Rare event simulation in immune biology: Models of negative selection in T-cell maturation

Corinna Ernst ¹  Ellen Baake ²

¹Genome Informatics
Institute of Human Genetics
University of Duisburg-Essen

²Biomathematics & Theoretical Bioinformatics
Faculty of Technology
Bielefeld University

10th International Workshop on Rare Event Simulation (RESIM)
Tinbergen Institute Amsterdam, 2014
Initiation of adaptive immune responses by activation of T-cells
Recognition of non-self antigens by \textit{Probabilistic Recognition} \\
van den Berg, Rand & Burroughs, 2001

- In absence of foreign antigens
- In presence of 2000 copies of a foreign antigen
Negative selection
Deletion of immature T-cells receiving high stimuli from self antigens
Contrasting models of antigen presentation in thymus

Mixture Model
Derbinsky et al., 2005

Emulation Model
Gillard & Farr, 2005

Thymus

uncorrelated expression and presentation

expression and presentation in tissue-specific subsets
- duration of binding $\mathcal{T}$ follows the $\text{Exp}(1/\tau)$ distribution
- $\mathcal{T}$ is translated into stimulation rate $W$ via

$$W = w(\mathcal{T}) = \frac{1}{\mathcal{T}} \exp \left(-\frac{1}{\mathcal{T}}\right)$$
Review: T-cell model without negative selection

Assumptions

- $n_s$ types of self-antigens are presented on an APC’s surface in (constant) copy number $z_s$
- total number $M$ of presented antigens remains constant, even in presence of foreign antigens
- at most one foreign antigen type per APC, presented in $z_f$ copies
- no a-priori difference between the stimulation rates induced by self and non-self antigens
- total stimulation rate received by an encountering T-cell is given by the sum over the stimulation rates induced by all presented antigens, i.e.,

$$G(z_f) = \left( \sum_{i=1}^{n_s} q z_s W_i \right) + z_f W_{n_s+1}$$

with $q := (M - z_f)/M$
Total Stimulation Rate & T-cell Activation

An encountering T-cell is activated, if the total stimulation rate

$$G(z_f) = \left( \sum_{i=1}^{n_s} q z_s W_i \right) + z_f W_{n_s+1}$$

exceeds a certain activation threshold $g_{act}$.

Crucial quantity is the activation probability $\mathbb{P}(G(z_f) \geq g_{act})$. 
An encountering T-cell is activated, if the total stimulation rate
\[ G(z_f) = \left( \sum_{i=1}^{n_s} q z_s W_i \right) + z_f W_{n_s+1} \]

exceeds a certain activation threshold \( g_{\text{act}} \).

Crucial quantity is the activation probability
\[ \mathbb{P}(G(z_f) \geq g_{\text{act}}). \]

For physiologically realistic values of \( z_f \) and \( g_{\text{act}} \)
\[ \mathbb{P}(G(z_f)) \geq g_{\text{act}} \gg \mathbb{P}(G(0) \geq g_{\text{act}}) \]
is required.
Importance sampling approach for the basic T-cell model
Lipsmeier & Baake, 2009

• density \( \mu^{\vartheta} \) is obtained from 'natural' density \( \mu \) of \( G(z_f) \) by exponential reweighting according to

\[
\frac{dP_{\mu^{\vartheta}}}{d\mu}(g) = \frac{e^{\vartheta g}}{E(e^{\vartheta G(z_f)})}
\]

• optimal choice of tilting parameter \( \vartheta \) (in the sense of asymptotic efficiency) is the solution of

\[
E_{\mu^{\vartheta}}(\bar{G}) = \frac{d}{d\vartheta} \log E(e^{\vartheta G(z_f)}) = g_{\text{act}}
\]

• independence of the summands of \( G(z_f) \) yields

\[
E(e^{\vartheta G(z_f)}) = E(e^{q^{\vartheta} z_s W})^{n_s} E(e^{\vartheta z_f W})
\]

• sampling estimate is given by

\[
\hat{P}(G(z_f) \geq g_{\text{act}}) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\{\bar{G}_i(z_f) \geq g_{\text{act}}\} \frac{d\mu}{d\mu^{\vartheta}}(\bar{G}_i)
\]
Including negative selection into the basic T-cell model

- negative selection of (immature) T-cell $i$ is equivalent to several ($R$) encounters with (thymic) APCs
- individual-based T-cell modelling required, i.e., each T-cell $i$ is given by a full set of $K$ stimulation rates
  
  $T_i := \{W_{i1}, W_{i2}, \ldots, W_{iK}\}$

- during an encounter with thymic APC $A_r, 1 \leq r \leq R$, T-cell $i$ receives total stimulation rate
  
  $G_i^{(r)}(0) = z_s \sum_{a \in A_r} W_{ia}$

- for a given *thymic threshold* $g_{\text{thy}}$, T-cell $i$ survives negative selection if $G_i^{(r)}(0) < g_{\text{thy}}, 1 \leq r \leq R$
Modelling thymic antigen presentation

Mixture Model
uncorrelated expression and presentation

- for sampling of fresh APC $A_r$, $n_s$ samples are drawn from $S := \{1, \ldots, K\}$ independently and without replacement

Emulation Model
expression and presentation in tissue-specific subsets

- partitioning of $S$ into $s$ distinct 'histologically related' subsets of size $|S|/s$
- for sampling of APC $A_r$, one subset is chosen as 'preferred' set with probability $1/s$
- antigens of the 'preferred' subset of $A_r$ are presented with probability $p$, others with probability $1 - p$
Thymic threshold $g_{\text{thy}}$

- neither survival or nor death of immature T-cells during negative selection are rare events
- obtaining $g_{\text{thy}}$ via simple sampling assuming survival probability $\mathbb{P}(\Omega) = 0.5$
Empirical post-selection densities of stimulation rates $W$

- negative selection reduces the self background by cutting away a significant part of the stimulation rate’s density

Parameters

- rounds of negative selection $R := 2000$
- number of self antigens $K := 1000$
- number of self antigens presented by one APC $n_s := 50$
- copy number of self antigens on the APC surface $z_s := 500$
Inclusion of negative selection turns the crucial quantity from activation probability \( \mathbb{P}(G(z_f) \geq g_{\text{act}}) \) into conditional activation probability \( \mathbb{P}(G(z_f) \geq g_{\text{act}} | \Omega) \) with \( \Omega \) denoting the event of survival of negative selection, i.e., \( \Omega := \{ G^{(r)}(0) < g_{\text{thy}}, 1 \leq r \leq R \} \).

\[
\mathbb{P}(G(z_f) \geq g_{\text{act}} | \Omega) = \frac{\mathbb{P}(G(z_f) \geq g_{\text{act}} , \Omega)}{\mathbb{P}(\Omega)}
\]
Sampling \( \mathbb{P}(G(z_f) \geq g_{\text{act}}, \Omega) \)

- for T-cell \( i \), sampling of ‘partially tilted’ T-cell
  \( T_i^* := (\bar{W}_{i1}, \ldots, \bar{W}_{in}, W_{i,n+1}, \ldots, W_{i,K}) \) due to tilting parameter \( \vartheta \)
- sampling of APC \( A_{R+1} := \{1, \ldots, n_s\} \) and \( \bar{W}_{i,K+1} \) (in case \( z_f > 0 \))
- \( T_i^* \) becomes activated by an encounter with \( A_{R+1} \) if

\[
(G_{i}^{(R+1)}(z_f))^* := qz_s \left( \sum_{a \in A_{R+1}} T_{ia}^* \right) + z_f \bar{W}_{i,K+1} \geq g_{\text{act}},
\]

where \( T_{ia}^* \) is the \( a \)-th component of \( T_i^* \)

- if \( T_i^* \) is activated
  - sample \( A_1, \ldots, A_R \) from \( S := \{1, \ldots, K\} \), due to the presumed model of thymic antigen presentation
  - T-cell \( i \) survives negative selection if

\[
(G_{i}^{(r)}(0))^* = \sum_{a \in A_r} T_{ia}^* < g_{\text{thy}} \text{ for } 1 \leq r \leq R
\]
Reduction of the self background by negative selection increases the activation probabilities of T-cells in presence of non-self antigens presented in copy numbers identical to those of self antigens.