35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Decoding of the Sound Frequency from the Steady-state Neural Activities in Rat Auditory Cortex*

Tomoyo I. Shiramatsu, Takahiro Noda, Ryohei Kanzaki, and Hirokazu Takahashi

Abstract— **In the auditory cortex, onset activities have been extensively investigated as a cortical representation of sound information such as sound frequency. Yet, less attention has been paid to date to steady-state activities following the onset activities. In this study, we used machine learning to investigate whether steady-state activities in the presence of continuous sounds represent the sound frequency. Sparse Logistic Regression (SLR) decoded the sound frequency from band specific power or phase locking value (PLV) of local field potentials (LFP) from the fourth layer of the auditory cortex of anesthetized rats. Consequently, we found that SLR was able to decode the sound frequency from steady-state neural activities as well as onset activities. This result demonstrates that the steady-state activities contain information about the sound such as sound frequency.**

I. Introduction

Neural systems are usually sensitive to transient rather than steady-state stimulus features. Onset activities have thus been characterized extensively to date. For example, in the auditory cortex, most existing studies used short test tones with durations of a few 10 or 100 ms to characterize neural activation at stimulus onsets as a function of test frequency and intensity [1, 2]. On the other hand, steady-state neural activities following the onset responses have received less attention because the neural activities easily adapt to continuous sounds lasting on the order of seconds. However, some specific neural features such as cross-correlation of the firing rates between two neurons may carry significant information about test tones in steady-state neural activities [3]. We thus attempt to decode steady-state activities in the auditory cortex by paying specific attention to, interaction between neural populations.

Microelectrode arrays that are able to densely map neural activities and investigate interactions between neurons are now available in routine experiments. However, thus-obtained neural data become very high dimension, making it difficult for us to find reliable neural activity patterns representing sound information. Thus, automatic identification of informative patterns may bring significant benefit to decode high-dimension neural patterns.

*Resrach supported by JSPS KAKENHI Grant Number 23135507.

T. I. Shiramatsu is with Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, 153-8904 Japan. (corresponding author to provide phone: +81-3-5452-5197; fax: +81-3-5452-5197; e-mail: isoguchi@brain.imi.i.u-tokyo.ac.jp)

T. Noda and R. Kanzaki are with Research Center for Advanced Science and Technology, The Univertisy of Tokyo, Tokyo, 153-8904 Japan.

H. Takahashi is with Research Center for Advanced Science and Technology, The Univertisy of Tokyo, Tokyo, 153-8904 Japan and JST PRESTO, 332-0012 Japan.

Previous studies demonstrated that machine learning such as support vector machine (SVM) and k-nearest neighbor method (KNN) successfully decoded sound stimulus frequency from firing rates of onset activity in the rat auditory cortex [4, 5]. However, it is known when the sample size is small or data dimension is high, these methods will be over-trained and insufficient for decoding. In this study, to handle more high-dimension data, we utilize the sparse logistic regression (SLR [6]) to decode steady-state neural activity. In SLR, logistic regression is extended to Bayesian framework using automatic relevance determination (ARD). ARD is an effective algorithm to compress a dimension of input characteristics vector, by reducing weights of non-reliable component of the input vector to zero [7]. Therefore, SLR is likely useful when input vectors sparsely represent information like neural activity patterns.

The objective of this study is to investigate whether the steady-state activity in the auditory cortex represent the sound information. In particular, we attempt to decode the frequency of the sound stimuli from the steady-state activity in rat auditory cortex. For neural data to be decoded, a microelectrode array with a grid of 10 × 10 within 4 × 4-mm recording area simultaneously recorded local field potentials (LFP) from the 4th layer of the auditory cortex of anesthetized rats. The band-specific power and the phase synchrony were extracted as neural features of the steady-state activity for decoding.

II. Materials and Methods

This study was carried out in strict accordance with "Guiding Principles for the Care and Use of Animals in the Field of Physiological Science" by the Japanese Physiological Society. The experimental protocol was approved by the Committee on the Ethics of Animal Experiments at Research Center for Advanced Science and Technology, the University of Tokyo (Permit Number: RAC07110). All surgery was performed under isoflurane anesthesia, and every effort was made to minimize suffering.

A. Electrophysiological Experiment

Seven Wistar rats, at postnatal week 8 – 10, with a body weight of 230 – 300 g, were used. Rats were anesthetized with isoflurane (3 % at induction and 1-2% for maintenance) and were held in place with a custom-made head-holding device. Atropine sulfate (0.1 mg/kg) was administrated at the beginning and at the end of the surgery to reduce the viscosity of bronchial secretions. A heating blanket was used to maintain body temperature at around 37 degrees C. Skin incision at the beginning of surgery was made under local anesthesia of xylocaine (0.3 - 0.5 ml). A needle electrode was subcutaneously inserted at a right forepaw and used as a

ground. A small craniotomy was made near the bregma landmark to embed a 0.5-mm-thick integrated circuit socket as a reference electrode, with an electrical contact to the dura mater. The right temporal muscle, cranium, and dura overlying the auditory cortex were surgically removed, and the exposed cortical surface was perfused with saline in order to prevent desiccation. Cisternal cerebrospinal fluid drainage was performed to minimize cerebral edema. The right eardrum, i.e., ipsilaternal to the exposed cortex, was ruptured and waxed to ensure unilateral sound inputs from the ear contralateral to the exposed cortex. Respiratory rate, heart rate and hind-paw withdrawal reflexes were monitored throughout the experiment in order to maintain an adequate anesthetic level as stably as possible.

A microelectrode array (Blackrock Microsystems, ICS-96) with a grid of 10 × 10 recording sites within an area 4 × 4-mm simultaneously recorded LFPs from the 4th layer of the auditory cortex, i.e., 600 μm in depth. LFPs were obtained with an amplification gain of 1000, digital filter bandpass of 0.3 – 500 Hz, and sampling frequency of 1 kHz (Cyberkinetics Inc.; Cerebus Data Acquisition System). A speaker (10TH800, Matsushita Electric Industrial Co. Ltd., Japan) was positioned 10 cm from the left ear, i.e., contralateral to the exposed cortex. Test stimuli were calibrated at the pinna with a 1/4-inch microphone (Brüel&Kjær, 4939) and spectrum analyzer (Ono Sokki Co., Ltd., CF-5210). The stimulus level is presented in dB SPL (sound pressure level in decibels with respect to 20 μPa).

First, we recorded onset activities responding to pure tone bursts (5 ms rise/plateau/fall) with frequencies from 1 to 50 kHz and intensities from 30 to 70 dB SPL, repeated 20 times for each combination of frequency and intensity. These test frequencies cover the entire audibility range of rat. Then, we recorded steady-state activities responding to continuous pure tone (30 second duration, 60 dB SPL) with frequencies of 16 and 40 kHz. Each of the pure tone was repeated 10 times and interleaved with a silent block of 30 seconds.

B. Calculation of the Characteristics of Neural Activity

Three characteristics were extracted from the recorded LFPs, i.e., amplitude of evoked potential from the onset activities, and band-specific power and phase locking value (PLV) from the steady-state activities.

1) Characteristic of the Onset Activities

It is known that spatial pattern of the amplitude of evoked potential depends on the sound frequency, and machine learning can discriminate them [4, 5]. Thus, SLR first tried to discriminate sound frequency from the amplitudes of evoked potentials. Fig. 1 (a) shows the representative plot of the recorded LFP. In the 4th layer of the auditory cortex, auditory evoked middle-latency potentials were recorded as negative amplitudes. The absolute values of the minimum peak of LFPs before 50 ms from the sound onset (4, 16 and 40 kHz, 60 dB SPL) were defined as the characteristic of the onset activities.

2) Characteristics of the Steady-state Activities

Fig. 1 (b)(i) shows the representative plot of recorded LFP in response to a continuous sound. As this plot indicates, in the LFP under anesthesia, some spindles occurred regardless of

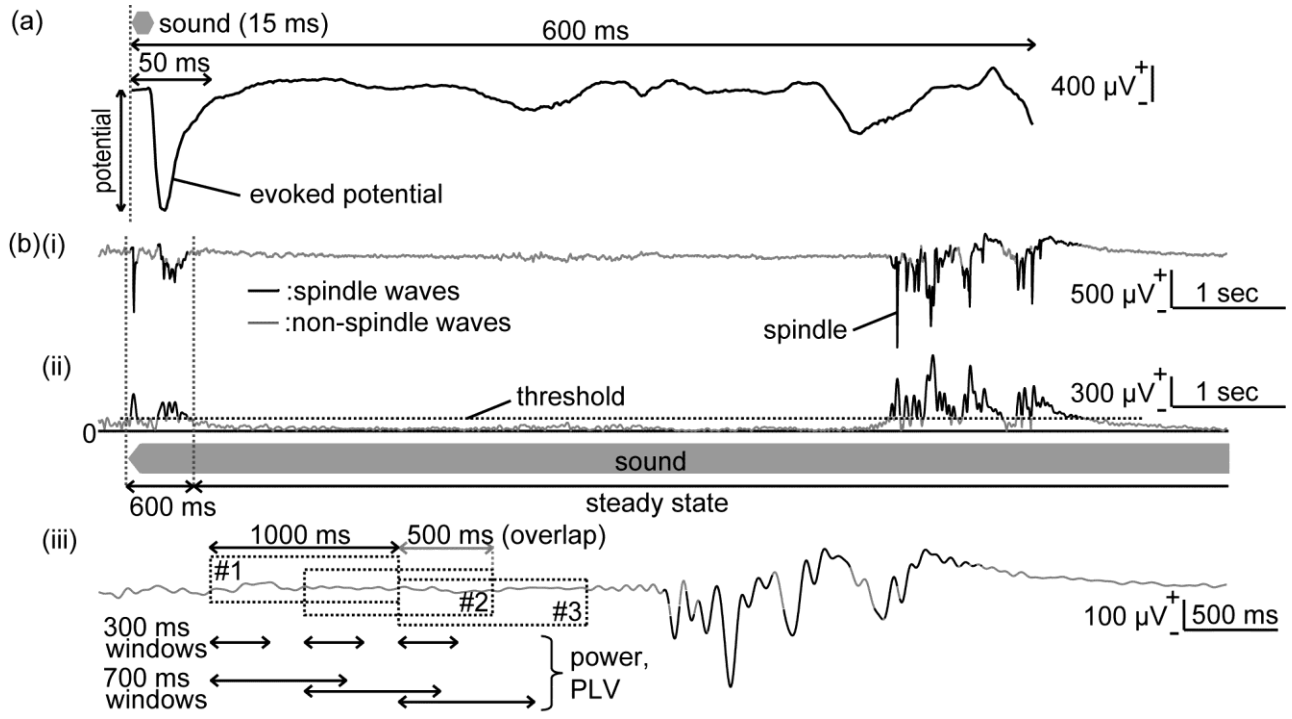

Figure 1. Tone-evoked local field potential (LFP)

(a) Onset LFP. The absolute value of the minimum peak before 50 ms from the sound onset was defined as the amplitude of evoked potential. The duration of tone was 15 ms. (b) Steady-state LFP. (i) Representative raw traces of LFP in response to a continuous sound (16 kHz, 60 dB SPL, 30 sec). (ii) Instantaneous amplitude of LFP after filtering with a passband of 11 – 16 Hz, where spindles were included. Time periods, during which the instantaneous amplitude exceeds a threshold in 25 and more recording sites, were classified as spindle waves (black line); others were classified as non -spindle waves (gray line) and the following analysis only included the non-spindle waves. (iii) Representative traces of LFP filtered by theta band (4 – 8 Hz). Black and gray lines represent the spindle and non-spindle waves. A length of time period was 1000 ms, and they had 500 ms overlap. A length of time window used in the analyses was chosen from 6 durations. These filtered LFP were used to quantify the band power and phase locking value (PLV) of steady-state neural activities.

the sound representation. Therefore, we eliminated the LFPs under spindle activities from analysis by the following way. First, LFPs were filtered with a passband of 11 – 16 Hz, where spindles were included, and instantaneous amplitude was extracted by Hilbert transform (Fig. 1 (b)(ii)). From the part of it, threshold to determine spindles was calculated as the summation of the average and three times the standard deviation of root mean square of the instantaneous amplitude. This threshold was calculated in each rat, then, it divided all LFPs into spindle and non-spindle waves.

For analysis, we cut time windows in 6 different lengths. First, we cut one-hundred 1000-ms time periods every 500 ms (Fig. 1 (b)(iii)) in each sound frequency. From these time periods, we obtain time windows with 6 different lengths (100, 300, 500, 700, 900, 1000 ms) from the beginning of the time periods (Fig. 1 (b)(iii)).

From the LFPs within these time windows, we extracted band-specific power and PLV [9] in 5 bands (theta, 4 – 8 Hz; alpha, 8 – 14 Hz; beta, 14 – 30 Hz; low-gamma, 30 – 40 Hz; high-gamma, 60 – 80 Hz). The band-specific power was calculated as the root mean square of the band-pass filtered LFPs within time windows at each recording sites. In addition, PLV between all pairs of recording sites was calculated according to (1).

$$\mathrm{PLV}_{(j,k)} = \frac{1}{|T|} \times \left| \sum_{t=T} e^{i\{\theta_j(t) - \theta_k(t)\}} \right| \quad (1)$$

In this equation, j and k indicate the recording site number, theta indicate instantaneous angle obtained by Hilbert transform of filtered LFP, T indicates the time included in the time window, and i is the imaginary unit.

3) Decoding of the Sound Frequency

SLR [6] attempted to decode the sound frequency from three characteristics of neural activities indicated above. Actually, we used SLR toolbox ver1.2.1 alpha [10] as the decoding software. The discriminations were executed separately in each rats, frequency band, and the length of time windows.

In the decoding from onset activities, SLR discriminated three sound frequencies (4 kHz, 16 kHz, and 40 kHz) and this discrimination was cross-validated for 5 times. In the decoding from the steady-state activities, SLR discriminated three stimulus conditions (first 30-second silence, 16-kHz pure tone presenting, and 40 kHz pure tone presenting) and this discrimination was cross-validated for 10 times.

After decoding, we evaluated the decoding performance of three characteristics. As the index of discriminating performance, accuracy rate, which indicates the percentage of successful discrimination from the test data, was compared.

III. Results

Fig. 2 (a) shows the representative spatial patterns of the amplitude of evoked potentials in onset activities. Because the activation focus at onset depended on sound frequency, high decoding performance was achieved with the decoding accuracy at 92.9 % (Fig. 2 (c)).

On the other hand, band-specific powers in the steady state responses did not show distinct spatial patterns (Fig .2 (b)). Nonetheless, the decoding accuracy increased with the length of time window; when the time window was 300 ms or longer, the accuracy was significantly better than the chance level in all bands tested. The highest accuracy was achieved in the high gamma band when the time windows were prolonged up to 900ms.

Fig. 3 (a) shows the PLV patterns in the steady-state activity. PLV patterns have an extremely high dimension, i.e., about 5000, without distinct patterns. Yet, these PLV patterns were again decodable with the accuracy above the chance level in all bands when the time window was 300 ms or longer. In addition, the accuracy was again best in the high gamma band with the time window of 900ms.

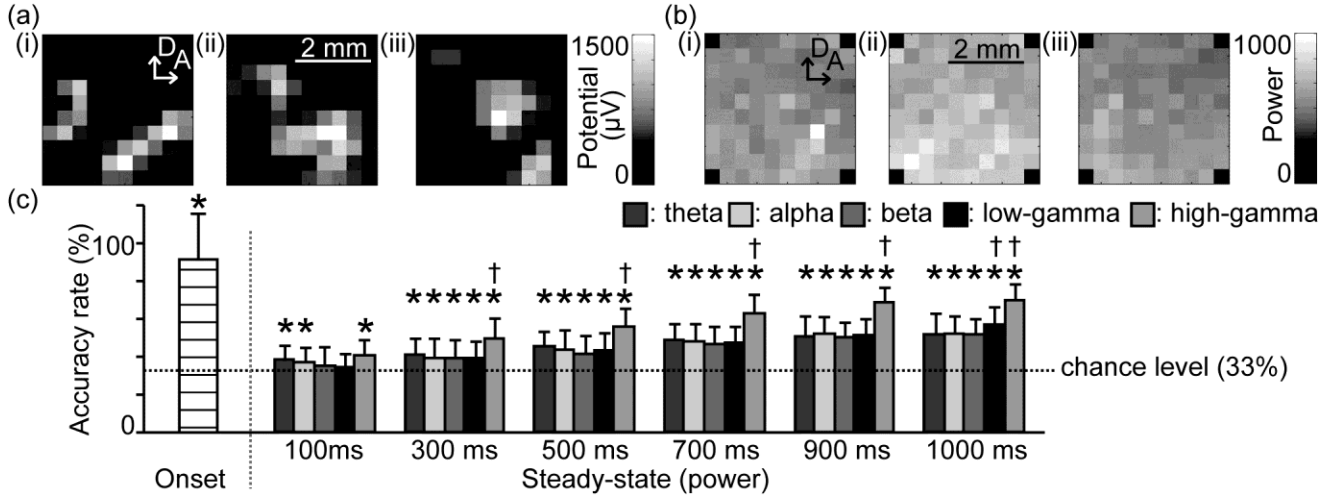

Figure 2. SLR decoding of test frequency from onset activity and from band-specific powers of steady-state activities

*(a) Representative spatial pattern of onset activity in response to varied test frequency: (i) 4 kHz, (ii) 16 kHz, and (iii) 40 kHz. (b) Representative spatial pattern of high-γ power of steady-state activity under varied stimulus conditions: (i) silence, (ii) 16 kHz, and (iii) 40 kHz. The duration of time window is 900 ms. (c) Decoding accuracy. For steady state activities, bands and the duration of time window served as parameters. Asterisks indicate that the decoding accuracy was better than the chance level: *, $p < 0.01$; two-sided one-sample t-test. Daggers indicate that the decoding accuracy was better than all lower bands: †, $p < 0.01$; two-sided two-sample t-test. Abbreviations: A, anterior; D, dorsal.*

IV. DISCUSSIONS

As expected, the decoding accuracy of onset activity was sufficiently high, i.e., 92.9 % (Fig.2 (c)). This is due to the tonotopic map in the auditory cortex [1], which generates the localized spatial pattern of evoked potentials [2]. Thus, some previous studies showed that the sound frequency can be easily discriminated from the spatial pattern of evoked potentials, using SVM or k-nearest neighbor algorithm [4, 5]. In these previous studies, the sample size was larger than the dimension of input vector. On the other hand, the sample size in this study was smaller than the dimension of the input vector. Under such condition, machine learning such as SVM does not perform well because of over-training resulting in low decoding accuracy [11] . The high accuracy in the present result thus indicates that SLR avoided the overtraining. Similar advantage of SLR was also demonstrated in fMRI study, where SLR showed high decoding performance under the small sample size [12]. Thus, SLR is useful for the decoding from neural activity, specifically when the sample size is small.

SLR decoding demonstrated that the sound frequency is represented not only in the onset activity but also in the steady-state neural activity (Fig. 2 (c), Fig. 3 (b)). In both of the band power and PLV, discrimination accuracy of high-gamma band was the best (Fig. 2 (c), Fig. 3 (b)). Discrimination accuracy increased with the length of time windows, suggesting that these features have some fluctuation on the order of 100 to 1000 ms. In addition, the best decoding performance was achieved in the high-gamma band. This band mainly reflects the activities from the cortical inhibitory interneurons [13]. These results are consistent with a notion that cortical interneurons play important roles in the sparse representation of sound information in the steady state.

V. CONCLUSION

In this study, we tried to extract the sound frequency information from the onset and steady-state neural activity in rat auditory cortex, using machine learning. As results, SLR could discriminate the sound frequency from steady-state neural activity, especially from the band power and PLV in high frequency band. From these results, steady-state neural activity in auditory cortex represents the sound information, such as sound frequency.

REFERENCES

[1] C. C. Lee, *et al*, "Tonotopic and heterotopic projection systems in physiologically defined auditory cortex," *Neuroscience*, vol. 128, 2004, pp. 871–887.

[2] H. Takahashi, *et al*, "Interfield differences in intensity and frequency representation of evoked potentials in rat auditory cortex," Hearing Research, vol. 210, 2005, pp. 9–23.

[3] R. C. deCharms and M. M. Merzenich, "Primary cortical representation of sounds by the coordination of action-potential timing," *Letter of Nature*, vol. 381, 1996, pp. 610–613.

[4] A. Funamizu, *et al*, "Decoding-accuracy-based sequential dimensionality reduction of spatio-temporal neural activities," *IEEJ Trans. EIS*, vol. 129, 2009, pp. 1648–1654. (in Japanese)

[5] A. Funamizu, *et al*, "Selection of Brain-Machine-Interface decoder depending on dispersiveness of neural activity," *IEEJ Trans. EIS*, vol. 129, 2009, pp. 1801–1807. (in Japanese)

[6] O. Yamashita *et al*, "Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns," *NeuroImage*, vol. 42, 2008, pp. 1414–1429.

[7] D. J. C. MacKay, "Baysian interpolation," *Neural Computation*, vol. 4, 1992, pp. 415–447.

[8] M. Steriade, *et al*, "Thalamocortical oscillations in the sleeping and aroused brain," *Science*, vol. 262, 1993, pp. 679–685.

[9] S. M.Doesburg, *et al*, "Asynchrony from synchrony: long-range gamma-band neural synchrony accompanies perception of audiovisual speech asynchrony," *Experimental Brain Research*, vol. 185, 2008, pp. 1–20.

[10] O. Yamashita: SLR toolbox ver.1.2.1 alpha (2009), http://www.cns.atr.jp/~oyamashi/SLR_WEB.html

[11] G.M. Foody, *et al*, "A relative evaluation of multiclass image classification by support vector machines," *Geoscience and Remote Sensing*, vol. 42, 2004, pp. 1335–1343.

[12] Y. Miyawaki, *et al*, "Visual image recognition from human brain activity using a combination of multiscale local image decoders," *Neuron*, vol. 60, 2008, pp. 915–929.

[13] M. Bartos, *et al*, "Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks," *Nature Reviews Neuroscience*, vol. 8, 2007, pp. 45–56.

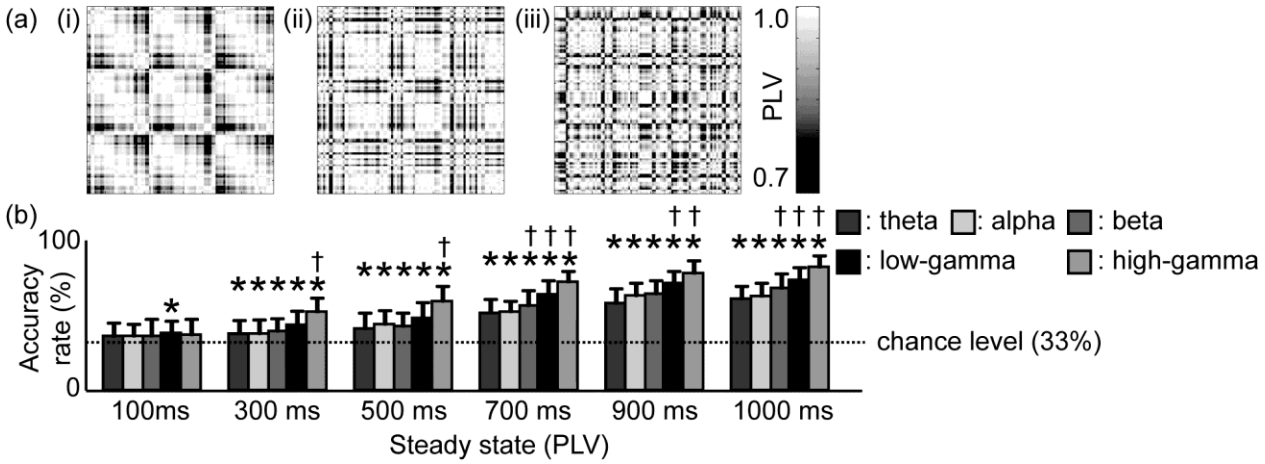

Figure 3. SLR decoding of test frequency from phase locking pattern of steady-state activities

*(a) Representative pattern of high-γ PLV of steady-state activity under varied stimulus conditions: (i) silence，(ii) 16-kHz，(iii) 40-kHz tone presentation. The duration of time window is 900 ms. (b) Decoding accuracy. For each band indicated, the accuracies were estimated in each duration of time windows separately. Asterisks indicate that the decoding accuracy was better than the chance level: *, $p < 0.001$; two-sided one-sample t-test. Daggers indicate that the decoding accuracy was better than all lower bands: †, $p < 0.01$; two-sided two-sample t-test.*