

## Supplementary Material

Parameters	Apical Dendrite	Basal Dendrite	Soma	Axon	Unit	Modfile
$C_m$	[0.5, 10]	[0.5, 10]	[0.5, 10]	[0.5, 10]	$\mu\text{Fcm}^{-2}$	pas
$g_{pas}$	$[10^{-7}, 10^{-2}]$	$[10^{-7}, 10^{-2}]$	$[10^{-7}, 10^{-2}]$	$[10^{-7}, 10^{-2}]$	$\text{Scm}^{-2}$	pas
$e_{pas}$	[-120, -60]	[-120, -60]	[-120, -60]	[-120, -60]	mV	pas
$R_a$	[50, 200]	[50, 200]	[50, 200]	[50, 200]	$\Omega\text{cm}$	pas
$I_h$	$[10^{-7}, 10^{-4}]$	$[10^{-7}, 10^{-4}]$	$[10^{-7}, 10^{-4}]$	-	$\text{Scm}^{-2}$	lh.mod
$NaT$	$[0, 10^{-1}]$	$[0, 10^{-2}]$	[0, 5]	-	$\text{Scm}^{-2}$	NaTs2_t.mod
$NaT$	-	-	-	[0, 10]	$\text{Scm}^{-2}$	NaTa_t.mod
$NaP$	-	-	[0, 1]	[0, 5]	$\text{Scm}^{-2}$	Nap_Et2.mod
$NaV$	-	$[10^{-7}, 10^{-1}]$	$[10^{-7}, 10^{-1}]$	$[10^{-7}, 5]$	$\text{Scm}^{-2}$	NaV.mod
$KT$	-	-	[0, 1]	[0, 1]	$\text{Scm}^{-2}$	K_Tst.mod
$KP$	-	-	[0, 1]	[0, 1]	$\text{Scm}^{-2}$	K_Pst.mod
$Kv3.1$	[0, 1]	[0, 1]	[0, 2]	[0, 2]	$\text{Scm}^{-2}$	Kv3_1.mod
$Kv2\text{-like}$	-	-	-	$[10^{-7}, 10^{-1}]$	$\text{Scm}^{-2}$	Kv2like.mod
$I_m$	$[10^{-7}, 10^{-2}]$	$[10^{-7}, 10^{-2}]$	-	-	$\text{Scm}^{-2}$	Im.mod
$I_{mv2}$	-	$[10^{-7}, 10^{-2}]$	-	-	$\text{Scm}^{-2}$	Im_v2.mod
$SK$	-	-	$[10^{-7}, 10^{-1}]$	$[10^{-7}, 10^{-1}]$	$\text{Scm}^{-2}$	SK.mod
$Ca_{HVA}$	-	-	$[10^{-7}, 10^{-3}]$	$[10^{-7}, 10^{-3}]$	$\text{Scm}^{-2}$	Ca_HVA.mod
$Ca_{LVA}$	-	-	$[10^{-7}, 10^{-2}]$	$[10^{-7}, 10^{-2}]$	$\text{Scm}^{-2}$	Ca_LVA.mod
$\gamma_{CaDynamic}$ $s$	-	-	$[5 \times 10^{-4}, 0.05]$	$[5 \times 10^{-4}, 0.05]$		CaDynamics.mo d
$decay_{CaDynamic}$ $s$	-	-	[20, 1000]	[20, 1000]	ms	CaDynamics.mo d

**Table S1:** Conductance density bounds for spiny/aspiny cells. Red and blue highlight across rows or columns represent specificity to excitatory (spiny) and inhibitory (aspiny) cells respectively. Note that none of the aspiny cell reconstructions in the Allen Cell-Types database has apical dendrite markings, thus the conductances on the apical dendrite in our models is specific to spiny cells.

Feature Name	eFEL Name	Type	Description
<b>voltage_base</b>	voltage_base	Numeric	The average voltage during the last 10% before stimulus onset in mV
<b>steady_state_voltage</b>	steady_state_voltage	Numeric	The average voltage after the stimulus
<b>voltage_deflection</b>	voltage_deflection_vb_sse	Numeric	The voltage deflection between voltage base and average voltage during the last 10% of the stimulus duration in mV.
<b>sag_amplitude</b>	sag_amplitude	Numeric	The difference between the minimal voltage and average voltage during the last 10% of the stimulus duration in mV.
<b>sag_ratio</b>	sag_ratio1	Numeric	The ratio between the sag amplitude and the maximal sag extend from voltage base.
<b>decay_time_constant_after_stim</b>	decay_time_constant_after_stim	Numeric	The decay time constant of the voltage right after the stimulus.
<b>Spikecount</b>	Spikecount	Numeric	Number of spikes in the trace, including outside of stimulus interval.
<b>spike frequency</b>	mean_frequency	Numeric	Mean frequency calculated as number of action potentials during stimulation divided by time between stimulus onset and last spike in Hz.
<b>time_to_first_spike</b>	time_to_first_spike	Numeric	Time to first spike in ms.
<b>AP_amplitude</b>	AP_amplitude_from_voltagebase	Numeric	Height at peak of action potential in mV from voltage base. Mean for all AP.
<b>AP_width</b>	AP_width	Numeric	Mean of width at -20 mV of AP in ms. Mean for all AP.

<b>AHP_depth</b>	AHP_depth	Numeric	Relative voltage values with respect to voltage_base at the first after-hyperpolarization. Mean for all AP.
<b>adaptation_index2</b>	adaptation_index2	Numeric	Normalized average difference of two consecutive ISI starting from second ISI.
<b>ISI_CV</b>	ISI_CV	Boolean	The coefficient of variation of the ISI.
<b>depolarization_block</b>	depol_block	Boolean	True if depolarization block is detected during spiking.
<b>check_AISInitiation</b>	check_AISInitiation	Boolean	True if time difference between same AP recorded at axon and soma is positive False otherwise

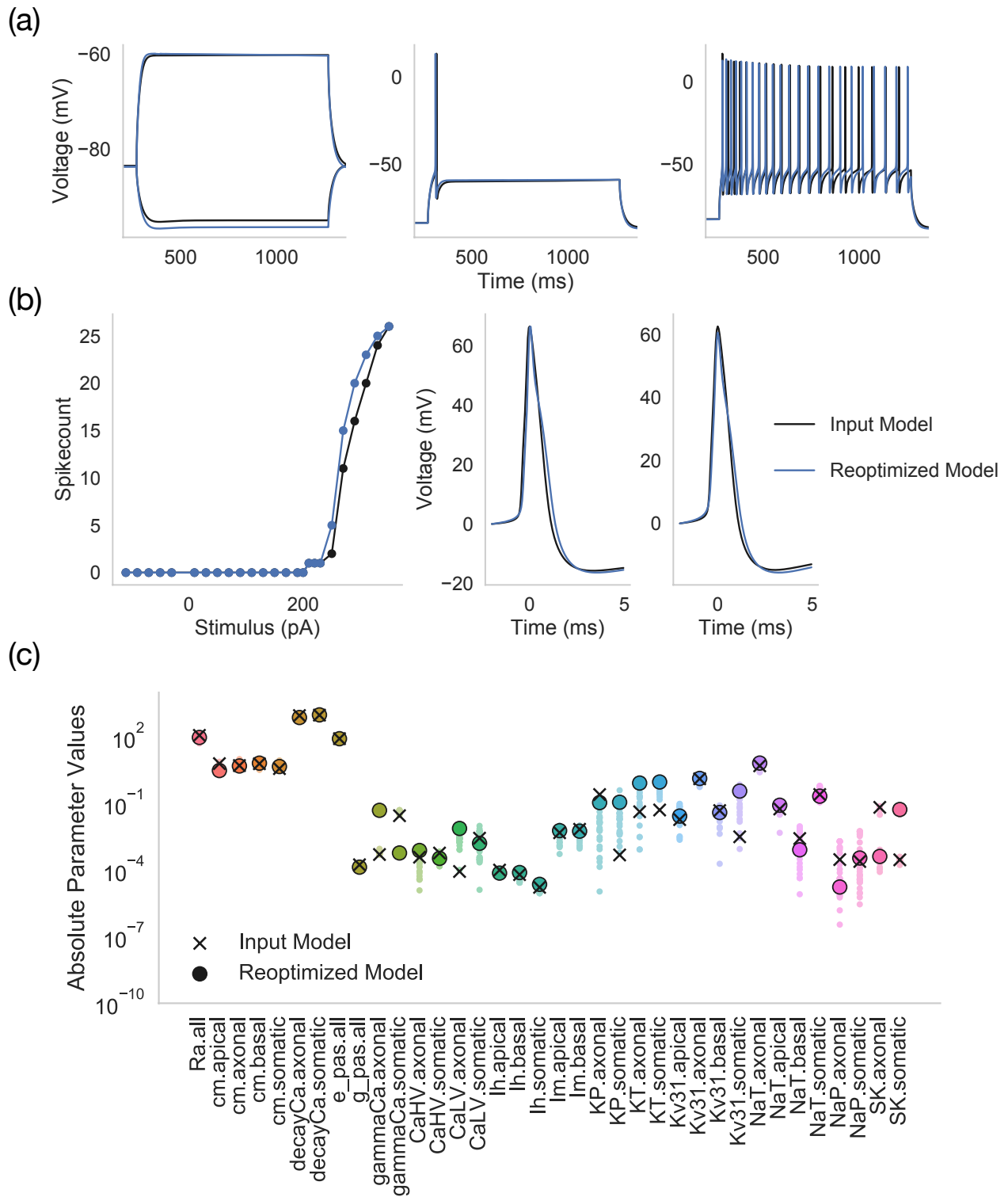
**Table S2:** Ephys features used at different stages during model generation workflow. sag\_ratio1 is not used in the optimization but used in the comparison between Nr5a1 and Rbp4 Cre-lines in Figure 6 of the main manuscript.

Stage	Parameters	Frozen Parameters/Tolerance	eFEL Features	Stimulus
<b>Stage 0</b>	<ul style="list-style-type: none"> <li><math>C_m</math></li> <li><math>e_{pas}</math></li> <li><math>g_{pas}</math></li> <li><math>R_a</math></li> </ul>	–, –	<ul style="list-style-type: none"> <li>voltage_base</li> <li>steady_state_voltage</li> <li>voltage_deflection_vb_ss_e</li> <li>decay_time_constant_after_stim</li> </ul>	All depolarizing subthreshold
<b>Stage 1</b>	<ul style="list-style-type: none"> <li><math>\bar{g}_{lh}</math></li> </ul>	None, $\pm 50\%$	eFEL Features at Stage 0+ <ul style="list-style-type: none"> <li>sag_amplitude</li> </ul>	All depolarizing subthreshold + all hyperpolarizing subthreshold
<b>Stage 2</b>	<ul style="list-style-type: none"> <li><math>\bar{g}_{NaT}</math>, <math>\bar{g}_{NaP}</math></li> <li><math>\bar{g}_{NaV}</math></li> <li><math>\bar{g}_{KT}</math>, <math>\bar{g}_{KP}</math></li> <li><math>\bar{g}_{Kv3.1}</math></li> <li><math>\bar{g}_{Kv2like}</math></li> <li><math>\bar{g}_{Im}</math>, <math>\bar{g}_{Imv2}</math></li> <li><math>\bar{g}_{SK}</math></li> <li>gamma/deca</li> <li>y_</li> <li>CaDynamics</li> </ul>	None, $\pm 50\%$	<ul style="list-style-type: none"> <li>voltage_base</li> <li>steady_state_voltage</li> <li>mean_frequency</li> <li>Spikecount</li> <li>time_to_first_spike</li> <li>AP_amplitude_from_voltagebase</li> <li>AP_width</li> <li>AHP_depth</li> <li>adaptation_index2</li> <li>ISI_CV</li> <li>depol_block</li> <li>check_AISInitiation</li> </ul>	×2 spiking traces with maximal amplitude + Rheobase trace + Maximal subthreshold

**Table S3:** Details of Optimization Workflow: We fit conductance densities ( $\bar{g}$ ) on the primary neurites-apical, basal dendrites, soma and axon, with the assumption that these conductances are distributed uniformly across each section of the reconstructed morphology belonging to the neurite.



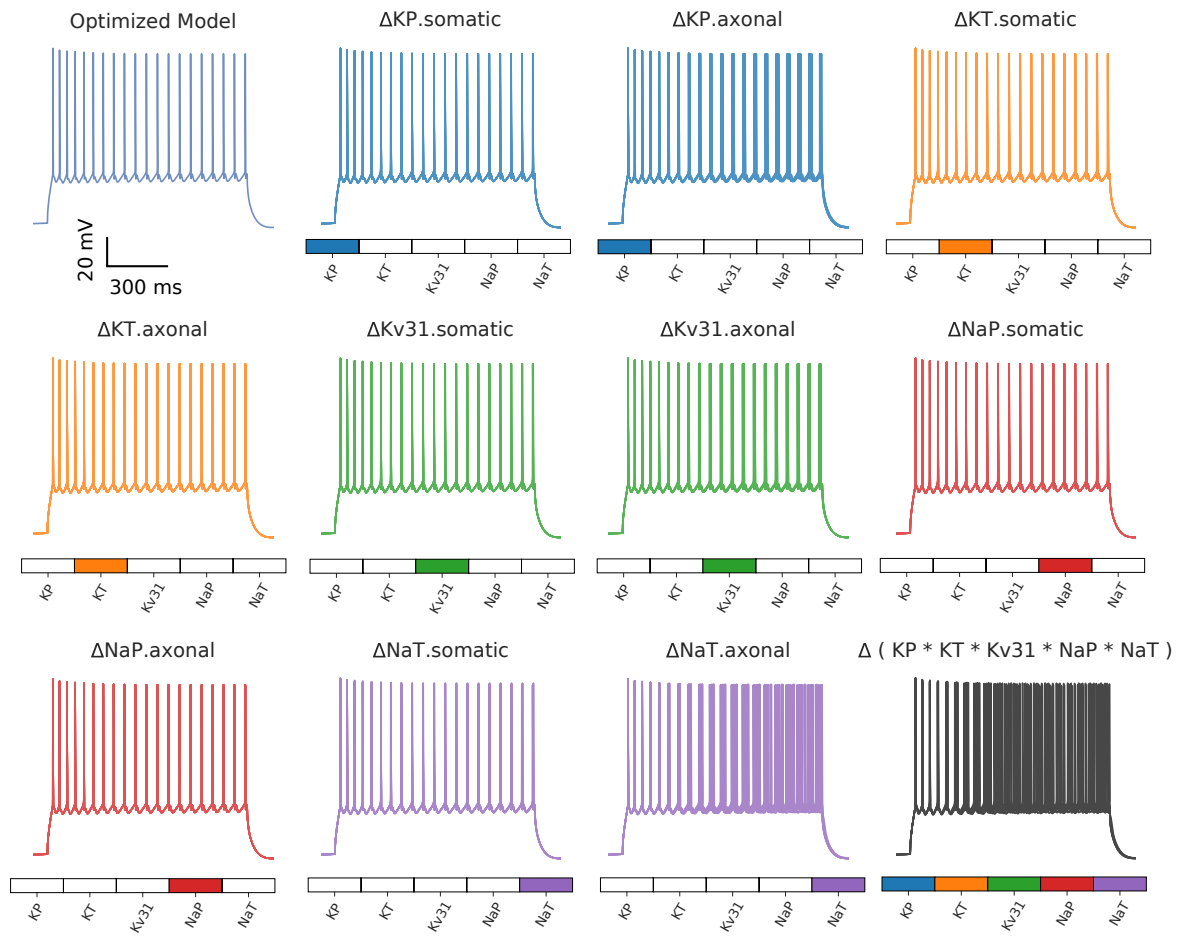




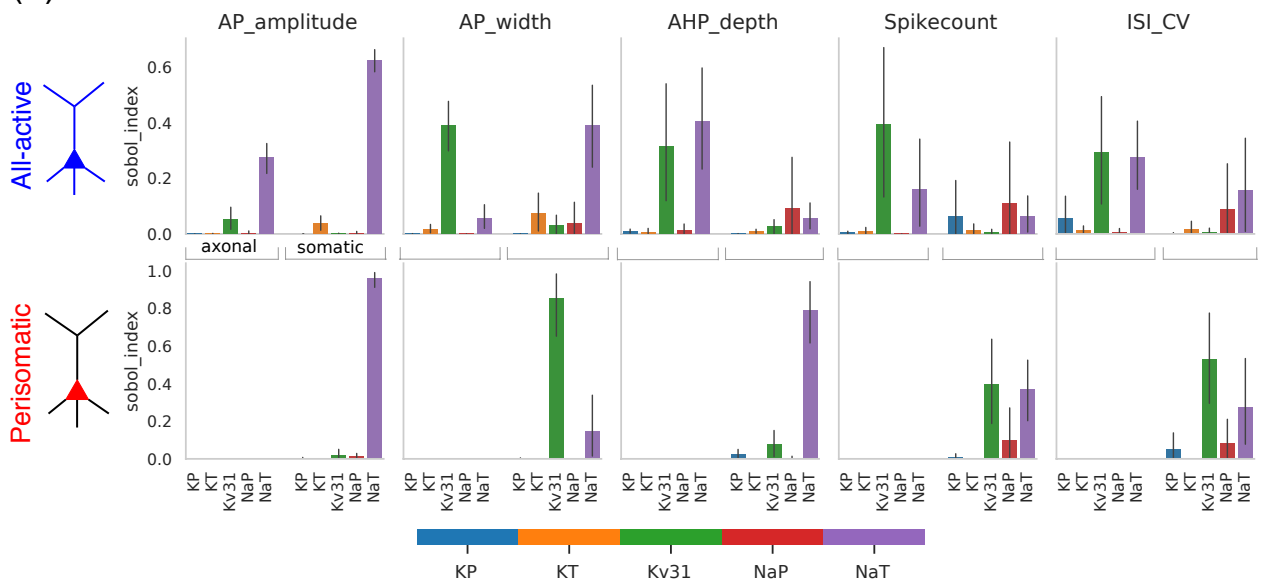
**Figure S2: Convergence of the optimization workflow for a ground truth solution.** Simulated response from the all-active model corresponding to CellID: 468193142 is fed to the workflow for a second round of optimization. **(a)** Comparison between subthreshold and suprathreshold responses as well as **(b)** fl curve and spike shapes for the input and reoptimized models. **(c)** The distribution of the hall of fame model parameters (circles) at the conclusion of the re-optimization. The best model (hof index =

0, full circle) is shown as well as the ground truth, input model (shown as 'x') overlaid to examine convergence. Note that the majority of the conductance parameters, especially the ones highlighted in the manuscript for downstream analysis in excitatory cells, i.e., passive parameters, h-channel conductance and axonal NaT, Kv31 have converged to their ground truth.

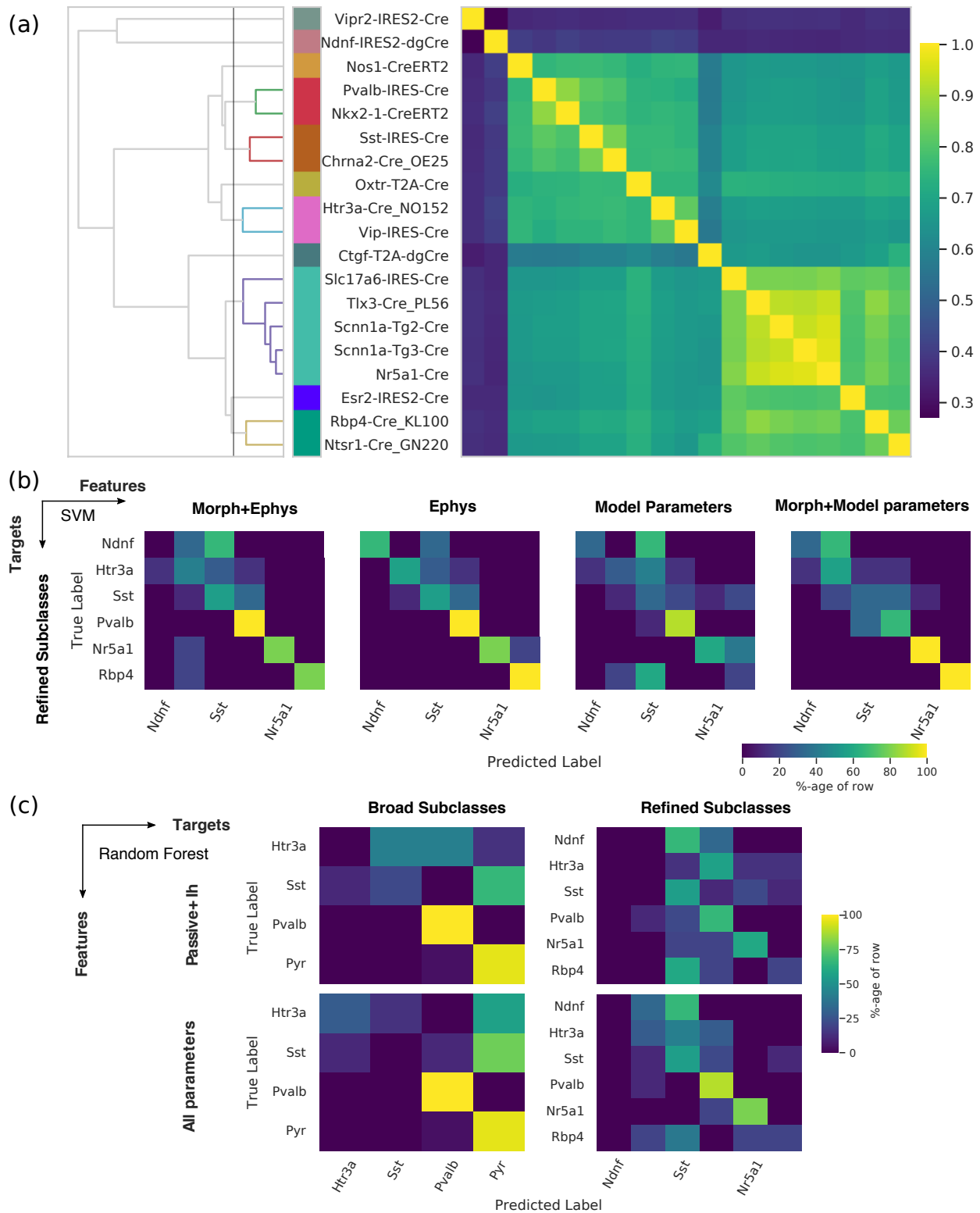
(a)



(b)

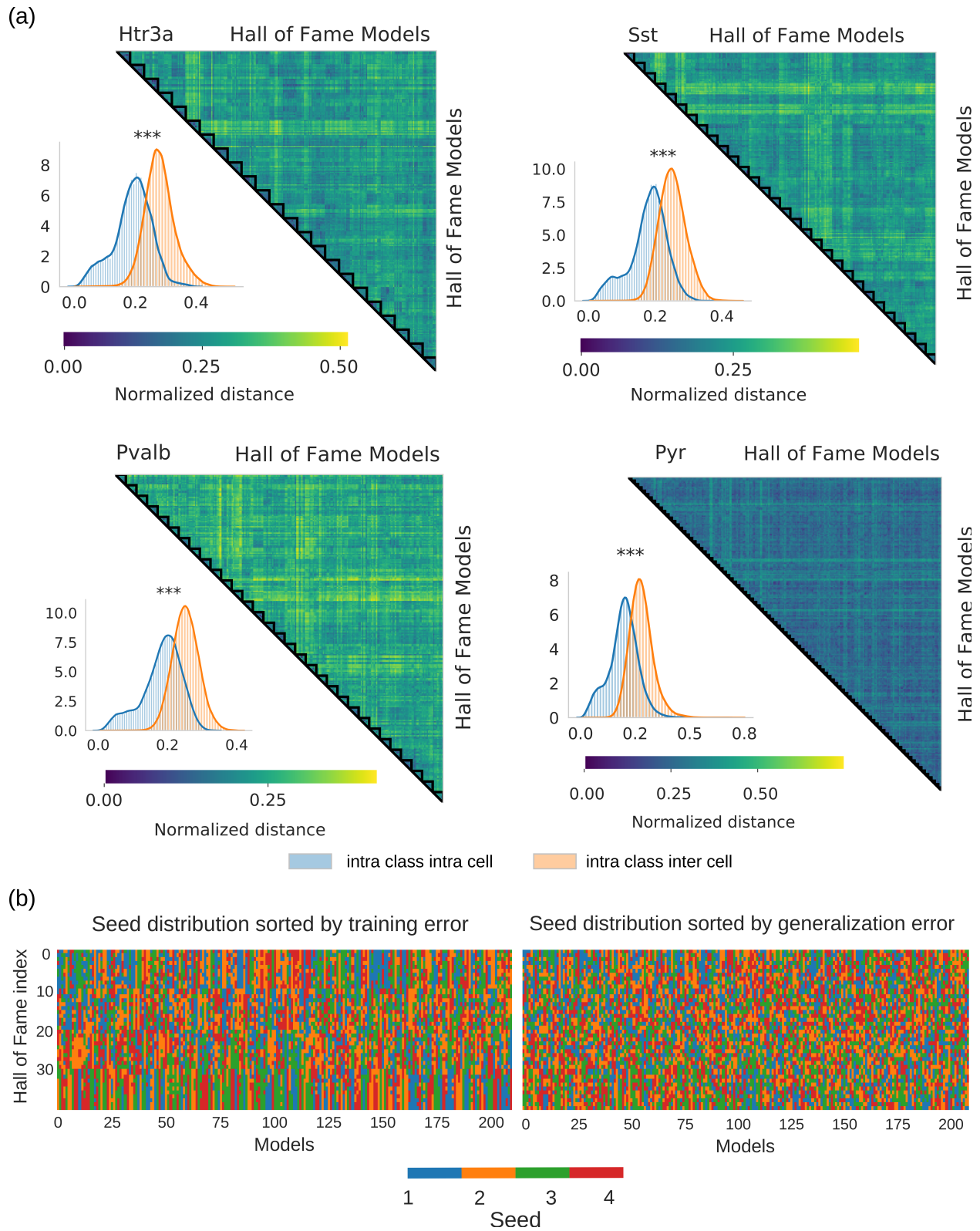


**Figure S3: Illustration of the sensitivity analysis and its utility in revealing mechanistic differences between all-active and perisomatic models.** **(a)** (top left) Output of the best model (least training error, i.e., hof index = 0) for Cell ID: 483101699 at 1s long stimulus of amplitude 270pA. In the following panels 100 simulations are shown for the same protocol with each of the selected  $K^+$ ,  $Na^+$  channel groups (indicated by the color bar at the bottom of each panel) individually perturbed within a  $\pm 10\%$  tolerance level about the optimized value. Finally, at bottom right the channels are perturbed together, equivalent to the sobol analysis computations and plot the simulations for 100 such parameter combinations. Only the parameters on the color bar are modified in the illustration, the rest of the channel densities remain unperturbed from their optimized value. **(b)** The comparison between the sobol indices as it relates to spike amplitude, width, AHP depth, spike count and ISI coefficient-of-variation for all-active and perisomatic models (10 excitatory cells each, bar: mean; error bar: std). Note the absence (due to the design of perisomatic models) of axonal conductances in the bottom panel, which play a key role for the all-active spiking features (top).



**Figure S4: Compatibility between Cre-line and transcriptomically defined cell-types and supervised classifier performance. (a) Correlation matrix between the transcriptomic types mapped to**


Cre-lines by single-cell RNA-sequencing data in a Monte Carlo fashion. The resultant Cre- to Cre-line correlation matrix (only Cre-lines from the model set are shown) is passed through hierarchical clustering resulting in similarity between the Cre-lines (left, 11 refined cell-types shown along with the dendrogram; Cre-lines of the same color are members of the same type). Confusion matrices for the **(b)** SVM classifier with 6 refined subclasses as targets and 4 feature sets (accuracy: 66%, 76%, 48% and 63% respectively) similar to **Figure 3b** and **(c)** Random forest classification (69% accuracy with both feature sets for broad subclasses; 40%, 53% respectively for reduced and all parameters respectively for refined subclasses) illustrated in **Figure 3d**.




**Figure S5: Parameter diversity across cortical cell classes.** (a) Similar to **Figure 6(d)** the dispersion matrix for the hall of fame parameters within and across cell for the 4 broad subclasses. The block diagonals for 'Pyr' are compressed due to numerical advantage in comparison to the other inhibitory subclasses. For



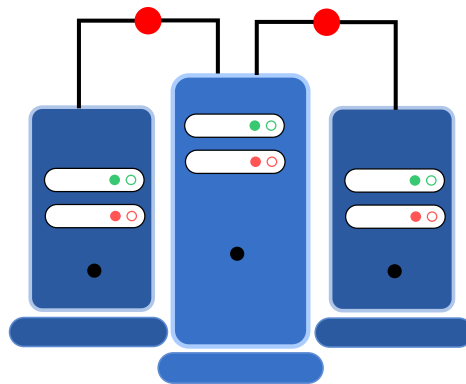
each broad subclass the intra-class intra-cell dispersion (blue) is less than intra-class inter-cell (orange). Mann-Whitney U-test; Statistical significance: \*: p-val < 0.05, \*\*: p-val <  $10^{-2}$ , \*\*\*: p-val <  $10^{-3}$ . **(b)** The hall of fame models for each cell is not dependent on the initialization. The single-color strands belonging to one seed when sorted to training error (left) is lost when the models are arranged according to the generalization error, i.e., performance on unseen stimulus set. This points to an existence of attraction basin for this highly constrained optimization formulation.




256 distributed cores  
2.2 GHz Intel Xeon E5-2630v4 Processor  
100 GB process memory  
PBS Job scheduler




16 c3.8xlarge instances  
16 cores per node  
Intel Xeon Processor  
60 GB RAM per node  
SGE Job scheduler



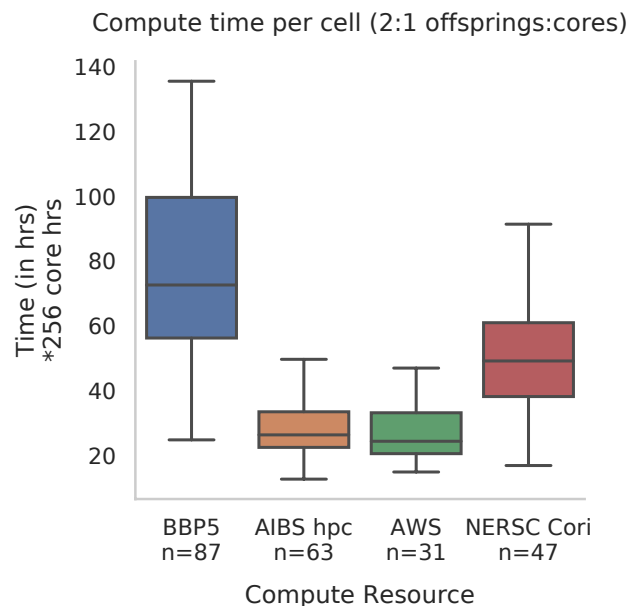
Parallelization details:  
Ipython ZMQ cluster  
ipyparallel : ipcontroller + ipengines  
load balanced view, database backend :sqlite



8 Haswell Compute Nodes  
2.3 GHz 16 cores X 2 per node  
Intel Xeon E5-2698 v3 Processor  
128 GB DDR4 RAM  
SLURM Job scheduler

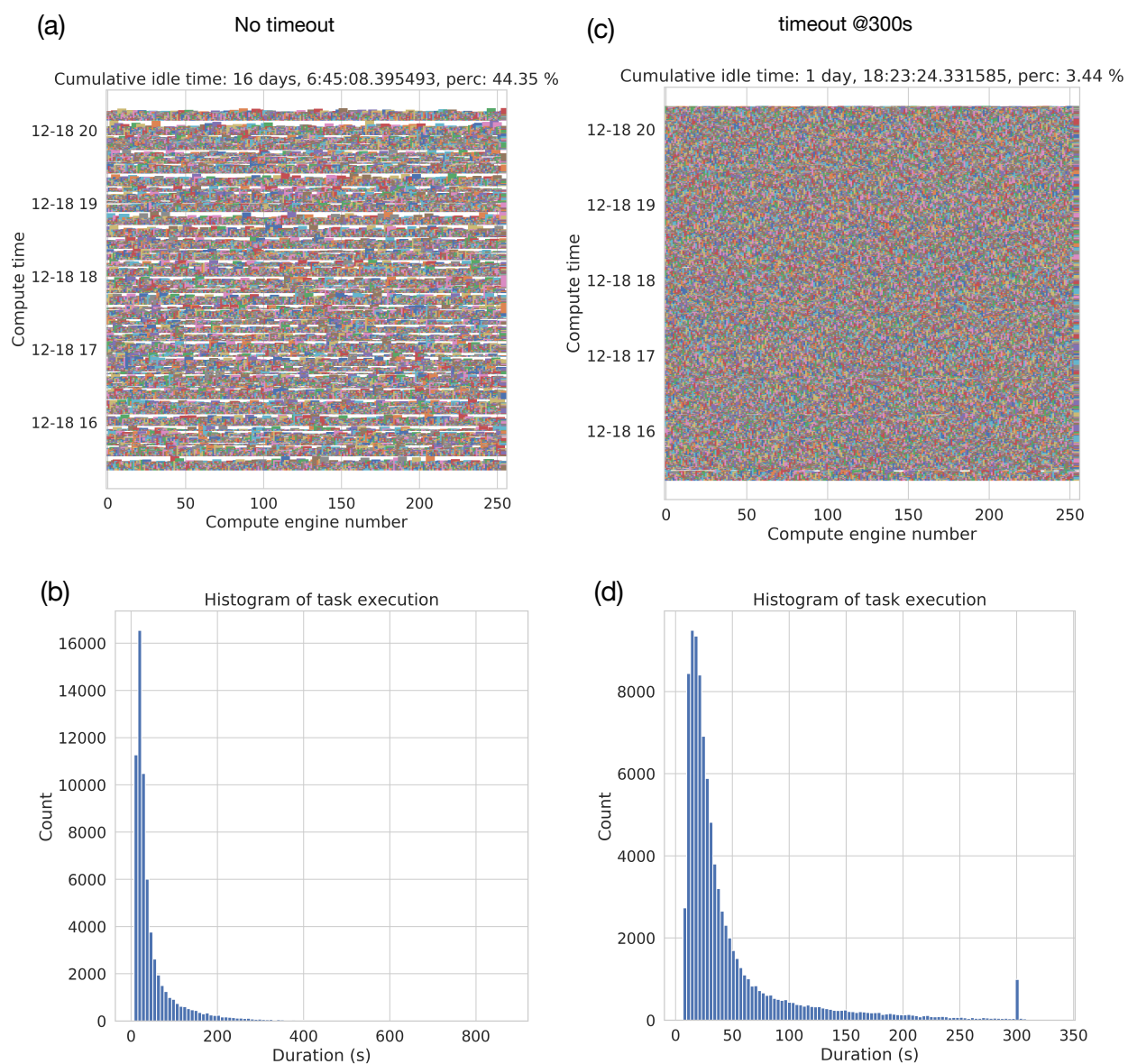


256 distributed cores  
2.3 GHz 18 cores X 2 per node  
Intel Xeon 6140 Processor  
384 GB memory per node  
SLURM Job scheduler



**Figure S6: Description of the computational resources employed during the model generation workflow.** Resources utilized thus far include the in house Allen Institute for Brain Science High Performance Computing (AIBS hpc) cluster, Amazon Web Services (AWS) EC2 instances in conjunction with Wasabi cloud storage as AWS S3 substitute, Cori supercomputers at National Energy Research Scientific Computing Center (NERSC) and Supercomputing infrastructure at Blue Brain Project (BBP5), École Polytechnique Fédérale de Lausanne, Switzerland. In total close to 3.5 million core hours are utilized across these four Linux based systems for the ~230 models. Reduction in compute time and subsequently core hours in AIBS hpc and AWS EC2 instances is mainly due to an algorithmic improvement in the latter

stages of the model generation timeline, where solutions were discarded for which evaluation exceeded a 300s cut-off. Number of requested cores at each job is equal across machines.



**Figure S7: Effect of adding timeout functionality within BluePyOpt.** (a) Without timeout the evolutionary algorithm occasionally encounters a pathological parameter combination resulting in a stiff differential equation. The blank spaces represent the idle time for the rest of engines in a distributed infrastructure. (b) The histogram of the task execution shows that the evaluation time for rare individuals go up to 900s. (c) In comparison adding a timeout of 300s enables us to run a controlled optimization in the HPC clusters with drastic reduction in idle time (44% to 3%). The jump in the histogram (d) at 300s means that the engines are reassigned to evaluate new individuals without completing the numerical integration when the cut-off time is hit. For this illustration we have used the stage 2 for model generation of CellID: 483101699 under identical parameter bounds.