14</b>	<b>Eq14</b>	
$k$	index		Tissue index	<b>Eq 14</b>	<b>Eq14</b>	
$\epsilon$	index		Exon index	<b>Eq 14</b>	<b>Eq14</b>	
$s_{e,c,k}$	Relative units		Log2 signal exon $\epsilon$ , transcript $c$ , tissue $k$	<b>Before Eq 14</b>	<b>Eq14</b>	
$m_c$			# exons in transcript $c$	<b>Eq 14</b>	<b>Eq14</b>	
$\pi_{\epsilon,c,k}$	probability	$\pi_{\epsilon,c,k} = \frac{s_{\epsilon,c,k}}{\sum_{i=1}^{m_c} s_{i,c,k}}$	Probability that exon $\epsilon$ is included in the final transcript	<b>Before Eq 14</b>		
$\sigma_{\epsilon,c,k}$	Index	$\sigma_{\epsilon,c,k} = s_{\epsilon,c,k} / g_{c,k}$	Splicing index for exon $\epsilon$ in transcript $c$ and tissue $k$ .	<b>Before Eq 14</b>	<b>Eq14</b>	
$\Gamma_c$		$\Gamma_c = \sum_{\epsilon=1}^{m_c} \left( \sum_k (\sigma_{\epsilon,k,c})^2 - \left( \sum_k \sigma_{\epsilon,k,c} \right)^2 \right) / k$	Variance based scoring metric	<b>Eq14</b>	<b>Eq14</b>	
$f_{\epsilon,k}$		$f_{\epsilon,k} = \sum_c \left( 100 (s_{\epsilon,c,k} - s_{1,c,k}) / s_{1,c,k} \right)$	FENAS (first-exon normalized signal)	<b>Eq15</b>	<b>F4 Eq15</b>	
		$n = \theta(\epsilon)$	Position of the $\epsilon^{th}$ exon	<b>Before Eq16</b>	<b>F3C-D</b>	
$h_{m,k}$	Relative units	$h_{m,k} = \sum_{r=1}^b \left( \sum_{\epsilon=1}^m s_{\epsilon,r,m,k} / m \right) / b$	Genome level average of transcripts with $m$ exons in $k^{th}$ tissue	<b>Eq16</b>	<b>F6 Eq16</b>	
$b(m)$			Number of transcripts with $m$ exons	<b>Eq16</b>	<b>Eq16</b>	