Higher-Order Interactions in Neuronal Function: From Genes to Ionic Currents in Biophysical Models

Maria Reva[a,b], Alexis Arnaudon[a], Michael Zbili[a,f], Henry Markram[a,c], Jean-Marc Goaillard[e], and Werner Van Geit[a,d]

[a]Blue Brain Project, École polytechnique fédérale de Lausanne (EPFL), Campus Biotech, 1202 Geneva, Switzerland
[b]Present address: Department of Basic Neurosciences, Medical Faculty, University of Geneva, CH-1211 Geneva, Switzerlando
[c]Laboratory of Neural Microcircuitry (LNMC), Brain Mind Institute, School of Life Sciences, École polytechnique fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland
[d]Present address: Foundation for Research on Information Technologies in Society (IT'IS), Zurich 8004, Switzerland
[e]Systems Approaches to Neuronal Excitability Institut de Neurosciences de la Timone UMR7289 CNRS & AMU
[f]Present address: Université Clermont Auvergne Clermont-Ferrand, France

Abstract

Neuronal firing patterns are the consequence of precise variations in neuronal membrane potential, which are themselves shaped by multiple ionic currents. In this study, we use biophysical models, statistical methods, and information theory to explore the interaction between these ionic currents and neuron electrophysiological phenotype. We created numerous electrical models with diverse firing patterns using Monte Carlo Markov Chain methods. By analyzing these models, we identified intricate relationships between model parameters and electrical features. Our findings show that neuronal features are often influenced by multiple ionic currents sharing synergistic relationships. We also applied our methods to single-cell RNAseq data, discovering gene expression modules specific to certain interneuron types. This research sheds light on the complex links between biophysical parameters and neuronal phenotypes.

Intorduction

While neurons are often categorized under broad classes, a number of studies have demonstrated that, even in seemingly homogeneous neuronal populations, neurons exhibit profound heterogeneity across multiple dimensions, ranging from genetic and molecular differences to variations in electrophysiological properties (1–5). Thus neuronal diversity appears as an intricate and multifaceted phenomenon, involving not only variations between neuronal types but also within neuronal types.

Genetic heterogeneity, for instance, is reflected in the differential expression of genes, particularly those involved in the regulation of ion channels. Interestingly, the landscape of ion channel genetic expression does not seem to be random. Several studies found correlations between the levels of expression of various ion channels (6–9), the patterns of these correlations giving rise to co-expression modules. Moreover, these co-expression modules appear to be specific of neuronal types. This suggests that, despite the variability in ion channel expression, neuronal phenotype might be defined by a specific co-expression module rather than by individual levels of expression of ion channels. A study utilizing single-cell multiplex RT-PCR on neocortical neurons identified 26 ion channel genes clustering with three calcium-binding proteins—calbindin, parvalbumin, and calretinin—suggesting a potential link between gene expression profiles and calcium dynamics (10). In fact, single-cell transcriptomics has consistently uncovered significant variations in mRNA and non-coding RNA levels between individual neurons of the same class, emphasizing the significance of molecular heterogeneity. Electrophysiological diversity adds yet another layer of complexity to neuronal phenotype. The combination of ion channel gene expression, along with the factors governing mRNA transcription, protein translation, and post-translational modifications, plays a crucial role in shaping the electrical phenotype of a neuron (11), such that variations in electrophysiological behaviors across neurons are often associated with genetic and molecular variations(12, 13).

To decipher this complex link, tools like information theory offer valuable insights by helping researchers analyze large-scale data and distinguish meaningful heterogeneity from noise. For example, the Partial Information Decomposition for Cell regulatory networks (PIDC) enables the identification of gene regulatory networks (GRNs) from single-cell data, moving beyond pairwise correlations to uncover higher-order relationships (14).High-order interactions in complex systems such as gene regulatory networks can exhibit both synergistic and redundant characteristics. Synergistic interactions occur when the combined effect of multiple components is greater than the sum of their individual effects. This means that certain information is only accessible when considering the joint state of a group of variables, rather than any single component alone. In contrast, redundant interactions involve multiple components providing overlapping or duplicate information, where the same information is present in different parts of a system. Synergistic and redundant interactions between genes play crucial roles in biological systems, influencing phenotype, drug response. Studies have demonstrated that synergistic interactions can enhance cooperativity in mutualistic assemblages and predict compound synergy in drug combinations, while redundancy in gene copies can provide advantages under conditions promoting rapid growth and translation, highlighting the complex interplay between these types of interactions in genetic networks (15, 16).

The focus of our research is to provide a direct link between variations in gene expression and variations in electrical phenotype. To achieve this, we first investigate conductance-based electrical models that replicate the electrophysiological behavior of excitatory and inhibitory neurons. By leveraging our previous work (17, 18), we generate populations of neurons that display similar electrophysiological properties but are

characterized by different sets of underlying parameters. This approach allows us to explore how high-order interactions among ionic currents relate to electrophysiological phenotypes. Using methods from information theory, such as Organisation information (O-info) (19), we examine these high-order interactions both in isolation and in relation to specific electrophysiological features. Additionally, we analyze publicly available Patch-seq data from interneuron populations, discovering contrasting results that suggest the prevalence of redundancy within single-cell transcriptomics data. This underscores the complexity and variability of the relationships between genetic expression and neuronal function.

Results

Synergistic high-order interactions of ionic currents in biophysically detailed models. To investigate the presence and impact of high-order interactions on the electrical phenotype (e-phenotype) of neurons, we used a detailed biophysical model of Layer 5 Pyramidal Cells (L5PC) of the somatosensory cortex (SSCx) (Fig. 1.A) (17). This model, incorporating region-specific distributions of ion channels and calcium dynamics across the soma, dendrites, and axon (present in the model as AIS followed by a myelinated compartment), underwent parameter exploration using a Markov Chain Monte Carlo (MCMC) approach (18). Through this process, we generated 5000 models capturing 60 electrical features of L5PC, presenting a wide range of ion channel conductances, calcium dynamics, and other biophysical properties. This process created a population of parameter sets representing a single e-phenotype of L5PC neuron. Importantly, this population of models display a variability in electrical features similar to the biological diversity observed in L5PC biological population (18).

The resulting parameter distributions revealed substantial variability in conductances such as gNaT and gK_Pst, as well as in the parameters reflecting calcium decay and free calcium levels across neuronal compartments (Fig. 1.B). This parameter variability underscores both variability in e-phenotype but also the fact that different parameter sets are able to achieve the same electrical output. Moreover, the parameter space was homogeneous and independent of the model fitness (see Sup. Figure 1.A). Next, we examined the relationships between these parameters by calculating pairwise mutual information (MI), which revealed significant interdependencies (Fig. 1.C). Notable interactions included axonal sodium conductance (gNaT in axon) and its decay constant along the apical trunk, correlations between axonal and somatic sodium conductances, and strong interactions between calcium conductances (gCa_HVA, gCa_LVA) and intracellular calcium levels. Overall, 17 out of 30 parameters demonstrated significant MI, indicating that extensive parameter interdependence is needed to build this e-phenotype.

To move beyond pairwise interactions, we applied high-order information measures to characterize the presence and nature of multi-parameter interactions (19). We first used O-information (O-info), a measure able to reveal redundant (O-info>0) or synergistic (O-info<0) interactions between variables. We computed O-info for combinations of 3 parameters, 4 parameters, and so forth up to 30 parameters, thereby spanning the full range of possible multi-parameter interactions (Fig.1.D). The O-info values decreased consistently as the tuple size increased, reaching a minimum of 25 parameters. Beyond this point, adding more parameters did not further change the O-info measure. The negative value of O-info indicated a strong prevalence of synergistic interactions between model parameters. Synergistic interactions, in this context, are akin to collective computations within the neuron, where multiple parameters interact in a non-additive way to create outcomes that would not be predictable by examining each parameter individually. These synergistic effects suggest that certain groups of ion channels and calcium dynamics act together to influence the neuron's behavior in a way that goes beyond the sum of individual influences. Notably, the O-info associated with redundancy (O-info>0) was minimal, approximately 10 times lower than that for synergy, highlighting the dominance of synergistic interactions (Fig.1.D).

To identify the specific parameters that contribute most significantly to the minimum O-info values, we analyzed the composition of parameter combinations associated with the highest levels of synergy. Specifically, we focused on the parameters that appeared most frequently in the top 5% of tuples with the lowest O-info values for each tuple order. For each tuple size (e.g., combinations of 3 parameters, 4 parameters, and so on), we calculated how often each parameter appeared in the most synergistic tuples, providing a frequency score that reflects each parameter's contribution to high-synergy interactions across different tuple orders (Fig.1.E) (for the redundancy related map see Sup. Figure 1.B). Key contributors to these synergistic interactions included parameters related to calcium dynamics, such as Free_Ca2+ (in both soma and dendrites) and Ca2+ decay, which frequently appeared in the most synergistic tuples. Additionally, potassium channel conductances, specifically gK_Pst and gKv3.1, were also frequently present in top synergistic tuples. This indicates that potassium conductance is another major factor shaping the collective dynamics, potentially influencing the neuron's capacity to generate or shape its electrical features. Other channels, including gNaT, gCa_LVA, and gCa_HVA, showed consistent but moderate contributions to synergy. Interestingly, the parameter that appeared least frequently in the most synergistic tuples was axonal gK_Tst. This aligns with our parameter sensitivity analysis (17), which showed that removing gK_Tst from the model had minimal impact on the model's output. This finding suggests that gK_Tst plays a limited role in driving the neuron's collective, synergy-based dynamics and may not be critical for the neuron's overall behavior.

Dependencies between electrical features and model parameters in biophysically detailed models. To map out the relationships between the electrical phenotype and the underlying ionic currents, we examined the relationship between the neuronal model outputs (electrical features, or "e-features") and the model parameters. First, we extracted e-features from voltage traces produced by each generated model (Fig.2.A left). We then computed MI between each pair of e-feature and biophysical parameters to quantify the dependencies. All e-features showed statistically significant MI values, ranging from 0.01 to 1. For further analysis, we focused on e-feature/parameter pairs with MI greater than 0.1 (Fig. 2.A right).

High MI values revealed distinct relationships between e-features and specific parameters. Calcium concentrations in

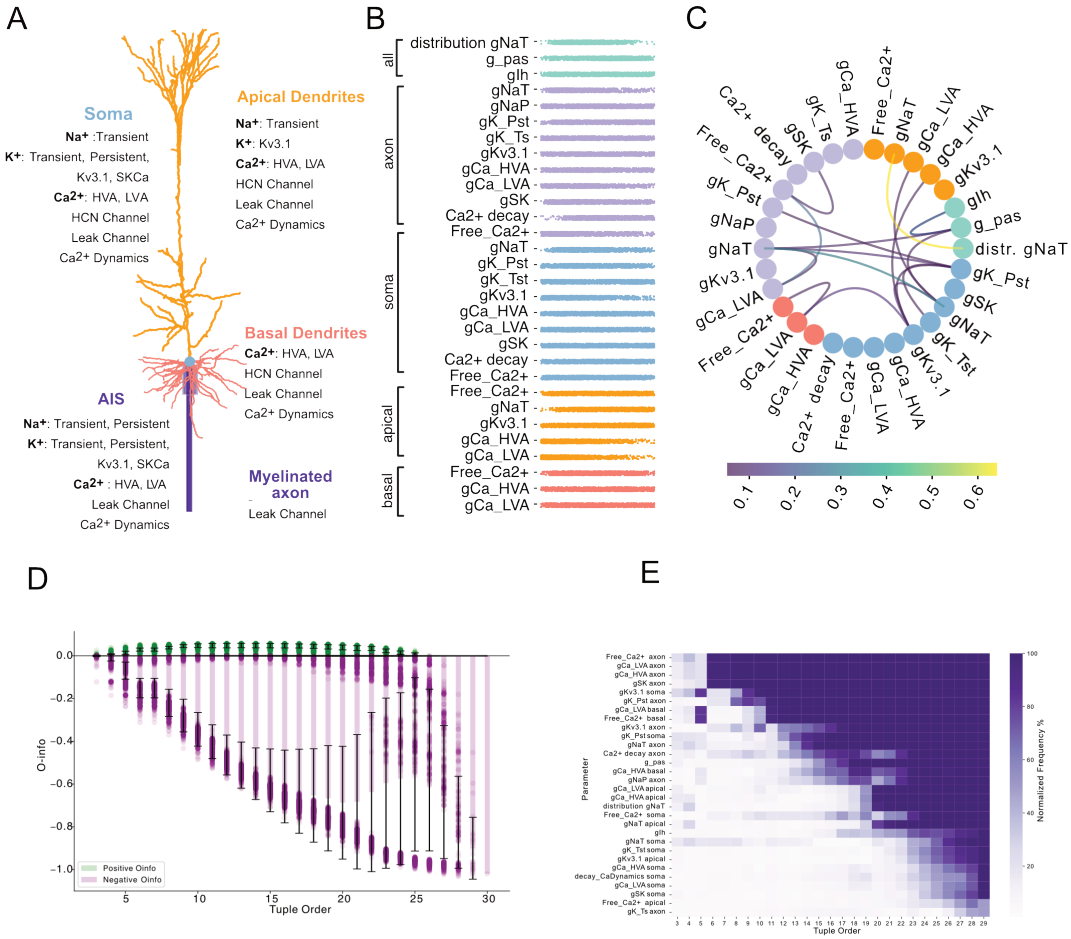

Fig. 1. Analysis of a biophysically detailed L5PC neuron model and parameter interaction. **A.** Morphology of the L5PC neuron model, showing the distribution of ion channels (Na+, K+, Ca2+) across soma, dendrites, AIS, and axon. **B.** Parameter distributions from 5000 models generated using MCMC. **C.** MI map between parameter pairs. The colors of the nodes correspond to the color of the compartments of the neural morphology from A and B. The color of the edges corresponds to the MI value. **D.** O-information plot showing the relationship between tuple number (number of parameters) and O-information values. Each dot represents the O-information value for a specific tuple. The bars illustrate the spread and variability of O-information values at each tuple order, with error bars indicating the range. The color encodes positive (green) and negative (purple) O-info values. **E.** Normalized frequency map of parameters frequently appearing in high-synergy tuples.

the axon and back-propagating action potential (bAP) features in the basal apical dendrites displayed strong associations with calcium-related parameters in the respective compartments. Somatic and axonal sodium conductances each exhibited dependencies with rheobase (I_Threshold), while action potential (AP) amplitude and afterhyperpolarization (AHP) depth were influenced by axonal sodium conductances. Mean spiking frequency showed a strong relationship with axonal calcium decay, whereas passive conductances, including g_pas, shared high MI with the time to first spike (the feature used is the inverse of time to first spike 1/TFS). Input resistance (Rin) and resting membrane potential (RMP) both displayed high MI with g_pas (Fig. 2.A right).

In order to determine whether specific sets of parameters collectively determine specific e-features, we moved beyond pairwise interactions and explored higher-order interactions between parameters and specific e-features. We focused on two e-features: resting membrane potential (RMP) and the first interspike interval (ISI). Intuitively, these e-features are expected to involve distinct parameter sets due to their different nature : RMP should mainly rely on passive currents while ISI is influenced by active currents. Additionally, while RMP shared high pairwise MI with two distinct parameters, the first ISI did not have any statistically significant MI with individual parameters. To quantify higher-order information between parameter sets in relation to each e-feature, we applied the Redundancy-Synergy Index (RSI) (20), which allows us to compute conditional mutual information and therefore understand which parameters interactions are responsible for a given e-feature. As for O-info, negative RSI values indicate synergistic interactions, while positive values reflect redundant interactions.

For RMP, the RSI reached its lowest value (-0.35) with just two parameters. With additional parameters, the RSI did not increase, remaining at -0.35, suggesting that any additional parameters do not contribute significantly to RMP's RSI (Fig. 2.B top). When examining the identity of these parameters, we found that they are gIh and g_pas (Fig. 2.B middle, bottom). This aligns with findings in pyramidal cells that demonstrated the implication of Ih and passive current in shaping somatic RMP (21). These parameters formed a synergistic module specifically for RMP, highlighting their central role in regulating this feature (Fig. 2.B bottom). In contrast, when examining higher-order information for the first ISI, we observed that RSI values continued to decrease as the tuple size increased, stabilizing around 10 parameters (Fig. 5.C top). The parameters constituting the maximum synergy for ISI included axonal calcium and sodium conductances, somatic potassium and sodium conductances, and apical calcium conductances (Fig. 5.C middle, bottom). This result suggests that ISI is influenced by a more complex module involving multiple conductances across various compartments.

Overall, these results demonstrate that the behavior of the L5PC neuron model is governed by complex, synergistic interactions among parameters, particularly those related to calcium handling and potassium conductance across different compartments. These synergistic modules are specific to certain e-features, and their complexity can be revealed by examining higher-order interactions. These findings underscore the importance of multi-parameter synergy in shaping distinct aspects of neuronal function, with RMP and first ISI being controlled by distinct, high-order parameter modules.

High-order interactions are specific to the electrical phenotype. So far, we have examined the high-order interaction modules for a single electrical phenotype, the L5PC neuron. However, an open question remains: would these high-order interaction modules differ in models producing different electrical phenotypes? To investigate this, we examined two distinct electrical models of cortical interneurons from the somatosensory cortex (SSC): the continuous accommodating cells (cAC) and the bursting non-accommodating cells (bNAC) (22). Both models were optimized to match the respective electrophysiological recordings obtained from these neuronal types (17) and share identical parameter sets and morphology, allowing for a meaningful comparison of optimized parameter values (Fig. 3.A).

When we compared the distributions of values of optimized parameters between the two models, we found that 19 out of 21 parameters exhibited similar distributions. The only significant differences were observed in the axonal and somatic sodium conductances (gNaT), which varied between the two models (Fig. 3.B). This indicates a high degree of similarity in the underlying parameter distributions between the cAC and bNAC models, with minimal differences confined to sodium conductances.

Furthermore, both models displayed a predominance of synergistic interactions across parameter tuples, as evidenced by the consistently negative O-info values. For both models, the O-info values stabilized at a tuple order of 14, indicating that beyond this point, additional parameters do not contribute significantly to the synergy or redundancy in the model (Fig. 3.C). This stabilization suggests that a similar number of parameters is required to capture the full scope of synergistic interactions in both electrical phenotypes (see Sup. Figure 2).

Given the similarity in the initial parameter space, we explored whether the same co-variation modules could drive electrical features in each model. Specifically, we focused on the features reflecting the firing pattern of the cells: mean firing frequency and inverse first ISI (1/ISI) features in response to a current injection of 200% of the rheobase. Notably, the mean firing frequency for the bNAC model was 20 Hz ($\pm$ 7.6), higher than the mean frequency of 7.4 Hz ($\pm$ 2.9) for the cAC model (Fig. 3.D). The RSI values for both models indicated strong synergy. For the bNAC model, the minimum tuple order that achieved the most negative RSI (indicating maximum synergy) was 5 (4 parameters), whereas for the cAC model, it was 8 (7 parameters). In the bNAC model, axonal LVA and HVA calcium channels, SK channels, and calcium dynamics formed the minimum parameter set that produced maximum synergy. For the cAC model, the minimum synergistic module included the same parameters as for bNAC, along with somatodendritic LVA calcium channels, axonal Kv3.1 channels, and free calcium. Interestingly, at a tuple order of 6, both models share the same core parameters, indicating close similarities between their underlying interaction structures.

To examine whether this similarity extends to other features, we tested if the co-variation modules for bNAC would be a subset of those for cAC for another feature—in this case, the inverse first interspike interval (ISI). Consistent with their electrical phenotypes, the inverse first ISI was larger for bNAC (29.9 $\pm$ 10.3) than for cAC (9.6 $\pm$ 6.7) (Fig. 3.E, right). For

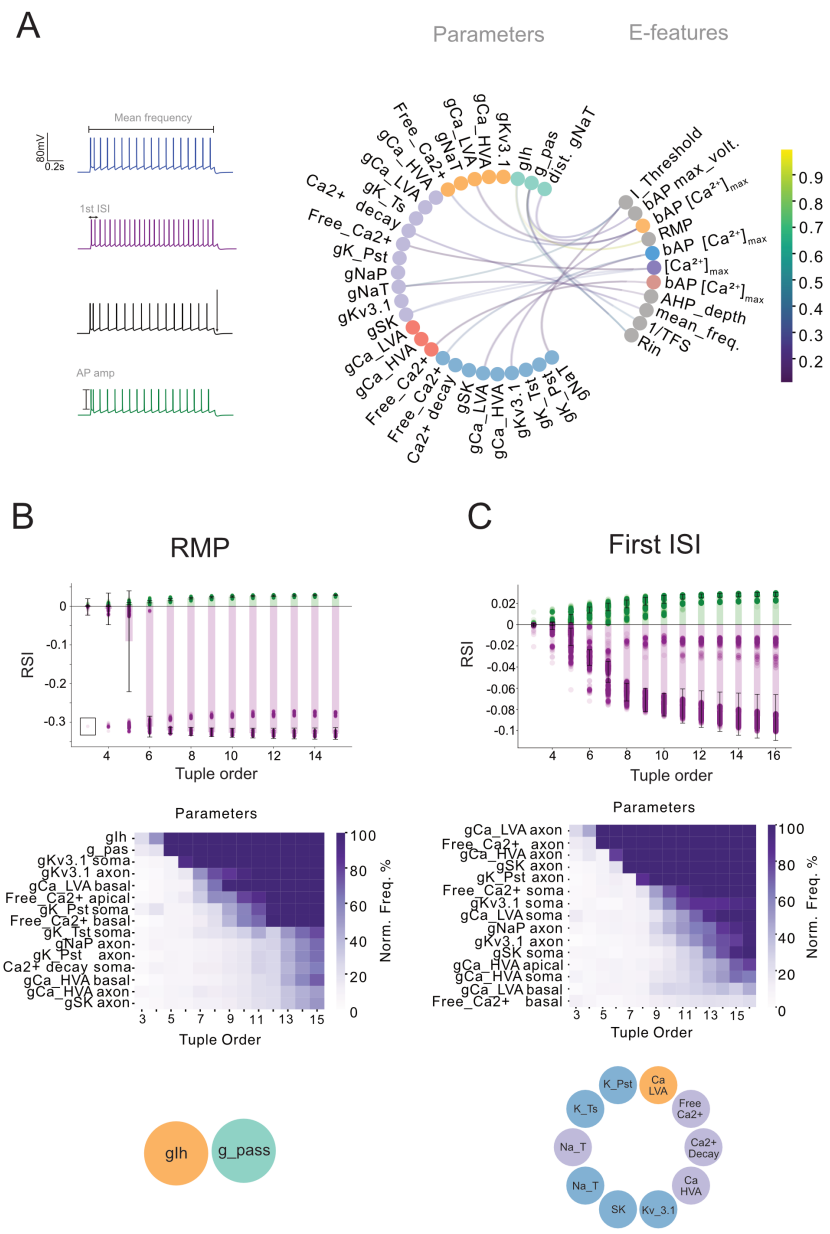

Fig. 2. Pairwise and high-order interactions between electrical feature and model parameters in L5PC model. **A**. Left: Example voltage traces from generated L5PC models in response to a step current injection (200% of their rheobase), along with examples of the extracted electrical features. Right: MI map showing the relationships between model parameters and electrical features (e-features). Parameters are color-coded according to Fig. 1.A, while electrical features are shown in grey, except for back-propagating action potential (bAP) features, which are colored to match the compartment where they were measured. **B.** Top: RSI values for each parameter combination as a function of tuple order for RMP. Middle: Parameters that most frequently appear in the top 5% of tuples with minimum RSI values, indicating their importance in synergy-driven interactions. Bottom: Schematic illustrating the parameters associated with minimum RSI values, with colors corresponding to their anatomical location in the neuron model. C. Similar analysis as in B, but for the First ISI (Interspike Interval) feature.

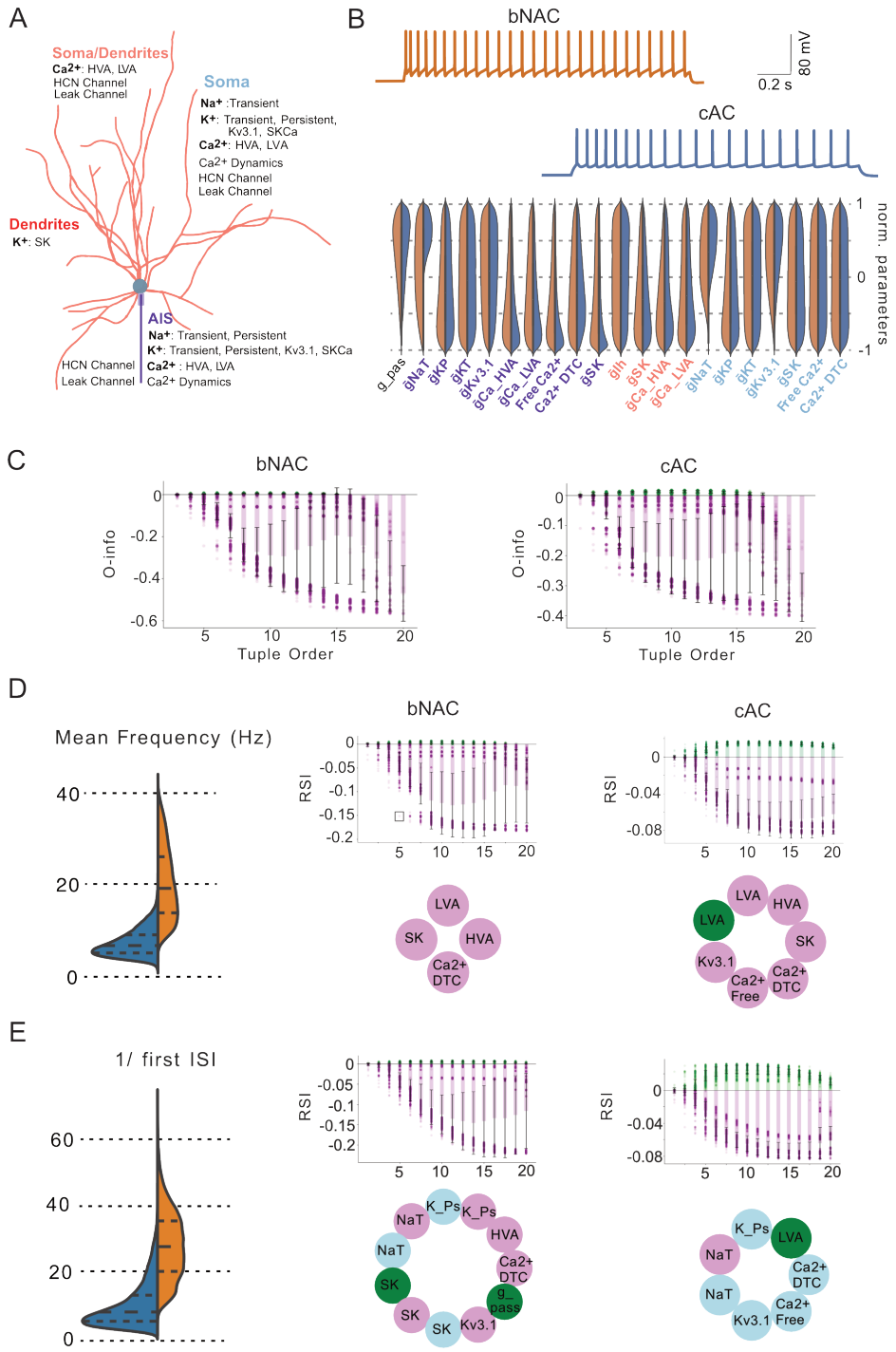

Fig. 3. High-order interaction in the models of bNAC and cAC interneurons. **A.** Morphology of the interneuron model used for optimization, illustrating the distribution of ion channels (Na+, K+, Ca2+) across the soma, dendrites, AIS, and axon used in the model. **B.** Top: Example voltage traces from the continuous accommodating (cAC) and bursting non-accommodating (bNAC) models (in response to IDrest protocol, 150%/140% from the rheobase). Bottom: Distribution of optimized parameters, with values for cAC shown in blue and bNAC in orange. **C.** O-information values as a function of tuple order, showing the progression of high-order interactions for bNAC (right) and cAC (left) models.**D.** Left: Mean firing rate distribution for bNAC (orange) and cAC (blue). Middle: RSI plot for bNAC, showing the minimum parameter set driving high synergy. Right: RSI plot and minimum parameter set for cAC, similar to the middle panel. **E.** Similar analysis as in D for the inverse first interspike interval (ISI) feature.

the bNAC model, we observed a steep and prolonged decay in RSI values, requiring 12 parameters to reach the minimum RSI, indicating significant synergy (Fig. 3.E, middle). In contrast, only 7 parameters were needed to achieve minimal RSI values in the cAC model (Fig. 3.E, left). Notably, only a quarter of the parameter sets overlapped between bNAC and cAC for this feature, indicating that the co-variation modules are highly specific to the electrical phenotype.

In summary, these results demonstrate that while cAC and bNAC models share similarities in their synergistic parameter modules for certain features, these modules are distinct and phenotype-specific for other features. The specificity of co-variation modules highlights that the complexity of electrical phenotypes emerges from unique sets of synergistic interactions, tailored to each neuron type's characteristic firing behavior.

Ion channel mRNAs as key elements identifying neuronal types. To investigate high-order interactions beyond in silico model populations, we leveraged patch-seq data (12) to explore the relationship between mRNA expression levels and electrical features in biological cortical GABAergic interneurons. Since our in silico results focused on interactions between ionic currents, our first step was to determine the role of ion channels in defining principal interneuron subtypes.

We began by isolating ion channel genes (a set of 150 among 45 768 mRNA) from the gene expression data (12) and performed clustering, aiming to capture the main interneuron types. When we compared our clustering results with the ground truth classifications provided by the dataset (Fig. 4.A), we found that most primary clusters were well-replicated using only these 150 ion channel genes. Notably, we observed distinct clusters for PV, SST, and LAMP5, indicating that ion channel gene expression alone effectively identifies these major interneuron subtypes. However, the clustering of ion channel genes merged SNCG and VIP into a single cluster, and an additional two clusters emerged—one corresponding to a mixed population of the primary cell types and the other to a distinct SST subtype.

To test whether this clustering was specific to ion channel genes, we repeated the clustering process with randomly selected mRNA subsets. We randomly sampled 150 genes 500 times and used these samples for clustering (Fig. 4.B), comparing the Adjusted Rand Index (ARI) of the resulting clusters with the ground truth (Fig. 4.C). The ARI of the ion channel gene clustering was 0.69, at least seven times higher than the average ARI obtained from random gene sets (mean ARI of 0.05; grey violin plot) (Fig. 4.D). This difference suggests that ion channel genes have substantial predictive power for interneuron subtypes compared to random mRNA subsets.

To further constrain our comparisons, we examined differentially expressed genes (DEGs) (see Methods) to see if a more specific subset of mRNAs could match the clustering power of ion channel genes. Using the ground truth, we identified DEGs across cell types and noticed that only a subset of ion channel mRNAs was comprised in these DEGs (Fig. 4.C). We then sampled 150 DEGs 500 times and performed clustering. The average ARI for DEGs (mean ARI of 0.4; blue violin plot) was higher than the ARI for randomly sampled genes from the full gene pool (mean ARI of 0.05), showing that DEGs generally have better clustering performance. Importantly, the ARI for ion channel genes was within the top 1% of the ARI distribution generated by DEG clusters, confirming that ion channel genes have a predictive power greater than 99% of DEG subsets (Fig. 4.D).

In summary, these findings demonstrate that ion channel gene expression alone provides a robust basis for distinguishing interneuron subtypes, outperforming both random mRNA samples and most DEG subsets. This result emphasizes the unique and predictive role of ion channels in defining neuronal identity and capturing the main structural and functional variations among interneurons.

High-order interactions between ion channel mRNAs specific to the electrical phenotype. To investigate whether ion channel-encoding genes share statistical interactions specific to neuronal subtypes, we focused on two electrically different interneuron types: Lamp5 and Pvalb. Pvalb interneurons are characterized by a fast-spiking pattern of activity, while Lamp5 neurons exhibit slower, delayed spiking (12). We selected 50 ion channel genes with non-zero median expression in both subtypes, including 2 HCN channels, 6 calcium channels and subunits, 6 chloride channels and subunits, 4 sodium channels, and 32 potassium channels.

First, we analyzed pairwise interactions between these ion channel genes (Fig. 5.A). The Lamp5 neurons exhibited higher mutual information (MI) values, reaching up to 0.26, compared to Pvalb neurons, which peaked at 0.19. In total, 58% of the ion channels in Lamp5 displayed statistically significant pairwise interactions, whereas only 34% of ion channels in Pvalb did so. In Lamp5 neurons, these interactions involved a broad range of potassium channels (Kcnq, Kcnk, Kcnj, Kcnt, Kcnc, Kcnn, and Kncd families), as well as calcium channels (Cacna1 family), chloride channels (Clcn family), and sodium channels (Scn family). In contrast, the Pvalb neurons exhibited a more focused pattern, with strong correlations primarily among potassium channels from the Kcn and Kcna families, along with select sodium channels from the Scn family.

We then examined high-order interactions among ion channel genes for both cell types. Both cell types showed a strong prevalence of redundant interactions (positive O-info), as indicated by negative O-information values fluctuating around zero. The O-info increased monotonically with tuple size, with potassium-encoding genes dominating the tuples that exhibited the highest redundancy (see Supp. Figure 3).

To explore the link between electrical features and ion channel gene expression, we analyzed electrical features from patch-seq data alongside gene expression profiles from the same individual cells for both Lamp5 and Pvalb neurons. We focused on features that distinguish these cell types, particularly action potential half-width (AP half-width), known to be shorter in Pvalb neurons (thus allowing fast-spiking). The AP width in Pvalb was threefold shorter (0.4 ± 0.1 ms) than in Lamp5 (1.1 ± 0.5 ms). Redundancy-Synergy Index (RSI) analysis again revealed predominantly redundant interactions among ion channel genes in both types. The tuple order yielding the highest RSI was similar between cell types (15 for Lamp5 and 14 for Pvalb). However, the genes driving maximal RSI differed: in Lamp5, potassium channel genes predominated, along with sodium channels and one chloride channel, while in Pvalb, we observed contributions from Hcn1, multiple calcium channels, and sodium channels. We further extended our analysis on the

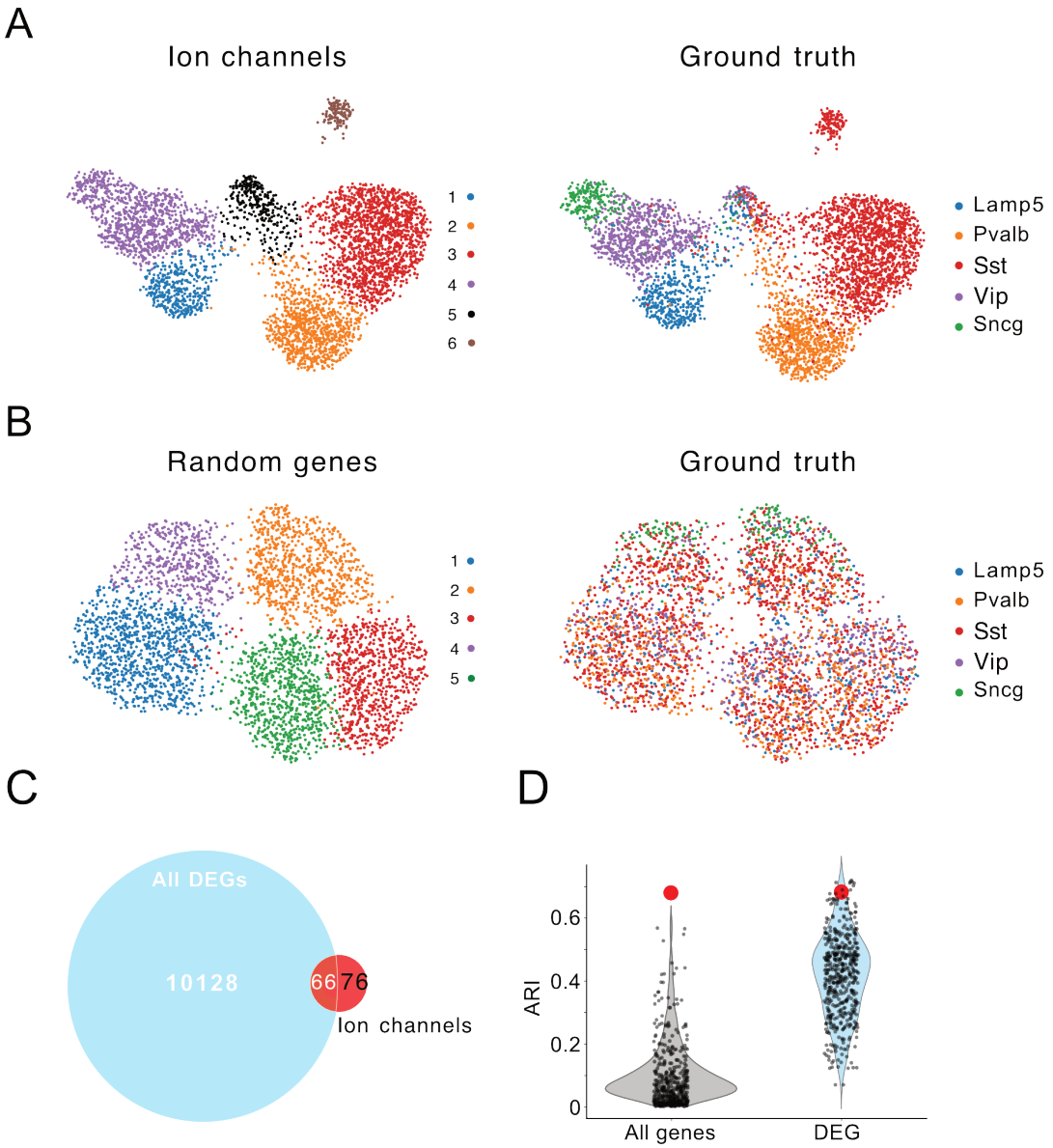

Fig. 4. Transcriptomic clustering of interneurons using Ion channel genes vs. random gene sets and differential expression analysis. **A.** UMAP representation of IN transcriptomic clustering based on ion channel-encoding genes. Each point represents single cell. Right. Clusters identified by ion channel gene expression, each cell is labeled with a unique identifier. Left. Same UMAP, with clusters labeled according to ground truth cell types for comparison. **B.** UMAP representation of IN transcriptomic clustering based on a randomly selected set of genes. Right. Each cell is labeled according to the results of clustering algorithm. Left. Same UMAP with ground truth labels. **C.** Venn diagram showing the overlap of DEGs identified through ground truth clustering and ion channel encoding genes. **D.** Distribution of the ARI values for clustering based on randomly selected genes. Grey distribution represents ARI values for clustering with random genes sampled from the entire gene pool, while blue distribution represents ARI values for clustering with random genes from the DEGs. Each dot corresponds to a single random sampling of the corresponding condition. The ARI value for clustering based on ion channel-encoding genes is marked by a red dot.

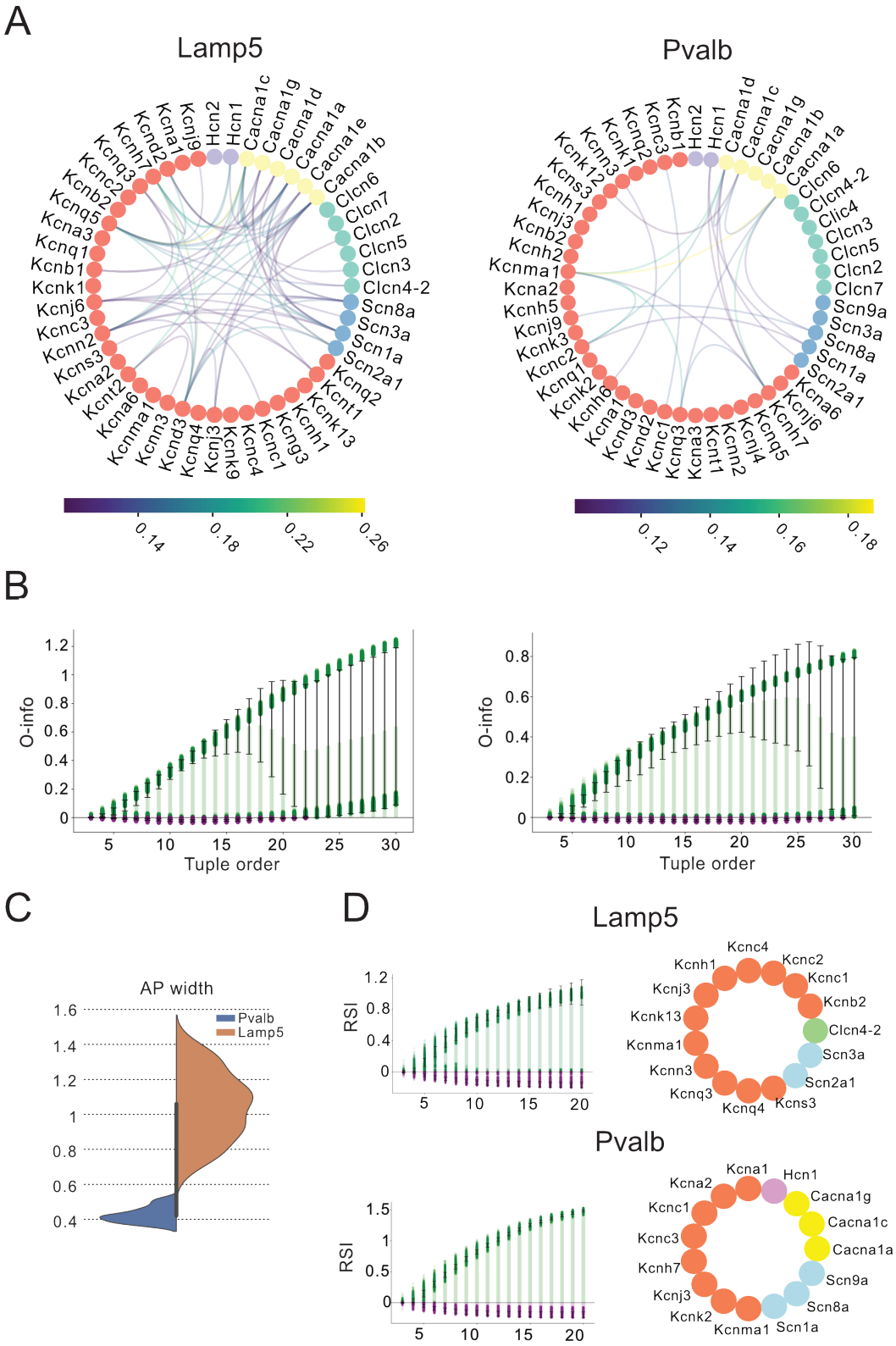

Fig. 5. High order interactions in the Ion Gene expression data. **A**. map of 50 ion channel-encoding genes. Node color represents the general current type: potassium (coral), sodium (blue), chloride (green), calcium (yellow), and HCN (purple) families. The color bar indicates the MI values. Right: MI map for Lamp5 ion channel gene expression. Left. Same for Pvalb. **B**.O-info values plotted as a function of tuple order for Lamp5 (right) and Pvalb (left) neurons. **C**.Distribution of action potential (AP) widths (in milliseconds) for cells from Lamp5 and Pvalb neuron types. **D**. RSI values as a function of tuple number (right) and corresponding set of the genes that result in highest value of RSI (left) for Lamp5 (top) and Pvalb (bottom) neurons.

rheobase feature and similarly found implication of distinct co-expression modules (see Sup. Figure 4 and 5).

These findings suggest that, although both cell types exhibit redundant high-order gene interactions, the specific ion channel genes contributing to these interactions differ between Lamp5 and Pvalb, reflecting their distinct electrophysiological properties. Moreover, while in silico populations displayed mostly synergistic interactions between ion channels, biological populations displayed strong redundant interactions.

Discussion

This study explores the intricate connection between biophysical parameters or gene expression and electrophysiological phenotypes in neurons. By employing biophysical models, statistical methods, and information theory, we aimed to decipher the complex interplay of ionic currents in shaping neuronal behavior and linked these insights to underlying transcriptomic profiles. Our approach involved the creation of numerous electrical models exhibiting diverse firing patterns using Monte Carlo Markov Chain methods. This enabled us to systematically investigate the relationship between model parameters, representing various ionic currents, and the resulting electrophysiological features of the simulated neurons. Our findings demonstrate that neuronal features are rarely determined by a single ionic current but often arise from the complex interactions of multiple ion channels. This highlights the importance of considering the integrated effects of ion channel activity when investigating neuronal phenotypes, which can only be identified using truly multivariate mathematical tools (O-info, RSI for instance). Furthermore, our analysis of single-cell RNA sequencing data uncovered gene expression modules specifically associated with certain interneuron types, supporting the link between transcriptomic profiles and distinct electrophysiological characteristics. The observation of high-order interactions among ion channel genes, particularly the prevalence of redundant interactions, suggests a level of robustness and functional compensation within neuronal systems. The dominance of potassium-encoding genes in tuples exhibiting high redundancy underscores the crucial role of potassium currents in maintaining neuronal stability and regulating excitability. Notably, the specific genes driving maximal redundancy differed between Lamp5 and Pvalb interneurons, reflecting their distinct electrophysiological properties. This difference was particularly evident in the action potential width, a key distinguishing feature between these two interneuron types. Pvalb neurons are characterized by significantly shorter AP widths and lower rheobase values compared to Lamp5 neurons, aligning with their fast-spiking capabilities and higher excitability (23, 24). The AP width influences the timing precision and frequency of neuronal firing, affecting synaptic transmission and the synchronization of neural networks (25). A narrower AP width in Pvalb neurons allows for rapid and high-frequency firing, which is essential for their role in controlling the timing of cortical circuitry (23). The shorter action potential width in Pvalb neurons, indicative of their fast-spiking nature, can be attributed to the specific combination of ion channels contributing to the maximal redundancy in this cell type. It is crucial to note that mRNA expression does not directly translate into protein expression (26). Future studies incorporating proteomic data would provide a more direct link between gene expression and functional ion channels.

The synergistic nature of the models and the redundancy observed in gene expression patterns raise interesting questions about the evolutionary advantages of these arrangements. Redundant elements in gene expression may provide robustness to the system, allowing for functional compensation if one component fails. In E.coli it was shown that the presence of multiple rRNA/tRNA gene copies can be advantageous under conditions that promote rapid growth and translation (27). This redundancy allows organisms to adapt more quickly to favorable environmental conditions (27). At the same time, redundancy offers significant advantages for both resilience and evolution. It has been shown that biological redundancy in protein-protein interaction networks enables robustness and increased tolerance for perturbations, enhancing their ability to maintain functionality despite genetic mutations or environmental stresses (28, 29). Conversely, synergetic interactions in the model might offer more efficient information processing but at the cost of increased vulnerability if a critical component is disrupted.

While our study demonstrates the significant predictive power of ion channel genes in distinguishing interneuron types, it is important to acknowledge that a comprehensive understanding of neuronal diversity requires consideration of synaptic genes and other groups of genes. Recent research has highlighted the strong predictive power of synaptic and GPCR genes in defining neuronal subtypes and their electrical properties (3).

This research contributes to the growing body of evidence linking mRNA expression to observable neuronal phenotypes (1, 7, 30). By integrating computational modelling, statistical analysis and transcriptomic data, we provide a framework for understanding the complex relationship between gene expression and the electrophysiological properties that define neuronal function. This approach holds significant potential for future studies investigating the molecular basis of neuronal diversity and the pathogenesis of neurological disorders arising from disrupted neuronal signalling.

Funding Acknowledgement. This study was supported by funding to the Blue Brain Project, a research center of the École polytechnique fédérale de Lausanne (EPFL), from the Swiss government's ETH Board of the Swiss Federal Institutes of Technology.

Supporting Information Appendix (SI).

Subsection for Method.

Detailed biophysical models . Detailed biophysical model was constructed based on the electrophysiological recordings of L5PC, cAC and bNAC recordings as previously described and available online (17). As in (17, 31) for each cell, a series of stimuli were administered: (1) IDrest, consisting of depolarizing steps with a sampling frequency of 10 kHz and duration of 2 seconds; (2) IDthresh, also with depolarizing steps at a 10 kHz sampling frequency and 2-second duration; (3) APWaveform, with depolarizing steps at a higher sampling frequency of 50 kHz for 50 ms; (4) IV, a sequence of current steps ranging from hyperpolarization to depolarization, sampled at 10 kHz over 3 seconds. A holding current was applied to maintain the membrane potential at -70 mV (prior to a liquid junction potential correction of 14 mV). In total, the features extracted from the recordings include voltage measurements, such as voltage after stimulus, baseline voltage, maximum voltage relative to baseline, and voltage deflections during stimulus onset and offset. Action potential (AP) properties, including AP amplitude, specific amplitudes (AP1, AP2, and last AP), AP half-width, and after-hyperpolarization (AHP) depth. Firing characteristics, such as mean firing frequency, total spike count, burst number, and ISI coefficient of variation. Timing-related features, such as time to first and last spikes and the inverse timing of successive ISIs, further describe the temporal aspects of neuronal firing. Additionally, resistance and current measurements (input resistance, holding current, and threshold current) and the decay time constant following stimulus were assessed to characterize the cell's intrinsic properties. The full description of the feature can be found at https://efel.readthedocs.io/en/latest/eFeatures.html.. These features were extracted for step protocols, the intensity of the step protocols reflects the fraction of the rheobase current (current necessary to elicit a single spike). The full list of protocols and features is can be found in Table 1.

The neuronal models were implemented in NEURON simulator (32). We used a set of ionic mechanisms described previously, the table with mechanisms specific for each electrical type can be found in Table 2.

Morphologies were manually reconstructed. The axons and their branches were replaced by a synthetic axon section consisting of an AIS (60 um) followed by a myelinated axon segment of 1,000 um.

To optimize these models, we employed an MCMC procedure as described in (18), with the cost computed as:

$$C(\mathbf{p}) = \max_i \frac{|f_i(\mathbf{p}) - f_{\text{exp},i}|}{\sigma(f_{\text{exp},i})} = \max_i z_i(\mathbf{p}), \qquad [1]$$

where z_i are the absolute z-scores for each feature i and simulated feature values are represented as a vector $\mathbf{f} = (f_0, \ldots, f_k)$, and the parameters are denoted as $\mathbf{p} = (p_0, \ldots, p_n)$.

From the resulting parameter sets, we selected models with an optimization cost within 3 standard deviations for cAC and bNAC, yielding 63,408 cAC models and 86,361 bNAC models. For L5PC, we selected models with an optimization cost within 5 standard deviations, resulting in 110,128 L5PC models.

Information theory measures and computations. Given a set of random variables $\mathbf{X} = \{X_1, X_2, \ldots, X_N\}$, the O-info is computed as (19):

$$O(\mathbf{X}) = (N-2)H(\mathbf{X}) - \sum_{i=1}^{N} H(X_i) + \sum_{i=1}^{N} H(\mathbf{X}_{\setminus i}),$$

where:

- $H(\mathbf{X})$ is the joint entropy of all variables,
- $H(X_i)$ is the marginal entropy of the i-th variable,
- $H(\mathbf{X}_{\setminus i})$ is the joint entropy of all variables except X_i.

The RSI, as defined by Chechik et al. (2001) (20), measures the net redundancy or synergy provided by a set of variables $S = \{X_1, X_2, \ldots, X_N\}$ about a target variable Y. It is given by:

$$\text{RSI}(S;Y) = I(S;Y) - \sum_{X_i \in S} I(X_i;Y),$$

where:

- $I(S;Y)$ is the mutual information between the set S and the target Y,
- $I(X_i;Y)$ is the marginal mutual information between each individual variable $X_i \in S$ and Y.

In the current work, we used Gaussian approximation of O-info and RSI, as it simplifies the computation of O-info and RSI by assuming the variables follow a multivariate Gaussian distribution. This assumption allows leveraging the covariance matrix Σ to compute entropy and mutual information analytically.

Gaussian Approximation for O-info. For a multivariate Gaussian system with covariance matrix Σ, the O-info is approximated as:

$$O(\mathbf{X}) = \frac{1}{2}\log\left(\frac{\prod_{i=1}^{N}\sigma_i^2}{|\Sigma|}\right) - \frac{N-1}{2}H,$$

where:

- σ_i^2 is the variance of the i-th variable (diagonal of Σ),
- $|\Sigma|$ is the determinant of the covariance matrix,
- $H = \frac{1}{2}\log((2\pi e)^N|\Sigma|)$ is the Gaussian entropy.

Gaussian Approximation for RSI. For RSI under Gaussian assumptions, the mutual information terms $I(S;Y)$ and $I(X_i;Y)$ are calculated directly from the covariance matrices:

$$I(S;Y) = \frac{1}{2}\log\left(\frac{|\Sigma_S|}{|\Sigma_{S|Y}|}\right),$$

$$I(X_i;Y) = \frac{1}{2}\log\left(\frac{\sigma^2_{X_i}}{\sigma^2_{X_i|Y}}\right).$$

Substituting into the RSI formula:

$$\text{RSI}_{\text{Gaussian}}(S;Y) = \frac{1}{2}\log\left(\frac{|\Sigma_S|\prod_{X_i\in S}\sigma^2_{X_i|Y}}{|\Sigma_{S|Y}|\prod_{X_i\in S}\sigma^2_{X_i}}\right).$$

E-type: bNAC, cNAC
Protocol: IDThresh/IDrest 150, 200, 250, 300 %   E-features: voltage_base, voltage_after_stim, AP_amplitude, APlast_amp, AHP_depth, inv_time_to_first_spike, time_to_last_spike, inv_first_ISI, inv_second_ISI, inv_third_ISI, inv_fourth_ISI, inv_fifth_ISI, mean_frequency
Protocol: APWaveform 360 %   E-features: AP_amplitude, AP1_amp, AP_duration_half_width, AHP_depth
Protocol: IV -100 %   E-features: voltage_deflection, voltage_deflection_begin
Protocol: IV -20 % **(Rin)**   E-features: ohmic_input_resistance_vb_ssse, voltage_base
Protocol: IV 0 % **(RMP)**   E-features: voltage_base, Spikecount
Protocol: RinHoldCurrent   E-features: bpo_holding_current
Protocol: Threshold   E-features: bpo_threshold_current
E-type: cAC
Protocol: IDThresh/IDrest 140, 200, 250, 300 %   E-features: voltage_base, voltage_after_stim, AP_amplitude, APlast_amp, AHP_depth, inv_time_to_first_spike, time_to_last_spike, inv_first_ISI, inv_second_ISI, inv_third_ISI, inv_fourth_ISI, inv_fifth_ISI, mean_frequency
Protocol: APWaveform 360 %   E-features: AP_amplitude, AP1_amp, AP_duration_half_width, AHP_depth
Protocol: IV -100 %   E-features: voltage_deflection, voltage_deflection_begin
Protocol: IV -20 % **(Rin)**   E-features: ohmic_input_resistance_vb_ssse, voltage_base
Protocol: IV 0 % **(RMP)**   E-features: voltage_base, Spikecount
Protocol: RinHoldCurrent   E-features: bpo_holding_current
Protocol: Threshold   E-features: bpo_threshold_current
E-type: cADPYR
Protocol: APWaveform 320 %   E-features: AP_amplitude, AP1_amp, AP2_amp, AP_duration_half_width, AHP_depth
Protocol: IV -100 %   E-features: voltage_deflection, voltage_deflection_begin
Protocol: IDrest & IDthresh 150, 200, 280 %   E-features: voltage_base, voltage_after_stim, AP_amplitude, APlast_amp, AHP_depth, inv_time_to_first_spike, time_to_last_spike, inv_first_ISI, inv_second_ISI, inv_third_ISI, inv_fourth_ISI, inv_fifth_ISI, mean_frequency
Protocol: SpikeRec 600 %   E-features: decay_time_constant_after_stim, voltage_after_stim, Spikecount
Protocol: IV -20 % **(Rin)**   E-features: ohmic_input_resistance_vb_ssse, voltage_base
Protocol: IV 0 % **(RMP)**   E-features: voltage_base, Spikecount
Protocol: RinHoldCurrent   E-features: bpo_holding_current
Protocol: Threshold   E-features: bpo_threshold_current

Table 1. Relevant protocols and e-features for bNAC, cAC, and cADPYR e-types in a single column format.

Detection of Stabilization Points. To identify the stabilization point in the O-info (or RSI) value series, we employed a derivative-based algorithm. This approach utilizes the first and second derivatives of the RSI curve to detect the point where the data stabilizes, as follows:

- **Data Extraction:** For each tuple, the maximum RSI value is extracted, forming a series of values $y = \{y_1, y_2, \ldots, y_n\}$ corresponding to tuple indices $x = \{1, 2, \ldots, n\}$.

- **Derivative Calculation:** The first derivative $y'(x)$ is computed as the rate of change between consecutive RSI values:

$$y'(x) = \frac{\Delta y}{\Delta x},$$

E-type: cADPYR
Soma
CaDynamics, HVA Ca, LVA Ca, Kv3.1, Ca-activated K, Persistent K, Transient K, Transient Na
AIS
CaDynamics, HVA Ca, LVA Ca, Kv3.1, Ca-activated K, Persistent K, Transient K, Transient Na, Persistent Na
Dendrites
CaDynamics, HVA Ca, LVA Ca, Kv3.1 (Apical), Transient Na (Apical; Decaying), HCN (Apical and Basal; Exponential increasing towards terminals)
E-types: bNAC, cAC
Soma
CaDynamics, HVA Ca, LVA Ca, Ca-activated K, Transient Na, Kv3.1, Persistent K, Transient K, HCN
AIS
CaDynamics, HVA Ca, LVA Ca, Ca-activated K, Transient Na, Kv3.1, Persistent K, Transient K
Dendrites
CaDynamics, HVA Ca, LVA Ca, Ca-activated K, HCN (Exponential increasing towards terminals)

Table 2. Active parameters and their compartmental placement for cAC, bNAC, and cADPYR e-types.

where $\Delta y = y_{i+1} - y_i$ and $\Delta x = x_{i+1} - x_i$. Subsequently, the second derivative $y''(x)$ is calculated to capture the rate of change in the slope:

$$y''(x) = \frac{\Delta y'(x)}{\Delta x}.$$

- **Stabilization Point Detection:** The stabilization point is identified as the index where the second derivative $y''(x)$ attains its most negative value, corresponding to the steepest concave point in the RSI curve. Mathematically:

$$x_{\text{stab}} = \arg\min y''(x).$$

This point marks the transition from steep decline to stabilization in the curve.

Clustering and analysis of the single cell mRNA from Patch-seq data. Data preprocessing and clustering were performed on single-cell RNA sequencing data containing expression profiles for thousands of genes across cells (12). Ground truth cell type labels were used to evaluate clustering performance. The analyses and clustering were implemented using Python libraries, scanpy (33) was used for bio-informatics analysis (preprocessing and clustering workflows).

Gene subsets were selected as follows:

- Ion Channel Subset: A predefined list of ion channel genes was used to extract a subset of the expression matrix.
- Differentially Expressed Gene (DEG) Subset: DEGs were identified using statistical tests (e.g., t-tests) to compare expression levels across conditions.
- Random Subsets: Control subsets were generated by selecting 1,000 random gene sets, each matching the size of the ion channel subset.

Normalization was applied to ensure consistency in expression data:

- Log-transformation ($log_2(counts + 1)$) was used to normalize skewed distributions.
- Standardization was performed to scale each gene's expression to have zero mean and unit variance.

Dimensionality reduction was conducted using Principal Component Analysis (PCA) to reduce data dimensionality while retaining features explaining most of the variance (e.g., 90%).

Clustering was performed using the Leiden algorithm, which is optimized for detecting communities in graphs derived from high-dimensional data. Key steps included:

- Graph Construction: A k-nearest neighbor (kNN) graph was constructed from the principal components, with edges indicating cell similarities.
- Community Detection: The Leiden algorithm partitioned cells into clusters, with the number of clusters determined adaptively from the graph structure.
- Evaluation: Clustering results were compared to ground truth cell type labels using the Adjusted Rand Index (ARI). ARI measures label similarity, correcting for chance, with a score of 1 indicating perfect agreement.

Clustering was used to assess the specificity of ion channel, randomly selected and DEG subsets. The same preprocessing and clustering steps were repeated for 1,000 random gene subsets. ARI scores were computed for each run and compared to the ground truth labels. The resulting distribution of ARI scores served as a baseline for comparison.

1. Y Bernaerts, et al., Combined statistical-mechanistic modeling links ion channel genes to physiology of cortical neuron types. *bioRxiv* pp. 2023–03 (2023).
2. J Eberwine, J Kim, Cellular deconstruction: finding meaning in individual cell variation. *Trends cell biology* **25**, 569–578 (2015).
3. RM Kendrick, S Linderman, SF Owen, Transcriptomically-measured gene expression predicts physiological variation across single neurons in humans and mice. *bioRxiv* pp. 2024–08 (2024).
4. AM Swensen, BP Bean, Robustness of burst firing in dissociated purkinje neurons with acute or long-term reductions in sodium conductance. *J. Neurosci.* **25**, 3509–3520 (2005).
5. DJ Schulz, JM Goaillard, E Marder, Variable channel expression in identified single and electrically coupled neurons in different animals. *Nat. neuroscience* **9**, 356–362 (2006).
6. T Kodama, et al., Graded coexpression of ion channel, neurofilament, and synaptic genes in fast-spiking vestibular nucleus neurons. *J. Neurosci.* **40**, 496–508 (2020).
7. M Tapia, et al., Neurotransmitter identity and electrophysiological phenotype are genetically coupled in midbrain dopaminergic neurons. *Sci. reports* **8**, 13637 (2018).
8. DJ Schulz, JM Goaillard, EE Marder, Quantitative expression profiling of identified neurons reveals cell-specific constraints on highly variable levels of gene expression. *Proc. Natl. Acad. Sci.* **104**, 13187–13191 (2007).
9. JM Goaillard, AL Taylor, DJ Schulz, E Marder, Functional consequences of animal-to-animal variation in circuit parameters. *Nat. neuroscience* **12**, 1424–1430 (2009).
10. M Toledo-Rodriguez, et al., Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. *Cereb. cortex* **14**,

1310–1327 (2004).
11. F Buettner, et al., Computational analysis of cell-to-cell heterogeneity in single-cell rna-sequencing data reveals hidden subpopulations of cells. *Nat. biotechnology* **33**, 155–160 (2015).
12. NW Gouwens, et al., Integrated morphoelectric and transcriptomic classification of cortical gabaergic cells. *Cell* **183**, 935–953 (2020).
13. F Scala, et al., Phenotypic variation of transcriptomic cell types in mouse motor cortex. *Nature* **598**, 144–150 (2021).
14. TE Chan, MP Stumpf, AC Babtie, Gene regulatory network inference from single-cell data using multivariate information measures. *Cell systems* **5**, 251–267 (2017).
15. J Wildenhain, et al., Prediction of synergism from chemical-genetic interactions by machine learning. *Cell Syst.* **1**, 383–395 (2015).
16. D Anastassiou, Computational analysis of the synergy among multiple interacting genes. *Mol. systems biology* **3**, 83 (2007).
17. M Reva, et al., A universal workflow for creation, validation, and generalization of detailed neuronal models. *Patterns* **4** (2023).
18. A Arnaudon, et al., Controlling morpho-electrophysiological variability of neurons with detailed biophysical models. *Iscience* **26** (2023).
19. FE Rosas, PA Mediano, M Gastpar, HJ Jensen, Quantifying high-order interdependencies via multivariate extensions of the mutual information. *Phys. Rev. E* **100**, 032305 (2019).
20. G Chechik, et al., Group redundancy measures reveal redundancy reduction in the auditory pathway. *Adv. neural information processing systems* **14** (2001).
21. W Hu, BP Bean, Differential control of axonal and somatic resting potential by voltage-dependent conductances in cortical layer 5 pyramidal neurons. *Neuron* **97**, 1315–1326 (2018).
22. H Markram, et al., Interneurons of the neocortical inhibitory system. *Nat. reviews neuroscience* **5**, 793–807 (2004).
23. H Hu, J Gan, P Jonas, Fast-spiking, parvalbumin+ gabaergic interneurons: From cellular design to microcircuit function. *Science* **345**, 1255263 (2014).
24. R Tremblay, S Lee, B Rudy, Gabaergic interneurons in the neocortex: from cellular properties to circuits. *Neuron* **91**, 260–292 (2016).
25. BP Bean, The action potential in mammalian central neurons. *Nat. Rev. Neurosci.* **8**, 451–465 (2007).
26. J Reimegård, et al., A combined approach for single-cell mrna and intracellular protein expression analysis. *Commun. biology* **4**, 624 (2021).
27. PK Raval, WY Ngan, J Gallie, D Agashe, The layered costs and benefits of translational redundancy. *Elife* **12**, e81005 (2023).
28. B Klein, et al., A computational exploration of resilience and evolvability of protein–protein interaction networks. *Commun. Biol.* **4**, 1352 (2021).
29. P Hunter, Understanding redundancy and resilience: Redundancy in life is provided by distributing functions across networks rather than back-up systems. *EMBO reports* **23**, e54742 (2022).
30. C Bomkamp, et al., Transcriptomic correlates of electrophysiological and morphological diversity within and across excitatory and inhibitory neuron classes. *PLoS computational biology* **15**, e1007113 (2019).
31. H Markram, et al., Reconstruction and simulation of neocortical microcircuitry. *Cell* **163**, 456–492 (2015).
32. NT Carnevale, ML Hines, *The NEURON book.* (Cambridge University Press), (2006).
33. FA Wolf, P Angerer, FJ Theis, Scanpy: large-scale single-cell gene expression data analysis. *Genome biology* **19**, 1–5 (2018).

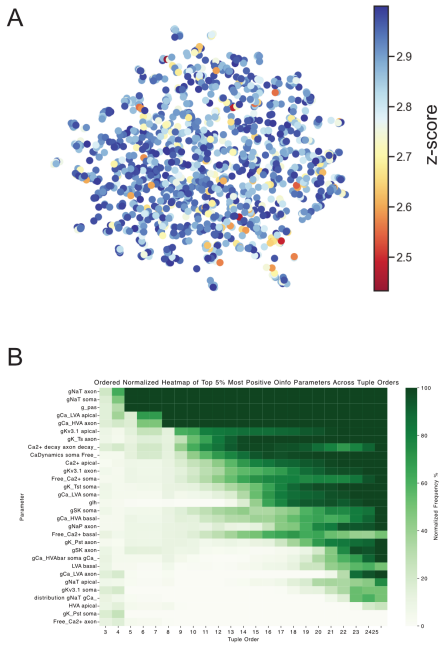

Figure 1: A. UMAP representation of model parameters. The colors indicate the fitness of the model, quantified as the z-score for the corresponding parameter set. B. Normalized frequency map highlighting parameters that frequently occur in high-redundancy tuples.

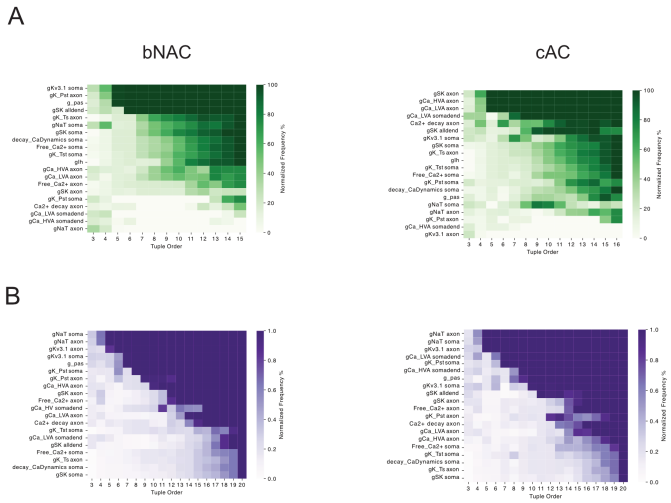

Figure 2: A. Normalized frequency map showing parameters that frequently appear in high-redundancy tuples based on Oinfo calculations for cAC and bNAC neurons. B. Normalized frequency map for parameters in high-synergy tuples, similar to panel A.

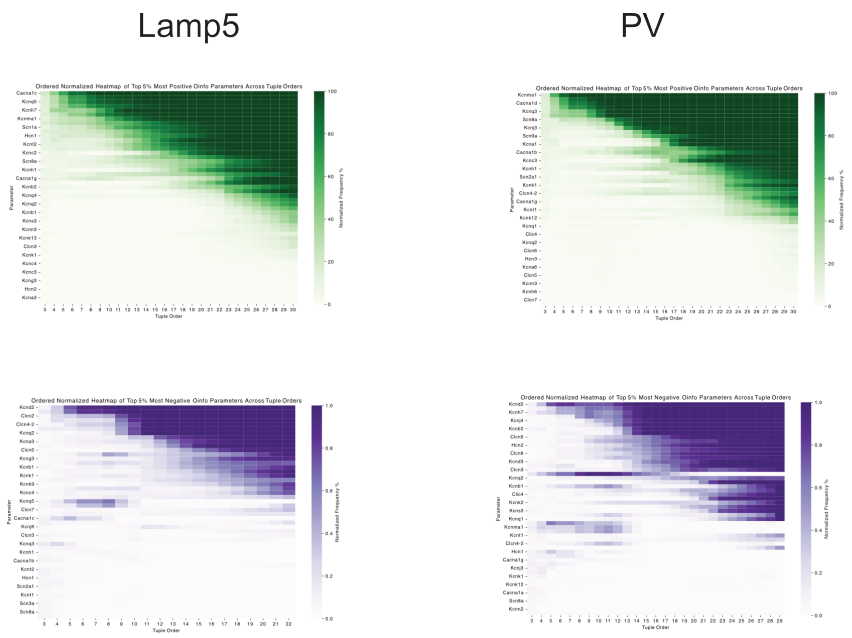

Figure 3: Normalized frequency maps of parameters frequently appearing in high-redundancy (top) and high-synergy (bottom) tuples for Lamp5 and Pvalb neurons.

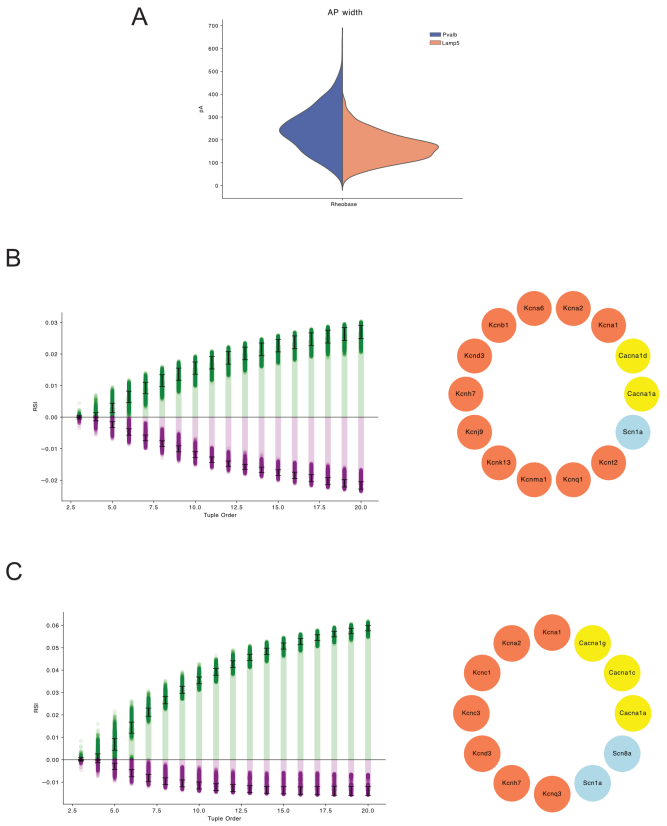

Figure 4: A. Distribution of mean frequencies (Hz) for cells from Lamp5 and Pvalb neuronal types. B. RSI values plotted against tuple numbers (right), along with the set of genes that yield the highest RSI value (left) for Lamp5 neurons. C. Similar RSI values and gene sets as in panel B, but for Pvalb neurons.

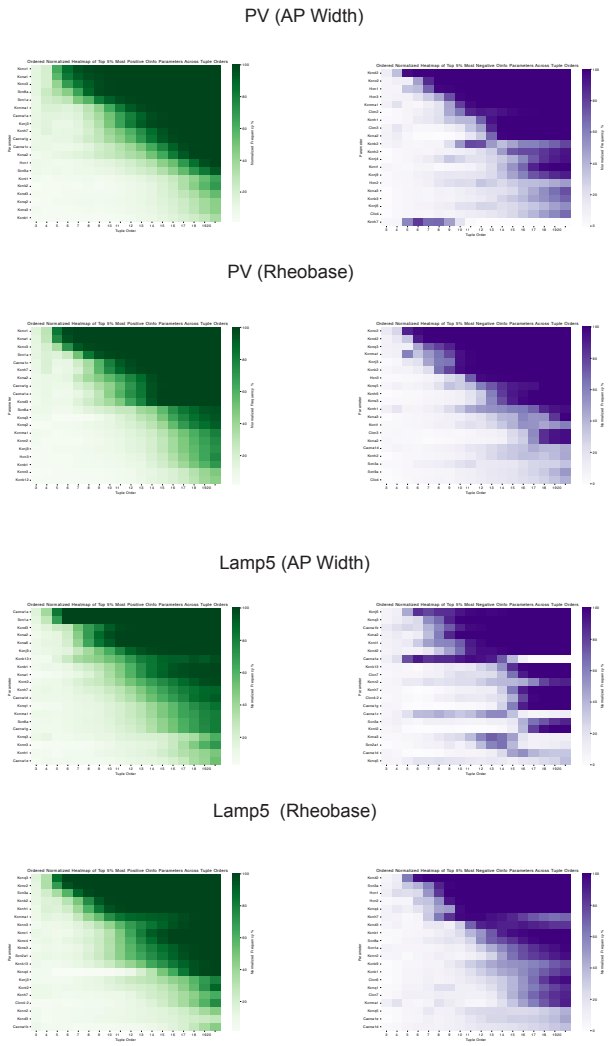

Figure 5: Normalized frequency maps of parameters frequently appearing in high-redundancy (left) and high-synergy (right) tuples. Data correspond to gene expression, action potential width (AP width), and mean frequency features for Lamp5 and Pvalb neurons.