A mechanistic model of genetic effects on excitation-inhibition balance in a RNN model of auditory processing

Rebeca Ianov Vitanov (Rebeca.Ianov-Vitanov@mrc-cbu.cam.ac.uk)
MRC Cognition and Brain Sciences Unit, University of Cambridge, 15 Chaucer Road
Cambridge CB2 7EF, United Kingdom

Kate Baker (Kate.Baker@mrc-cbu.cam.ac.uk)
MRC Cognition and Brain Sciences Unit, University of Cambridge, 15 Chaucer Road
Cambridge CB2 7EF, United Kingdom
Department of Medical Genetics, University of Cambridge, Addenbrooke's Treatment Centre, Cambridge Biomedical Campus
Cambridge CB2 0QQ, United Kingdom

Duncan Astle (Duncan.Astle@mrc-cbu.cam.ac.uk)
MRC Cognition and Brain Sciences Unit, University of Cambridge, 15 Chaucer Road
Cambridge CB2 7EF, United Kingdom
Department of Psychiatry, University of Cambridge, Forvie Site, Robinson Way
Cambridge CB2 0SZ, United Kingdom

Jascha Achterberg (Jascha.Achterberg@mrc-cbu.cam.ac.uk)
MRC Cognition and Brain Sciences Unit, University of Cambridge, 15 Chaucer Road
Cambridge CB2 7EF, United Kingdom
Intel Labs
Santa Clara, CA 95054, United States

Danyal Akarca (Danyal.Akarca@mrc-cbu.cam.ac.uk)
MRC Cognition and Brain Sciences Unit, University of Cambridge, 15 Chaucer Road
Cambridge CB2 7EF, United Kingdom

Abstract:

Neural networks offer a powerful tool for testing mechanistic hypotheses about cognition. We explored the utility of this approach in a monogenic developmental disorder, by integrating a recurrent neural network (RNN) model and MEG task data from individuals with *ZDHHC9*-associated intellectual disability and age-matched controls. Given experimental evidence of *ZDHHC9* implication in inhibitory synapse formation, we tested whether reducing inhibition levels in a RNN model of auditory processing trained on neurotypical evoked responses recapitulates case group neurophysiology. We show that stronger reductions in recurrent, inhibitory weights resulted in increased peak amplitude and peak latency of RNN prediction relative to the pre-perturbation predictions, similar to case-control empirical trends. In contrast, increasing network excitation via the excitatory weights failed to consistently recapitulate these trends. Together, these results suggest that reduced synaptic inhibition is a plausible mechanism by which loss of *ZDHHC9* function alters cortical dynamics during sensory processing.

Keywords: recurrent neural network; monogenic disorders; ZDHHC9; intellectual disability, epilepsy, language, MEG.

Introduction

Combining the study of rare single gene disorders related to intellectual disability (ID) with neural network models of the brain can provide insights into specific mechanisms contributing to developmental cognitive difficulties. We trial this approach by employing a recurrent neural network (RNN) model of auditory processing as a tool for mapping genetically-driven, local alterations to systems-level activity, in a group of individuals with *ZDHHC9*-associated ID. The *ZDHHC9* gene encodes a palmitoylation enzyme, ZDHHC9, which promotes inhibitory synapse formation (Shimell et al., 2019). Loss-of-function variants have been associated with X-linked ID, rolandic epilepsy and language difficulties (Baker et al., 2015). In primary neuronal cultures, *ZDHHC9* variants lead to shorter and less complex dendritic arbours and an increase in the ratio of excitatory-to-inhibitory synapses (Shimell et al., 2019).

The aims of this study were two-fold: 1) to characterise neurophysiological differences between individuals with *ZDHHC9* variants and neurotypical controls and 2) to test the hypothesis that weaker network inhibition in a RNN model of auditory processing recapitulates these differences.

MEG data were recorded from participants in the case (N = 8) and control (N = 7) groups during a passive roving oddball task, to enable assessment of auditory change detection via MEG mismatch negativity (mMMN) (Cowan et al., 1993; Garrido et al., 2009; Kirihara et al., 2020; Näätänen et al., 2014). The empirical MEG data was then integrated with the RNN to test whether qualitative differences between the two groups in terms of auditory evoked field responses (AEFs) can be explained by *ZDHHC9*-associated synaptic alterations, i.e. reduced network inhibition.

Methods

Experimental task

The roving task involved the repeated presentation of standard stimuli of a particular frequency (250Hz, 500Hz, or 1000Hz) 3-12 times, followed by a frequency switch (the deviant, which, through repetition, became the new standard).

The model

We implement a discrete-time RNN with four hidden layers using the Tensorflow package (Abadi et al., 2016). Standard inputs (S) to the RNN were designed as three consecutive waveforms (250, 500 or 1000Hz); deviants (D) were represented by a change in frequency of the third waveform (Figure 1a). The RNN was trained in a supervised fashion, and the labels (targets) were empirically derived.

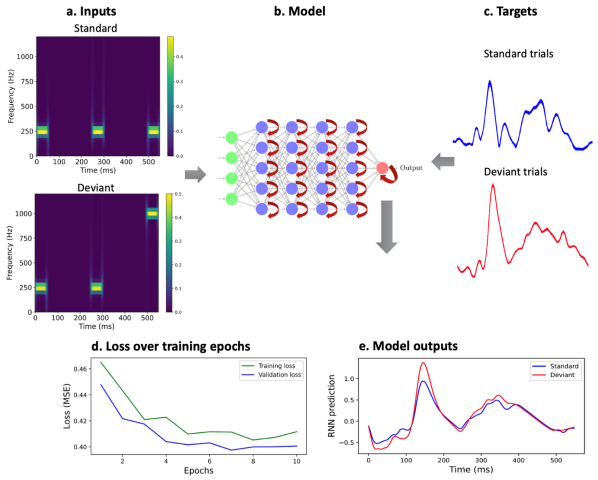


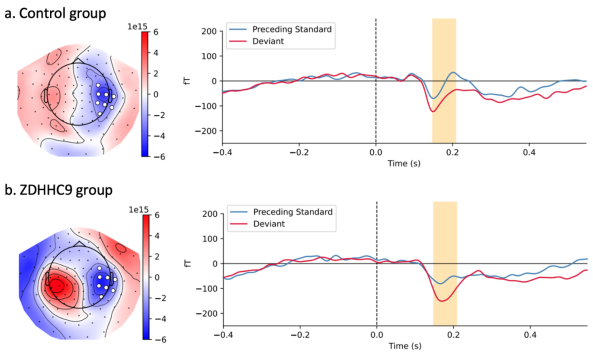
Figure 1. a. Spectrograms of an example standard (S) and deviant (D) input. b. Simplified diagram of the RNN architecture. Input layer (green) had 63 recurrent units, each hidden layer had 64 units and the output layer had 1 recurrent unit. c. Targets were simulated AEFs obtained by adding Gaussian white noise (s.d. = 0.6) to the control group-level post-stimulus AEF in response to S and D tones, respectively. d. The RNN was trained for 10 epochs. e. RNN predictions to S and D inputs.

The targets were robust-scaled and flipped so that most values are positive. The model was optimised to minimise mean-squared-error (MSE) loss between predictions and targets, using gradient descent and the Adam optimiser. Dropout regularisation with a rate of 0.15 was used in the hidden layers.

After training, the RNN was perturbed to test the impact of alterations mimicking the *ZDHHC9* loss-of-function phenotype, reduced synaptic inhibition, on the RNN output. We conducted perturbation experiments at the level of the recurrent connections in the hidden layers by weakening inhibitory connections relative to task-optimized values for 8 levels between 0.5%-4% and assessed the effects on the network's predictions. Two control experiments in which we either increased excitatory weights or concomitantly reduced and increased a random set of inhibitory and excitatory weights, respectively, for the same levels, were performed.

Results

We compared S and D trial responses between the control and ZDHHC9 groups (Figure 2). AEFs in the ZDHHC9 group showed increased amplitude, increased peak latency and increased mismatch negativity (MMN) compared to controls, suggesting stronger and slower AEF responses (Figure 2).



c.

	Peak amplitude (fT)		Peak latency (ms)		MMN
	S	D	S	D	
Control group	-70.09	-123.27	148	148	59.56
ZDHHC9 group	-80.85	-151.03	164	168	67.48
Ratio ZDHHC9:Control	1.15	1.23	1.1	1.14	1.13

Figure 2. a. Averaged AEF responses to all deviants (D) and their corresponding preceding standard (S) at the 8 sensors where significant S-D differences were found in both the control and ZDHHC9 groups. In yellow, the timeframe of significant differences is shown. a. Control group response (p-value = 0.0008). b. ZDHHC9 group response (p-value = 0.0015). c. The values from the plots in a. and b. (absolute values for peak amplitudes) and mismatch negativity (MMN) calculated as mean absolute error between standard-evoked responses and deviant-evoked responses in the significant time window.

Next, we compared RNN predictions before and after perturbations to the empirical trends. Inhibition reduction experiments resulted in predicted AEFs and MMN with linearly increasing amplitudes (Figure 3a) relative to baseline levels, which mirrors the trend observed empirically between the control and ZDHHC9 groups (Figure 2). AEF peak latencies also increased from baseline. Increased excitation resulted in exponential increases in AEF peak amplitude from 2% weight increases onwards (Figure 3b), accompanied by peak latencies in the latter half of the AEF window. Perturbing a random set of excitatory and inhibitory weights resulted in opposite polarity AEFs with peak amplitudes and MMN varying minimally across the perturbation levels, and a constant peak latency at 300ms (Figure 3c).

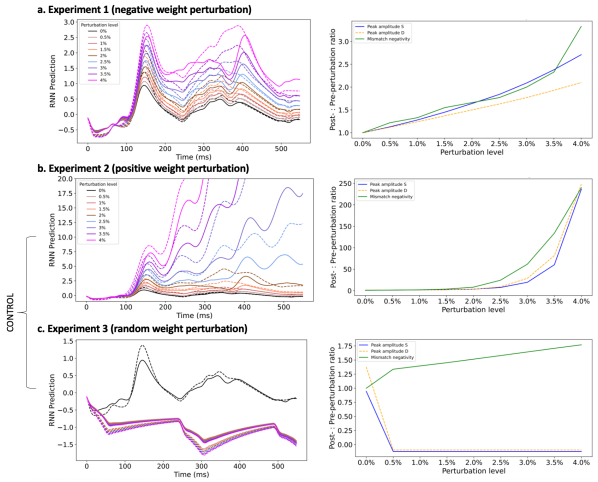


Figure 3. Predicted AEFs for standard (solid lines) and deviant inputs (dotted lines; left panel) and relative increases from initial RNN predictions (pre-perturbation) after each perturbation experiment (right panel). a. Experiment 1 (negative weight perturbation), b. Experiment 2 (positive weight perturbation) and c. Experiment 3 (random weight perturbation).

Conclusions

The current study serves as a proof-of-concept for using neural networks to investigate mechanistic origins of developmental cognitive disorders. Our results support a causal link between reduced cortical inhibition, increased amplitude and peak latency of electrophysiological responses in *ZDHHC9*-associated intellectual disability.

This work opens the door to the development of more complex neural network models of sensory processing to study a range of monogenic cognitive disorders. These will aid in the development of novel theories about genetic influences on cortical processing and offer a deeper, multi-level understanding of neurodevelopmental conditions.

Acknowledgments

R.I.V., K.B., D.Akarca, J.A. and D.Astle are supported by UKRI MRC funding. D.Akarca and D.Astle are funded by the James S. McDonnell Foundation Opportunity Award for Understanding Human Cognition. J.A. is funded by a Gates Cambridge Scholarship and Intel Labs. Data collection for this study was funded by the Wellcome Trust/Academy of Medical Sciences (Starter Grant for Clinical Lecturers to KB). This work was supported by UKRI/MRC University Unit Strategic Partnership Funding (programme awards MC_UU_00030/3; MC_UU_00030/2).

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. et al. (2016). Tensorflow: Large-scale machine learning on heterogenous distributed systems. Retrieved from: https://www.tensorflow.org/.

Baker, K., Astle, D. E., Scerif, G., Barnes, J., Smith, J., Moffat, G., Gillard, J., Baldeweg, T. and Raymond, F. L. (2015). Epilepsy, cognitive deficits and neuroanatomy in males with *ZDHHC9* mutations. *Annals of clinical and translational neurology* 2(5), 559–569.

Cowan, N., Winkler, I., Teder, W. and Näätänen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (erp). *Journal of experimental psychology: Learning, Memory, and Cognition* 19(4), 909.

Garrido, M. I., Kilner, J. M., Stephan, K. E. and Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clinical neurophysiology* 120(3), 453–463.

Kirihara, K., Tada, M., Koshiyama, D., Fujioka, M., Usui, K., Araki, T. and Kasai, K. (2020). A predictive coding perspective on mismatch negativity impairment in schizophrenia. *Frontiers in psychiatry* 11, 660.

Näätänen, R., S Sussman, E., Salisbury, D. and L Shafer, V. (2014). Mismatch negativity (MMN) as an index of cognitive dysfunction. *Brain topography* 27, 451–466.

Shimell, J. J., Shah, B. S., Cain, S. M., Thouta, S., Kuhlmann, N., Tatarnikov, I., Jovellar, D. B., Brigidi, G. S., Kass, J., Milnerwood, A. J. et al. (2019). The X-linked intellectual disability gene *ZDHHC9* is essential for dendrite outgrowth and inhibitory synapse formation. *Cell reports* 29(8), 2422–2437.