meaRtools: an R Package for the Comprehensive Analysis of Neuronal Networks Recorded on Multi-Electrode Arrays

Sahar Gelfman[1,2], Quanli Wang[1,2], Yi-Fan Lu[1,2,3], Diana Hall[1,2], Chris Bostick[1,2], Ryan Dhindsa[1,2], Matt Halvorsen[1,4], K. Melodi McSweeney[1,2,5], Ellesse Cotterill[6], Tom Edinburgh[6], Mike Beaumont[1,7], Wayne N. Frankel[1,2], Slavé Petrovski[1,8], Michael J. Boland[1,9], Andrew S. Allen[1,10], David B. Goldstein[1,2] and Stephen J. Eglen[6]

[1]Institute for Genomic Medicine, Columbia University Medical Center, New York, NY, 10032, USA

[2]Department of Genetics and Development, Columbia University Medical Center, New York, NY, 10032, USA

[3]Department of Biology, Westmont College, Santa Barbara, CA, 93108, USA

[4]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

[5]University Program in Genetics and Genomics, Duke University, Durham, North Carolina 27708, USA

[6]Cambridge Computational Biology Institute, University of Cambridge, Cambridge, CB3 OWA United Kingdom

[7]Axion BioSystems, Inc., Atlanta, GA, 30309, USA

[8]Department of Medicine, Austin Health and Royal Melbourne Hospital, University of Melbourne, Melbourne, Australia

[9]Department of Neurology, Columbia University, New York, NY, 10032, USA

[10]Department of Biostatistics and Bioinformatics, Duke University, Durham, NC 27708, US

Abstract

Here we present an open-source R package 'meaRtools' that provides a platform for analyzing neuronal networks recorded on Multi-Electrode Arrays (MEAs). Cultured neuronal networks monitored with MEAs are now being widely used to characterize *in vitro* models of neurological disorders and to evaluate pharmaceutical compounds. meaRtools provides core algorithms for MEA spike train analysis, feature extraction, statistical analysis and plotting of multiple MEA recordings with multiple genotypes and treatments. meaRtools functionality covers novel solutions for spike train analysis, including algorithms to assess electrode cross-correlation using the spike train tiling coefficient (STTC), mutual information, network burst synchronization and entropy within cultured wells. Also integrated is a solution to account for bursts variability originating from mixed-cell neuronal cultures. The package provides a statistical platform built specifically for MEA data that can combine multiple MEA recordings and compare extracted features between different genetic models or treatments. We demonstrate the utilization of meaRtools to successfully identify epilepsy-like phenotypes in neuronal networks from *Celf4* knockout mice. The package is freely available under the GPL license (GPL>=3) and is updated frequently on the

CRAN web-server repository. The package, along with full documentation can be downloaded from: https://cran.r-project.org/web/packages/meaRtools/.

Introduction

The MEA platform is now increasingly being used to study the response of neuronal networks to pharmacological manipulations and the spontaneous activity profiles of neural networks originating from genetic mouse models and derived from human pluripotent stem cells[1-4]. Recent studies aim to not only evaluate wild-type and mutation associated phenotypes, but also to recapitulate the *in vivo* response to various molecules, compounds and drug therapies [5, 6]. Capturing the many and varied activity features from a cultured neuronal network is critical for the full and accurate characterization of that network. However, MEA data are complex to handle. Moreover, an MEA experiment can last several weeks and incorporate many recordings and various treatments. Taken together, there is a genuine need for methods that can adequately characterize those neuronal networks and also provide valid assessments of phenotypic differences between genotypes and [drug] treatments in an experiment lasting many days *in vitro* (DIV).

The meaRtools package provides tools to identify complex phenotypes for assessing the effect of mutations and the screening of compounds in a multi-well MEA platform. The algorithms described here improve on existing methods through calculation of cross-correlation and mutual information between electrodes, as well as enhanced identification of synchronized network bursts (including entropy phenotypes for each well). The latter algorithm are shown here to identify recapitulation of *in vivo* epilepsy phenotypes in cultured neurons of the *Celf4* knockout mouse model. Incorporated into the package is also an algorithm that uses electrode-level burst features distributions to identify burst activity variations originating from neuronal subtypes in primary neuronal cultures. An earlier version of the package was recently used to examine the effects of microRNA-128 deficiency on the activity of cortical neural networks [6]. Last, the package provides functions to combine many recordings from multi-DIV experiments and perform rigorous statistical tests of phenotypic differences between genotypes and/or drug treatments.

Table 1. Abbreviations used in text

Abbreviation	Meaning
aE	Active electrodes
Celf4$^{-/-}$	*Celf4* knockout
DIV	Days *in vitro*
EMD	Earth mover's distance
IBI	Inter-burst interval
ISI	Inter-spike interval
MW	Mann–Whitney test
MD	Maximum distance
MEA	Multi-electrode array
MFR	Mean firing rate

NS	Network spike
NB	Network burst
STTC	Spike train tiling coefficient algorithm

Design and Implementation

meaRtools's objective is to provide a comprehensive characterization of electrode-level and network activity on a MEA plate, that is composed of one or more wells, each well consisting of multiple electrodes. The package enables a rigorous examination of differences between various genotypes and/or treatments cultured on the same plate over time. To achieve this purpose, the package provides functions to perform four major analyses (figure 1 and supplementary table S1):

1) Identify simple and complex single-electrode and network (multiple electrodes) activity phenotypes. Activity attributes that are extracted include spike and burst features on an electrode-level, as well as features of network synchronization on the multi-electrode or well-level, such as network spikes, network bursts, cross-correlation and entropy calculations.
2) Combine information from multiple recordings of the same experiment. An MEA experiment can have multiple recordings along several weeks while the neuronal culture is viable. Using meaRtools functions, many recordings of the same plate can be incorporated into a complete dataset, which can be used to test temporal replicability of results with the added advantage of enhanced statistical power.
3) Perform reproducible case/control based statistical analysis. Statistical tests are required when using the MEA platform to characterize and identify differences between genetic models and various treatments. This need intensifies when incorporating data from many recordings and comparing wells grouped by identifiers (e.g. drug or other treatment). The statistical testing scheme presented here is designed to handle this problem specifically.
4) Visualize the results in presentable ready-to-use graphs and charts. The complex picture arising from MEA recordings can be examined on various resolutions: electrode-level activity, well-level synchronization and genotype or treatment activity grouped across several wells and throughout several recordings. These different resolutions can all be visualized using designated functions. Furthermore, comparisons between genotypes and treatments are visualized along with statistical test results.

Figure 1

Input and Data Organization

The package can analyze multiple recordings with the same plate layout. The input format is similar to Axion Biosystems 'spike_list.csv' format (supplementary table S2), and can be loaded using the *read_spikelist* function. The input format is a comma separated file with a row for each spike holding spike time, electrode name and spike amplitude (mV). The input files are loaded into an R object of class 'spike.list' that holds the following information: electrode names and positions, spike trains and recording information (start/end time, machine version). The feature extraction functions introduced below add layers to this primary 'spike.list' object (supplementary table S1).
The package requires a layout scheme for each experiment that lists all wells used in the experiment and, when applicable, the treatment of each well (supplementary table S3). This layout scheme is necessary to

group wells for statistical comparisons of extracted features. A treatment label can represent multiple aspects used to distinguish a specific well, such as a genotype model, a drug treatment exposure, a change in the culture medium, an external stimulus, etc. The term 'treatment' will be used henceforth to relate to any of the above.

Identifying Simple and Complex Single and Multi-Electrode Activity Phenotypes

The package provides functions for identifying numerous features that characterize various network activity attributes. Features are calculated at three levels: electrode-, well- and treatment-level. Well-level calculations combine the information from all the electrodes within the well. Treatment-level calculations further group wells together by their assigned treatment label.

For the extraction of features for spikes, bursts and network spikes described below, the package utilizes code from the open-source Sjemea R package [7] and the algorithms of Eytan and Marom [8] and Legendy and Salcman [9].

Spikes, bursts and network spikes

Spikes, or single action potentials, are not directly detected by meaRtools. The package considers that spikes in the input file have already been filtered by a spike detection algorithm and it does not provide additional spike-detection/noise-reduction functions.
To calculate spike activity measurements, the package provides the function *calculate_spike_features*, which extracts spike information separately from the spike times of each electrode. Among the extracted features are various spike statistics per electrode and well, such as: 1) The number of active electrodes (aEs) per well, 2) The number of spikes, 3) Mean Firing Rate (MFR, in Hz) and 4) Inter-Spike Interval (ISI), which is the time between two sequential spikes.
Overall, the package extracts eleven spike features, among which are the Spike Train Tiling Coefficient (STTC) algorithm, as well as the mutual information and entropy algorithms discussed below (supplementary table S4).

Bursts are short periods of time with high spiking frequencies. The package provides two different algorithms for detecting these rapid spiking periods using the function *calculate_burst_features*; the Maximum Interval[10] and the Poisson Surprise[9] algorithms. The use of both heuristic and statistical modeling approaches for data comparison allows for an enhanced identification of features [11].

Using either the Maximum Interval or the Poisson Surprise algorithms produces the same burst features, among which are statistics per electrode and per-well for: 1) The number of bursts, 2) Burst durations, 3) Burst rates, 4) Spike rates within bursts, 5) IBIs and 6) ISIs within bursts.
Overall, the package extracts 19 burst features (supplementary table S4).

Network spikes (NS) are synchronization events of neuronal populations in a short period of time. The meaRtools package detects these events using an algorithm developed by Eytan and Marom [8] that was obtained from the sjemea package [7]. The function *calculate_network_spikes* detects NS within each well. A NS is detected when there are at least a user-defined number of aEs (default 4 aEs, or 25% of total number of electrodes) detected within a user-defined time window (default 10 ms).

NS features are next extracted using the function *summarize.network.spikes*, and include basic statistics for: 1) NS number, 2) ISIs in NS, 3) number and 4) percentage of spikes participating in NS.
Overall, the package extracts a total of ten NS features (supplementary table S4).

Novel algorithms assess network synchronization and cell-type specific burst patterns

Spike Train Tiling Coefficient

We have implemented STTC to evaluate pairwise correlations between a pair of electrodes[12]. Given N active electrodes within a well, we ignore any possible dependence of distance upon correlation and simply calculate the average STTC from all N(N-1)/2 pairwise correlations.
Average well-level STTC values per well can be computed using the *compute_mean_sttc_by_well* function.

Entropy and mutual information

We adapted two different metrics for assessing MEA data from information theory. These metrics have utility for characterizing recordings of a single well, a set of wells per treatment, and comparing treatments from the same plate.

Entropy was used as a broad measurement of the amount of disorder measured at an MEA electrode, as well as across a well. For entropy calculations, we used the standard equation for calculating the entropy of a system:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\, log(p(x_i))$$

We define H(X) as the calculated entropy measurement for electrode X. i represents a time interval bin, out of n total separate equally sized time interval bins (default is 0.1s bin size) from start to end of the recording time. In the probability distribution, Xi is set as the number of spikes in the i'th bin of electrode X divided by the total number of spikes observed in the full recording. For a well we use a mean statistic across all electrodes within well. When testing a full plate, we use the mean entropy across a group of wells (combined by treatment) as a test statistic representing the 'orderliness' of the firing patterns. Sets of mean entropies per treatment can be later compared to determine if there is evidence of a shift in the distribution between two different treatments. Normalized entropy statistics can be collected per well with the *calculate_entropy_and_mi* function.

The second metric, mutual information, was used to compare patterns from two separate MEA electrodes, and could be extended to represent the network level activity of neuronal firing in a particular well. We start with the generalized equation for mutual information (MI):

$$I(X,Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

We define I(X,Y) as the information shared between electrodes X and Y. We define a number of equally distributed time interval bins in the time period and for each electrode, count the spikes in each b_i. As

before, we transform X and Y into separate probability mass functions, where the probability of a spike falling in a particular time interval bin equals the count at that bin divided by the total number of spikes that detected in the recording.

Given that the number of spikes observed during recording time can vary between electrodes, it was important to further transform X and Y to take this into account. For this we take the spike count in a time interval to infer the presence or absence of a burst, and as such we can classify each time interval at an electrode as either a burst member or non-member. To do this we transformed each input vector X such that the value X_i equals 1 if the spike count is greater than the 75th percentile of spike counts across all bins in X, and set as 0 otherwise. Such a simple transformation of the transformation of the data means that the probability mass function for X is collapsed down to p(X=0) and p(X=1). Subsequent MI calculations are far more efficient since in terms of combinatorics, only 4 outcomes at a given time interval are possible: X=0/Y=0, X=1/Y=0, X=0/Y=1 and X=1/Y=1. As such we explicitly compute the mutual information between electrodes X and Y as:

$$I(X,Y) = \sum_{x,y \in \{0,1\}} p(x,y)\ log_2 \frac{p(x,y)}{p(x)p(y)}$$

We produce a distribution of these pairwise statistics per well, and aggregate the statistics per treatment. A Mann–Whitney (MW) test can be later performed on these sets of values between two treatments to determine if there is evidence for neurons in one treatment having a higher level of coordinated network level firing than in another treatment. Average well-level pairwise mutual information values per well can be computed using the *calculate_entropy_and_mi* function.

Burst distributions

Burst distributions are a novel way to calculate, compare and visualize certain burst features by looking at their distributions along a recording. The reason behind constructing this method is that primary cultures contain multiple neuronal subtypes (e.g. GABAergic and glutamatergic), which demonstrate different activity signatures [13, 14]. Thus, merely extracting a mean and standard deviation of a feature will misrepresent the activity fluctuations that arise from the combined activity of neuronal subtypes in the network. For example, spike frequencies within bursts may differ between GABAergic interneurons and glutamatergic neurons [15], with GABAergic neurons often exhibiting narrower spike wave-forms and faster-spike activity [16, 17]. Furthermore, certain anti-epileptic or anti-psychotic drugs selectively target specific neuronal subtypes. However, this selective effect may not be observed when comparing the average change of an entire cultured network. The package provides an algorithm that offers a way to handle multiple cell-type variability.
The function *calc_burst_distributions* calculates empirical distributions for bursting features and compares them between treatments using two independent methods (figure 2A, see Statistical Testing and Visualization).

Figure 2

Distributions are calculated for five burst features: IBIs, ISIs within bursts, number of spikes in bursts, burst durations and spike frequencies within bursts (firing rate, Hz). For each feature, the algorithm

adjusts for variability between electrodes in a well. This is done by calculating the histogram of a feature in each electrode separately (figure 2A, left panel) and normalizing it to values between 0-1 (figure 2A, middle panel). Next, all normalized histograms are grouped and averaged by treatment labels. The algorithm permits performing this step also by grouping electrodes first by wells and then averaging well information by treatment. To later test for differences between treatments, the package provides a function which performs distribution comparison tests, permutes electrode labels and plots the results (figure 2A, right panel, see Statistical Testing and Visualization).

Network bursts

Network bursts (NB) are longer and more intense synchronization events than NS and correspond to electrode-level burst activity that is synchronized across electrodes in a well. The underlying reason to identify NBs is that, while the NS detection algorithm identifies short network synchronized activity lasting tens of milliseconds, synchronized bursting events were shown to last tenths of seconds to seconds in MEA experiments [11, 18, 19]. To catch these long network synchronization events, a method was constructed that investigates bursting patterns within wells and also between wells clustered based on treatments.

The function *calculate_network_bursts* combines burst information at the electrode-level into well-level data as the presentation of network synchronization of bursts across a well (figure 2B). First, spike time within spike trains from all electrodes is binned using a bin size of 2ms to guarantee that at most one spike is called within each bin. Next, a Gaussian filter with user-defined window sizes (defaults are: 10, 20 and 50 ms) is applied to smooth the binned spike trains from each electrode. The smoothed signal is then further standardized to have a maximum signal value of 1. All smoothed signals at the electrode-level are then combined and smoothed again using the same Gaussian filter. The final result from this step is a smoothed signal at each given window size that measures the overall synchronization of all electrodes in a well, with larger values indicating higher level of synchronized bursting activities. Then, the Otsu global thresholding method is applied to the well-level signal to automatically detect burst intervals [20]. This method was chosen for its simplicity and parameter free nature, although other methods, such as adaptive thresholding, can be utilized. Last, based on the network burst intervals obtained from Otsu thresholding, network burst information is collected at the well-level.

The algorithm extracts statistics for: the number and rate of NBs, the number and percentage of spikes participating in NBs and the spike intensities within NBs, which is the spike rate within NBs.

Overall, the package extracts 11 network burst features for each time window to a total of 33 features (supplementary table S4).

Combining Multiple Recordings

MEA experiments are constructed from multiple recordings of the same plate over a certain period of time. Correctly assessing the activity and differences over time requires analyzing several recordings as one set of information with various time points. The package provides functions for combining several recordings and filtering wells from the combined dataset based on inactivity measurements. All feature extraction functions store the extracted information in the same ‘spike.list’ object. The function *aggregate_features* uses the information stored in the ‘spike.list’ object to combine data from all the analyzed recordings into an aggregated table for each feature. The aggregated tables have recording labels as columns and well-labels as rows, and can be printed as csv files or used later for treatment comparisons. The package also provides the function *filter_wells* to exclude inactive wells from these

aggregated tables. An active well is measured using a minimum number of aEs (default 4, or ¼ of the total number of electrodes). The function *filter_wells* considers whether a well has been active in more than a certain percentage of recordings in the experiment (default is 50%). Inactive wells that fail to meet this criterion are not used when comparing treatments.

Statistical Testing and Visualization

The combined tables from multiple recordings can next be used to compare treatments along the experiment. The function *permute_features_and_plot* performs all the necessary statistical tests between treatment labels and plots the results of all features in .pdf format in a designated output directory. The tests are performed as follows: First, a MW test is performed to compare distributions of each feature between treatments, and the resultant p-value is recorded. Next, a permutation scheme is performed where the treatment labels of the active wells are randomly shuffled X times (default 100) while the observations within each well are kept intact. This preserves correlations between time points within wells while breaking any relationship with treatment and subsequent outcome. A permutation p-value for the original MW test is computed as the proportion of permuted data MW p-values that were less than or equal to the MW p-value of the original un-permuted dataset. Last, for each feature, a graph and a table (csv format) are printed with the mean and standard error (SEM) of the measured features for each of the recordings.

Comparing and plotting burst distributions, is done using the function *dist_perm*. The algorithm works as follows: For each of the five burst features, the function *calc_burst_distributions* generates a normalized histogram per electrode (supplementary table S5). Next, the function *dist_perm* groups all distributions by treatment labels and compares them between treatments using two methods: 1) The Earth Mover's Distance (EMD), using R package emdist (http://www.rforge.net/emd) and 2) the Maximum Distance (MD) between the cumulative distributions of the normalized histograms.
Once the test results for the EMD and MD algorithms are computed for the original dataset, a permutation scheme is performed where the treatment labels of the active wells are randomly shuffled X times (default 100) while the distributions within each well are kept intact. A permutation p-value for EMD is then computed as the proportion of permuted data EMD values that are equal to or greater than the original EMD value from the un-permuted dataset. A permutation p-value for MD is defined similarly. Last, both the normalized and cumulative histograms are plotted with the final permuted p-value.

Results

The meaRtools package can be used as an analysis pipeline on an experiment composed of data from several recordings and with several treatments. The package is shown here to extract novel biological insights by identifying neuronal phenotypes of a *Celf4* epilepsy mouse model[21] and was also recently utilized to assessing the effects of miRNA deficiency on neuronal activity [6], both of which demonstrate the ability of the algorithms presented here to identify and provide robust statistical measurements for simple and complex network-associated phenotypes. The pipeline step-by-step workflow and results over an exemplary dataset consisting of three recordings is available online with the package vignette. Below, we present various experimental examples to illustrate the diverse features and uses of meaRtools.

General Activity Information

To gain a preliminary view of plate activity, the package generates graphs of activity measurements for three levels of data: electrode-, well- and plate-levels. Presented here are a subset of these graphs for a single recording in a 48-well plate (16 electrodes per well) containing cultured cortical neural networks from the brains of postnatal wild-type mice. The lowest resolution map of the plate shows a matrix of all aEs for each well in the plate (figure 3A). A higher resolution graph shows electrode activity per well as the MFR of all aEs (figure 3B). Even higher resolution shows the MFR of each electrode in each well (figure 3C represents a 900s recording). The latter is plotted for each well separately.

Figure 3

In order to demonstrate network behavior before and after each NS, we show the number of electrodes participating in NS around the peak of a network event (figure 3D). For example: some networks present a decline in participating electrodes before the event takes place (wells B1, C1 and D1 in figure 3D), while others exhibit a gradually increasing number of electrodes participating in a NS before a fast accumulation of electrodes leading to the NS peak (wells F1 and F8, figure 3D).
The full set of graphs that can be printed by the package is available as supplementary information and includes log ISI statistics, spike statistics within bursts and other network information. These graphs provide an overall view of the activity of a single recording.

Intraplate Treatment Comparisons

For every extracted feature a comparison can be made between treatments in a multiple-recording experiment. For every activity attribute (i.e. spikes, bursts, NSs and NBs), an output folder is created with .csv files for every feature, holding average values per well for every recording analyzed. An example is shown here for Spike Intensity within NBs representing DIV 11-14 analysis of two genotypes of a genetic mouse model of epilepsy; which are heterozygous (+/-) or homozygous (-/-) for a specific mutation (figure 4A). Also printed are graphs showing mean and standard error for each feature. These graphs compare the treatments over all the analyzed recordings (figure 4B-4E). The output directories include tables and graphs for a total of 70 features: eight spike features, 19 burst features, 10 NS features and 33 NB (see supplementary table S4).

Figure 4

The comparison analysis is flexible and can be performed for subsets of recordings and treatments. The example in figure 4B and 4C shows a comparison of network activity between the two genotypes (+/- and -/-). Under each graph are the results of multiple MW and permutation tests performed between the genotypes. The Spike Intensity within Network Bursts feature (spikes per NB per sec) analyzed here illustrates the synchronization level of a network. A comparison of treatments, using five recordings from an experiment comprised of 14 DIV, shows a trend for higher spike intensities in the homozygous genotype, which is not significant after a permutation test (figure 4B). However, inclusion of data from three more DIV shows a significantly increased network synchronization in the homozygous genotype (figure 4C).
The feature comparison analysis is not limited to the number of recordings or treatments. In figure 4D, the number of aEs are shown for an experiment spanning 27 DIV. This experiment had three treatments: a vehicle treated control and two concentrations (0.1nM and 1nM) of a sodium channel blocker (figure 4D-4E). The treatment was administered on DIV 20 and the plate was recorded for 7 DIV following drug administration. Analysis of the number of aEs shows no significant difference between groups before the

drug was added, and a dramatic decrease in number of aEs in the wells treated with 1nM but not those treated with 0.1nM (figure 4D, green vs red lines, respectively). As expected, MFR analysis indicates a significant decrease in MFR for both 0.1nM and 1nM drug treatments[22]. Moreover, the kinetics of MFR decrease are relative to the drug concentration, as is the time it takes the MFR to return to untreated values (figure 4E, red and green lines).

Overall, these results present the flexibility of the functions to handle varying number of recordings and provide true biological insights. The ability to test treatment differences over many recordings, presented here for the first time, provides a strong and valuable tool to assure drug, compound or genotype effect.

Neuronal Network Analysis of the *Celf4* deficiency-induced seizure mouse model

We utilized the full capabilities of meaRtools to identify and compare complex activity phenotypes of a *Celf4* knockout (*Celf4*$^{-/-}$) mouse seizure model. [21, 23]. Here, we demonstrate that the MEA platform, analyzed with meaRtools, can identify epilepsy-like phenotypes in neuronal networks from *Celf4*$^{-/-}$ mice.

Seizures are often characterized by hypersynchronous discharges that may occur at a specific region of the cortex and spread into neighbouring brain areas. The cellular mechanism of seizure initiation is thought to be the network hyper-synchronization and high frequency bursts consisting of increased density of action potentials, presumably due to an excitation/inhibition imbalance [24]. Sufficiently synchronized bursts may pass the threshold of surrounding inhibition and activate neighboring neurons leading to broader recruitment, network propagation and ultimately seizures. While the *in vitro* manifestations of seizures are not fully understood, it is thought that both increased synchronicity of network firing and increased bursting are analogous to the *in vivo* phenotype [25]. As mentioned, *Celf4* deficiency is known to cause neurological phenotypes in mice including, most prominently convulsive seizures [21, 23]. Here, we use *Celf4*$^{-/-}$ mice as “proof of concept” for whether meaRtools synchronization algorithms can identify aberrant excitability phenotypes that are often characteristic of mouse seizure models.

We examined the various features that meaRtools computes and found that *Celf4*$^{-/-}$ neurons consistently showed significant elevation in several features compared to wild-type (*Celf4*$^{+/+}$) neurons, among which were: increased frequency of spikes in bursts (permutation p value < 0.01), increased mutual information between electrodes (permutation p value < 0.01) and an increase in various NB phenotypes (permutation p values < 0.01 - 0.02, figure 5 and supplementary table S6). Together, these features offer a view of the change in network synchronization of *Celf4*$^{-/-}$ mice, and point to a higher density of spikes in bursts and a higher synchronization of bursts between neurons belonging to the same neuronal network. This phenomenon was tested across an experiment lasting 23 DIV, where *Celf4*$^{+/+}$ and *Celf4*$^{-/-}$ genotypes were cultured on 12 and 11 active wells, respectively. The final statistical tests were performed on recordings from DIV 15-23, starting when the networks were developed and reached a stable MFR and ending with cell death. The results were replicated in another experiment lasting 23 DIV, where each genotype was cultured in nine wells.

We have also examined the change in MFR between the genotypes, and find a slight increase in MFR in the *Celf4*$^{-/-}$ neurons (supplementary Table S6). However, this increase did not survive the meaRtools permutation scheme and could not explain the significant increase in synchronization phenotypes observed in figure 5.

Figure 5

Burst Distributions Comparisons

In addition to the comprehensive examination of synchronization between electrodes, meaRtools provides a novel approach to identify cell subtype-specific burst patterns. Burst features distributions can account for the differing behaviors of neuronal subtypes in a primary culture. For example, burst activity might have differing durations, fluctuating spike rates and other features that are influenced by different activity profiles of specific cell types. While these varying activity properties might not be caught using simple statistical measurements, they can be identified using empirical distributions of features.

Burst features distributions can be compared between treatments in single or multiple recordings. Single recording treatment comparisons are tested using the Kolmogorov-Smirnov test (K-S test) for comparing probability distributions. For instance, when comparing the effects of cell density on network behavior, a significantly higher proportion of low Spike Frequencies is observed at low cell density of 25,000 (25k) cells relative to higher densities (figure 6A, red line), suggesting that higher cell densities have higher spike frequencies within bursts. In a separate experiment, we compared the effects of two different microbial light sensitive membrane proteins: Channelrhodopsin-2 (Chr2) and Archaerhodopsin-T (ArchT) on burst duration. We observed a higher ratio of long burst durations for 'Chr2' that is not significant between treatments (figure 6B, red line).

Figure 6

Burst distribution comparisons can be performed over multiple recordings. The analysis of number of spikes in a burst is presented here for an experiment with 24 recordings that were combined using the methods above. The differences between treatments are calculated using two measurements of distances (figure 6C and 6D), and followed by permutation tests. This specific analysis successfully validates that a treatment using a specific compound treatment at 25µM (figure 6C-6D, red line) has significantly more bursts with low number of spikes than the 100uM treatment (blue line).

In conclusion, the meaRtools package extracts a detailed report of over 70 activity phenotypes (features), both previously existing and novel. The package provides a single platform to handle multiple recordings of the same experiment, and the tools to perform statistical comparisons between treatments on these multiple recording experiments.

Availability and Future Directions

The meaRtools package is open-source and freely available under the General Public License version 3.0 (GPL>=3). The package is available on the CRAN web-server repository (https://cran.r-project.org/web/packages/meaRtools/index.html). Updated source-code can be found at https://github.com/igm-team/meaRtools. The package provides a step by step vignette for running an MEA analysis pipeline using exemplary datasets in an effort to make MEA analysis accessible to all.

Here we explain the major analyses that can be done using meaRtools and focus on several features to perform detection of phenotypic differences. However, the current version of the package detects 73 features and five feature distributions that can be used in the holistic evaluation of an MEA experiment.

Users are encouraged to explore all the features and capabilities the package entails with the help of the package vignette.

The package is updated regularly; each version incorporates additional capabilities to detect and test additional phenotypes. Current work focuses on adding machine learning algorithms to distinguish between treatments, graphical representation of network burst patterns and activity pattern recognition algorithms within and between wells.

Figure and table legend

Figure 1. A general scheme of an analysis workflow for several MEA recordings.

Figure 2. Schemes for computing network bursts and burst distributions. **A)** Creating burst features distributions. First, burst feature histogram is calculated for each electrode (left panel). In this example, it is calculated for burst duration. Next, histograms are normalized to number of values, resulting in a 0-1 value. Last, all electrodes are averaged to create a normalized distribution plot (top right panel) and a cumulative plot (bottom right panel) for each tested treatment. **B)** Detecting network bursts. Spike data from a raster plot (upper panel) showing the spikes (x-axis) for each active electrode (y-axis) is binned and combined through a weighted Gaussian kernel smoothing method to generate well level signal (blue lines in lower panel) and the Otsu global thresholding algorithm[20] is applied to identify intervals above the threshold (red horizontal line) as network bursts.

Figure 3. General information of plate activity. **A)** A matrix representing a 48-well plate, each consisting of 16 electrodes. aEs are represented by name consisting of column+row position in each well ("11" for first electrode to "44" for the last). **B)** MFR (Hz) for aEs per well in a 48-well plate. Title for each well has the genotype label of the well: +/- (wild-type), -/- (homozygous), +/- (heterozygous), NA (Not available). **C)** Average MFR of all 16 electrodes of well A6, presented for each second of a 900s recording. **D)** Average number of electrodes participating in NSs around the peak of a network event. The x-axis represents user-defined time bins (default is 100 ms) before and after a NS peak (-10 equals 1s before the peak). Title for each well consists of well-name and number of identified NSs.

Figure 4. Comparing features between treatments. In panels A-C, treatment refers to genotype (heterozygous, +/- and homozygous, -/-). In panels D and E, treatment refers to drug treatment. **A)** Output tables are printed for every feature, holding average values (in this case, Spike Intensity per aEs) per well (rows) and for every recording analyzed (columns). **B)** Spike Intensity per aEs for two genotypes by DIV on a subset of five recordings. This graph corresponds to the full aggregated table of A). **C)** Same as B), but for a subset of eight recordings. **D)** Comparing the number of aEs for wells treated with a vehicle or one of two channel blocker treatments (0.1nM and 1nM) along a 27 DIV experiment. **E)** MFR for the same experiment as in D), showing differences in the effects of drug concentrations.

Figure 5. *Celf4*$^{-/-}$ neurons show elevated network synchronization phenotypes. **A)** Spike frequency in bursts for *Celf4*$^{+/+}$ (blue) and *Celf4*$^{-/-}$ (red) neural networks. **B)** Mutual information between electrodes is increased in *Celf4*$^{-/-}$ networks (red). **C)** The percent of spikes participating in NBs is increased in *Celf4*$^{-/-}$ neurons (red), suggesting that ratio of spikes that participate in synchronized network events is higher in the *Celf4*$^{-/-}$ networks. **D)** Raster plots show network spikes (vertical green lines) and bursts (horizontal red lines) for two adjacent wells: a *Celf4*$^{-/-}$ well (upper panel) and a *Celf4*$^{+/+}$ well (lower panel). *Celf4*$^{-/-}$

exhibits few sporadic spikes between highly synchronized events, while *Celf4*$^{+/+}$ exhibits less network events and more sporadic spikes. Raster plots present 60 seconds of 900 seconds recordings.

Figure 6. Burst features distributions. **A)** Frequencies (y-axis) of spike-rate in bursts (x-axis) are calculated for four different cell density cultures in a single recording. A user defined maximum of 300 Hz is set. **B)** Frequencies (y-axis) of burst durations (x-axis) are presented when introducing two different microbial light sensitive membrane proteins vs. untreated cells in a single recording. A user defined maximum for burst duration was set to 3 seconds. **C)** Combining burst features distributions of number of spikes in bursts from 24 sequential same-plate recordings. Treatments were automatically tested for difference using the EMD test and results were displayed after permutations. **D)** Cumulative distributions of the same data as in C), treatments were automatically tested for difference using the MD test and results were displayed after permutations.

Funding

This research was supported by funding from R01 NS091118 (W.N.F.), Wellcome Trust (E.C), CURE AHC and the Robbins Family Gift (M.J.B. and D.B.G.).

References

1. Bateup HS, Johnson CA, Denefrio CL, Saulnier JL, Kornacker K, Sabatini BL. Excitatory/inhibitory synaptic imbalance leads to hippocampal hyperexcitability in mouse models of tuberous sclerosis. Neuron. 2013;78(3):510-22.
2. Cao Z, Hulsizer S, Tassone F, Tang HT, Hagerman RJ, Rogawski MA, et al. Clustered burst firing in FMR1 premutation hippocampal neurons: amelioration with allopregnanolone. Hum Mol Genet. 2012;21(13):2923-35.
3. Chailangkarn T, Trujillo CA, Freitas BC, Hrvoj-Mihic B, Herai RH, Yu DX, et al. A human neurodevelopmental model for Williams syndrome. Nature. 2016;536(7616):338-43.
4. Wainger BJ, Kiskinis E, Mellin C, Wiskow O, Han SS, Sandoe J, et al. Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. Cell Rep. 2014;7(1):1-11.
5. Valdivia P, Martin M, LeFew WR, Ross J, Houck KA, Shafer TJ. Multi-well microelectrode array recordings detect neuroactivity of ToxCast compounds. Neurotoxicology. 2014;44:204-17.
6. McSweeney KM, Gussow AB, Bradrick SS, Dugger SA, Gelfman S, Wang Q, et al. Inhibition of microRNA 128 promotes excitability of cultured cortical neuronal networks. Genome Res. 2016;26(10):1411-6.
7. Eglen SJ, Weeks M, Jessop M, Simonotto J, Jackson T, Sernagor E. A data repository and analysis framework for spontaneous neural activity recordings in developing retina. Gigascience. 2014;3(1):3.
8. Eytan D, Marom S. Dynamics and effective topology underlying synchronization in networks of cortical neurons. J Neurosci. 2006;26(33):8465-76.
9. Legendy CR, Salcman M. Bursts and recurrences of bursts in the spike trains of spontaneously active striate cortex neurons. J Neurophysiol. 1985;53(4):926-39.
10. Technologies N. Neuroexplorer Manual 2014. Available from: http://www.neuroexplorer.com/downloads/NeuroExplorerManual.pdf.
11. Cotterill E, Charlesworth P, Thomas CW, Paulsen O, Eglen SJ. A comparison of computational methods for detecting bursts in neuronal spike trains and their application to human stem cell-derived neuronal networks. J Neurophysiol. 2016;116(2):306-21.

12. Cutts CS, Eglen SJ. Detecting pairwise correlations in spike trains: an objective comparison of methods and application to the study of retinal waves. J Neurosci. 2014;34(43):14288-303.
13. Molyneaux BJ, Arlotta P, Menezes JR, Macklis JD. Neuronal subtype specification in the cerebral cortex. Nat Rev Neurosci. 2007;8(6):427-37.
14. Zeisel A, Munoz-Manchado AB, Codeluppi S, Lonnerberg P, La Manno G, Jureus A, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science. 2015;347(6226):1138-42.
15. Weir K, Blanquie O, Kilb W, Luhmann HJ, Sinning A. Comparison of spike parameters from optically identified GABAergic and glutamatergic neurons in sparse cortical cultures. Front Cell Neurosci. 2014;8:460.
16. Mann EO, Paulsen O. Role of GABAergic inhibition in hippocampal network oscillations. Trends Neurosci. 2007;30(7):343-9.
17. Bean BP. The action potential in mammalian central neurons. Nat Rev Neurosci. 2007;8(6):451-65.
18. Belykh I, de Lange E, Hasler M. Synchronization of bursting neurons: what matters in the network topology. Phys Rev Lett. 2005;94(18):188101.
19. Gullo F, Manfredi I, Lecchi M, Casari G, Wanke E, Becchetti A. Multi-electrode array study of neuronal cultures expressing nicotinic beta2-V287L subunits, linked to autosomal dominant nocturnal frontal lobe epilepsy. An in vitro model of spontaneous epilepsy. Front Neural Circuits. 2014;8:87.
20. Otsu N. A threshold selection method from gray-level histograms. Automatica. 1975;11(285-296):23-7.
21. Yang Y, Mahaffey CL, Berube N, Maddatu TP, Cox GA, Frankel WN. Complex seizure disorder caused by Brunol4 deficiency in mice. PLoS Genet. 2007;3(7):e124.
22. McConnell ER, McClain MA, Ross J, Lefew WR, Shafer TJ. Evaluation of multi-well microelectrode arrays for neurotoxicity screening using a chemical training set. Neurotoxicology. 2012;33(5):1048-57.
23. Wagnon JL, Mahaffey CL, Sun W, Yang Y, Chao HT, Frankel WN. Etiology of a genetically complex seizure disorder in Celf4 mutant mice. Genes Brain Behav. 2011;10(7):765-77.
24. In: Bromfield EB, Cavazos JE, Sirven JI, editors. An Introduction to Epilepsy. West Hartford (CT)2006.
25. Lignani G, Raimondi A, Ferrea E, Rocchi A, Paonessa F, Cesca F, et al. Epileptogenic Q555X SYN1 mutant triggers imbalances in release dynamics and short-term plasticity. Hum Mol Genet. 2013;22(11):2186-99.

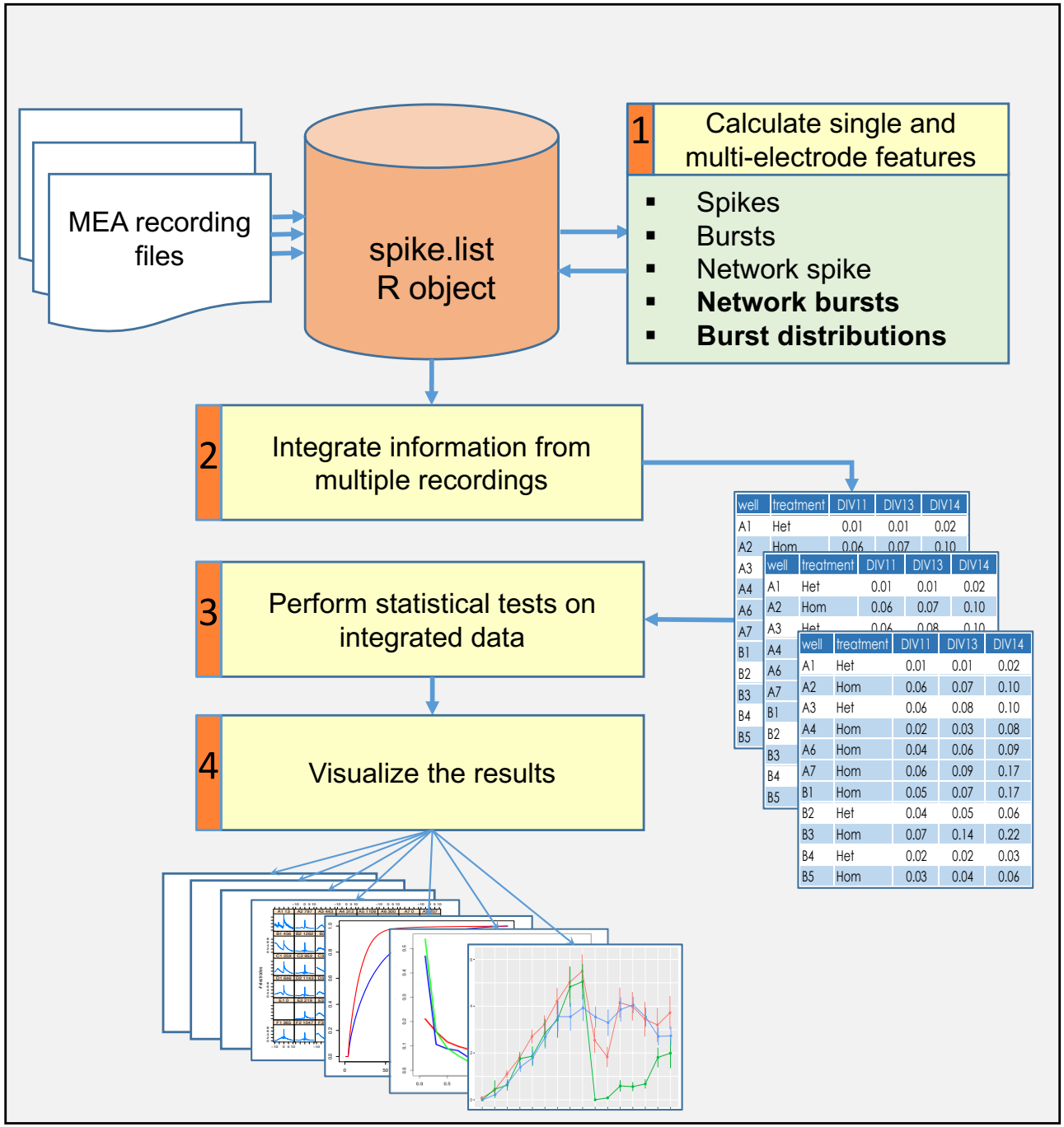
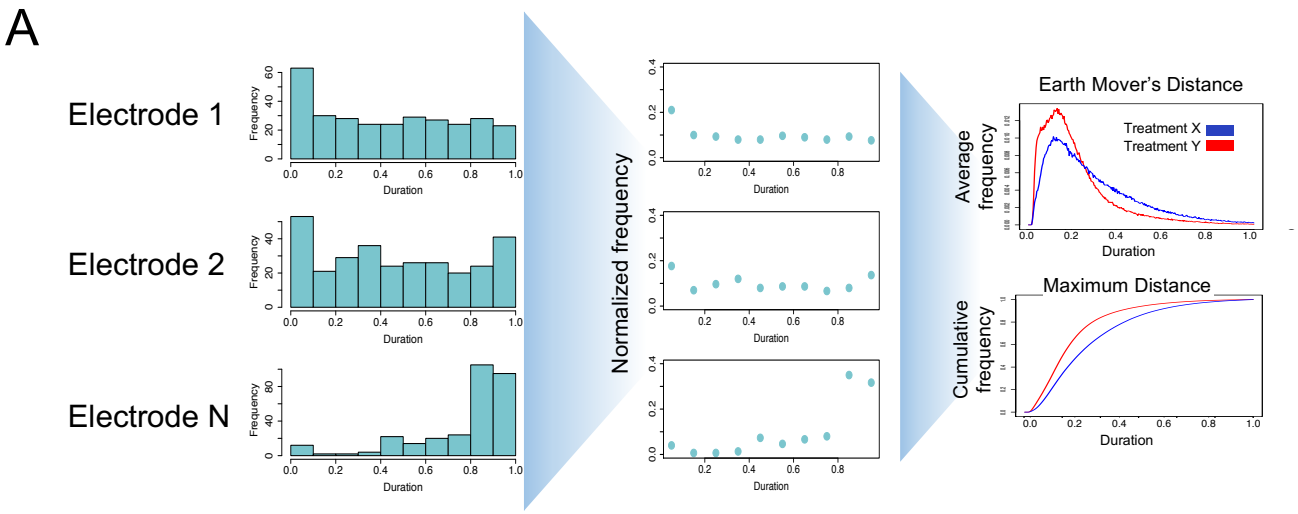

MEA recording files
spike.list R object
1
Calculate single and multi-electrode features
Spikes
Bursts
Network spike
Network bursts
Burst distributions
2
Integrate information from multiple recordings
3
Perform statistical tests on integrated data
4
Visualize the results
well treatment DIV11 DIV13 DIV14
A1 Het 0.01 0.01 0.02
A2 Hom 0.06 0.07 0.10
A3 Het 0.06 0.08 0.10
A4 Hom 0.02 0.03 0.08
A6 Hom 0.04 0.06 0.09
A7 Hom 0.06 0.09 0.17
B1 Hom 0.05 0.07 0.17
B2 Het 0.04 0.05 0.06
B3 Hom 0.07 0.14 0.22
B4 Het 0.02 0.02 0.03
B5 Hom 0.03 0.04 0.06

Figure 1


A
Electrode 1
Electrode 2
Electrode N
Frequency
Duration
Normalized frequency
Earth Mover's Distance
Average frequency
Treatment X
Treatment Y
Maximum Distance
Cumulative frequency

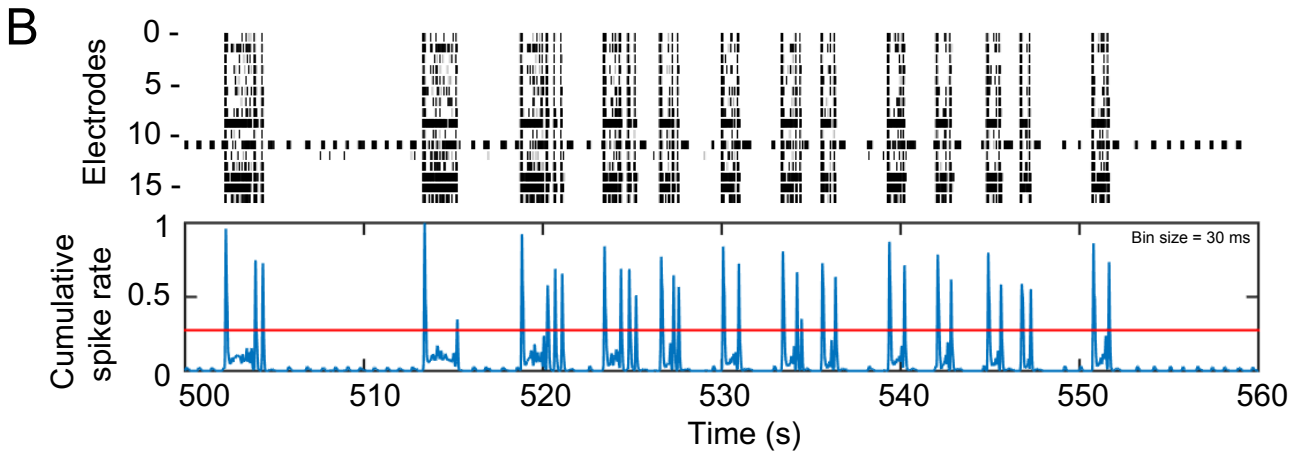

B
Electrodes
Cumulative spike rate
Bin size = 30 ms
Time (s)

Figure 2

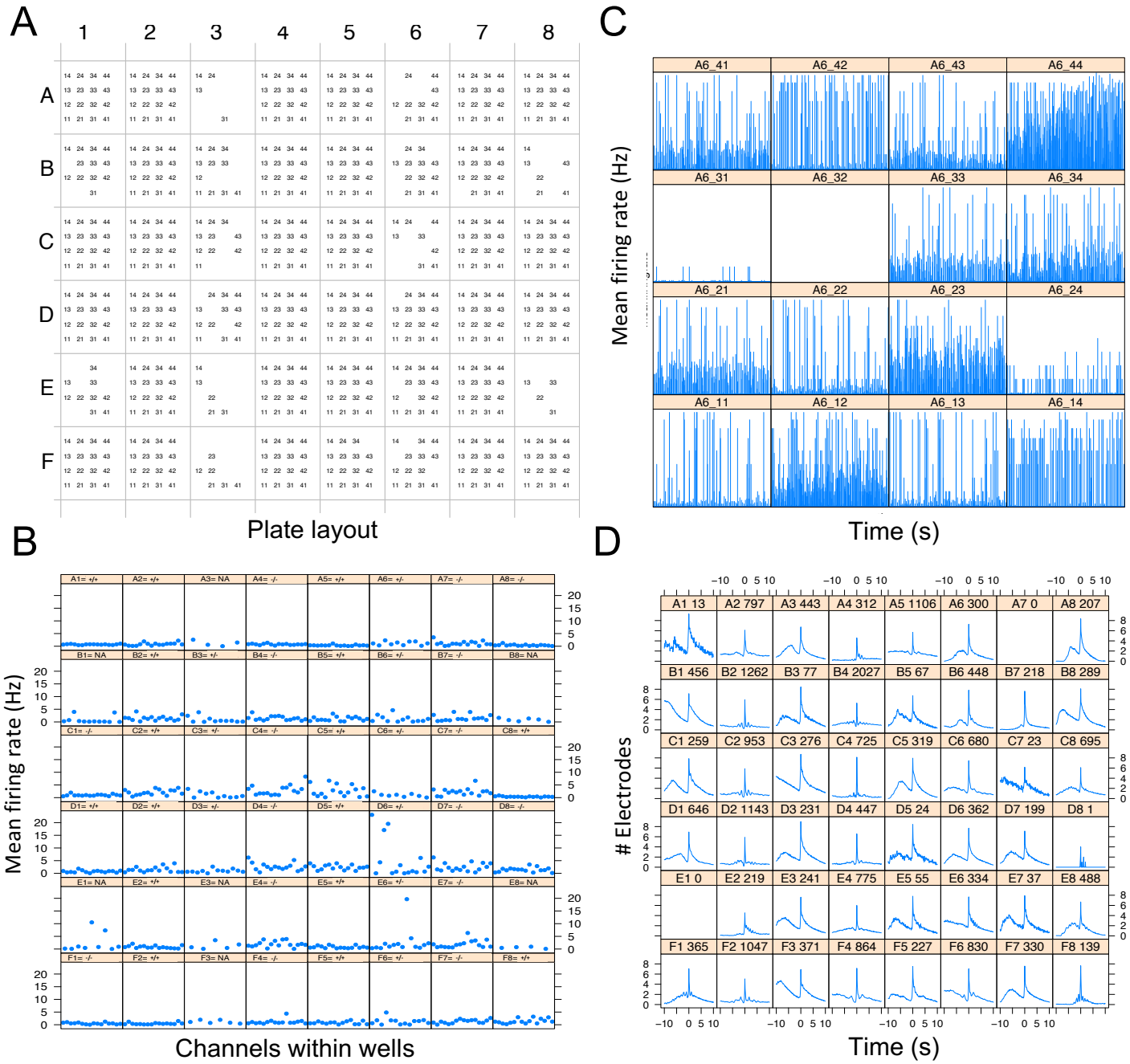
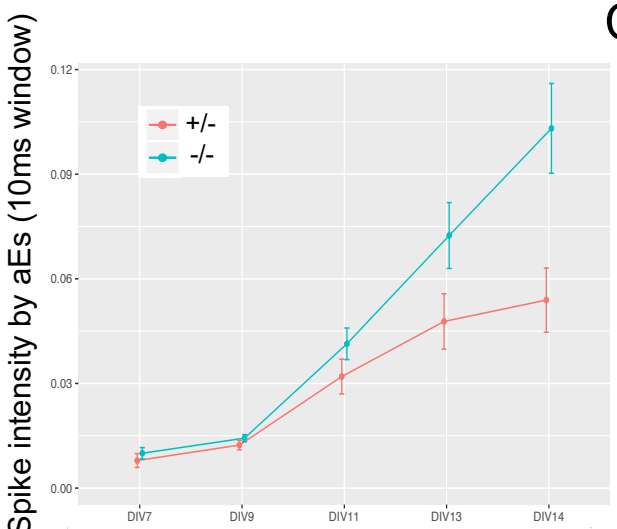

A
1 2 3 4 5 6 7 8
A B C D E F
Plate layout
C
A6_41 A6_42 A6_43 A6_44
A6_31 A6_32 A6_33 A6_34
A6_21 A6_22 A6_23 A6_24
A6_11 A6_12 A6_13 A6_14
Mean firing rate (Hz)
Time (s)
B
Mean firing rate (Hz)
Channels within wells
D
A1 13 A2 797 A3 443 A4 312 A5 1106 A6 300 A7 0 A8 207
B1 456 B2 1262 B3 77 B4 2027 B5 67 B6 448 B7 218 B8 289
C1 259 C2 953 C3 276 C4 725 C5 319 C6 680 C7 23 C8 695
D1 646 D2 1143 D3 231 D4 447 D5 24 D6 362 D7 199 D8 1
E1 0 E2 219 E3 241 E4 775 E5 55 E6 334 E7 37 E8 488
F1 365 F2 1047 F3 371 F4 864 F5 227 F6 830 F7 330 F8 139
Electrodes
Time (s)

Figure 3

A

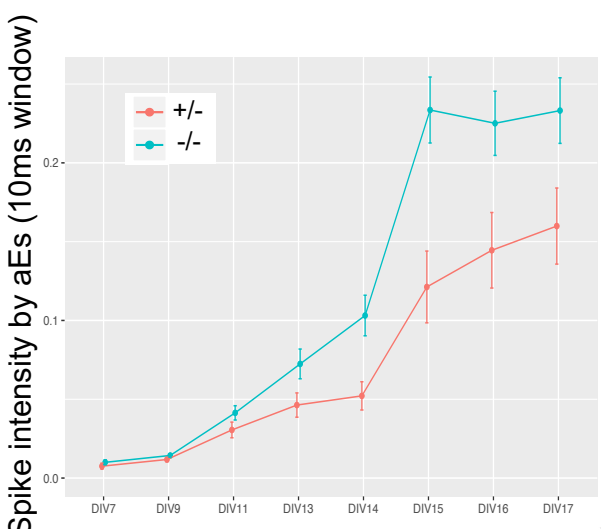
well	treatment	DIV11	DIV13	DIV14
A1	+/-	0.01	0.01	0.02
A2	-/-	0.06	0.07	0.10
A3	+/-	0.06	0.08	0.10
A4	-/-	0.02	0.03	0.08
A6	-/-	0.04	0.06	0.09
A7	-/-	0.06	0.09	0.17
B1	-/-	0.05	0.07	0.17
B2	+/-	0.04	0.05	0.06
B3	-/-	0.07	0.14	0.22
B4	+/-	0.02	0.02	0.03
B5	-/-	0.03	0.04	0.06

B



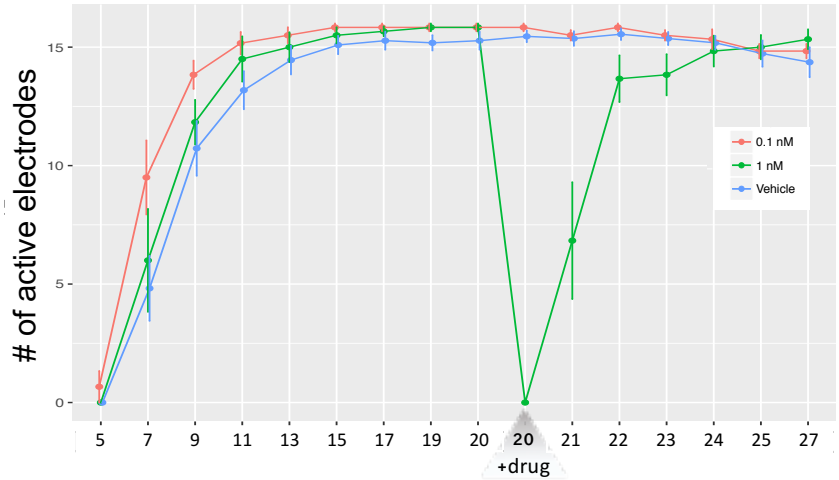
	Treatment/Genotype	perm.pval	MW pval
1	+/- vs. -/-	0.08	0.0146

C



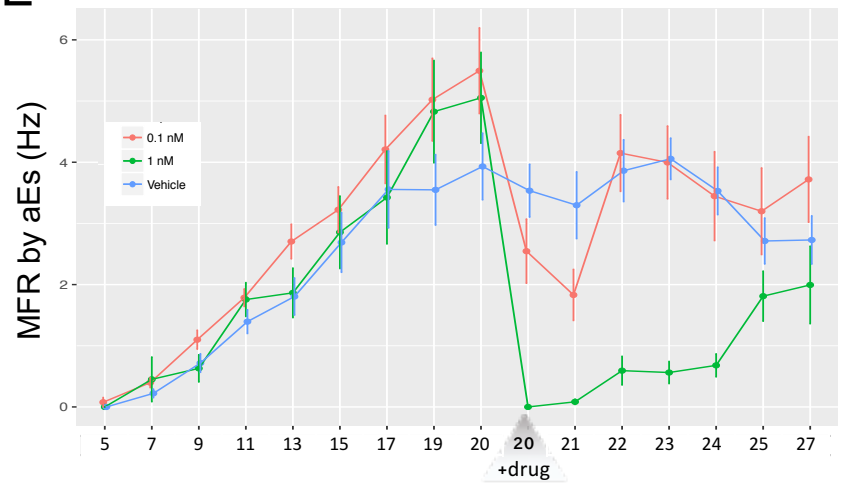
	Treatment/Genotype	perm.pval	MW pval
1	+/- vs. -/-	0.03	0.000105

D



	Treatment/Genotype	perm.pval	MW pval
1	Vehicle vs. 0.1 nM	0.39	0.0506
2	Vehicle vs. 1 nM	0.41	0.136

E



	Treatment/Genotype	perm.pval	MW pval
1	Vehicle vs. 0.1 nM	0.4	0.196
2	Vehicle vs. 1 nM	0.01	7.46e−06

Figure 4

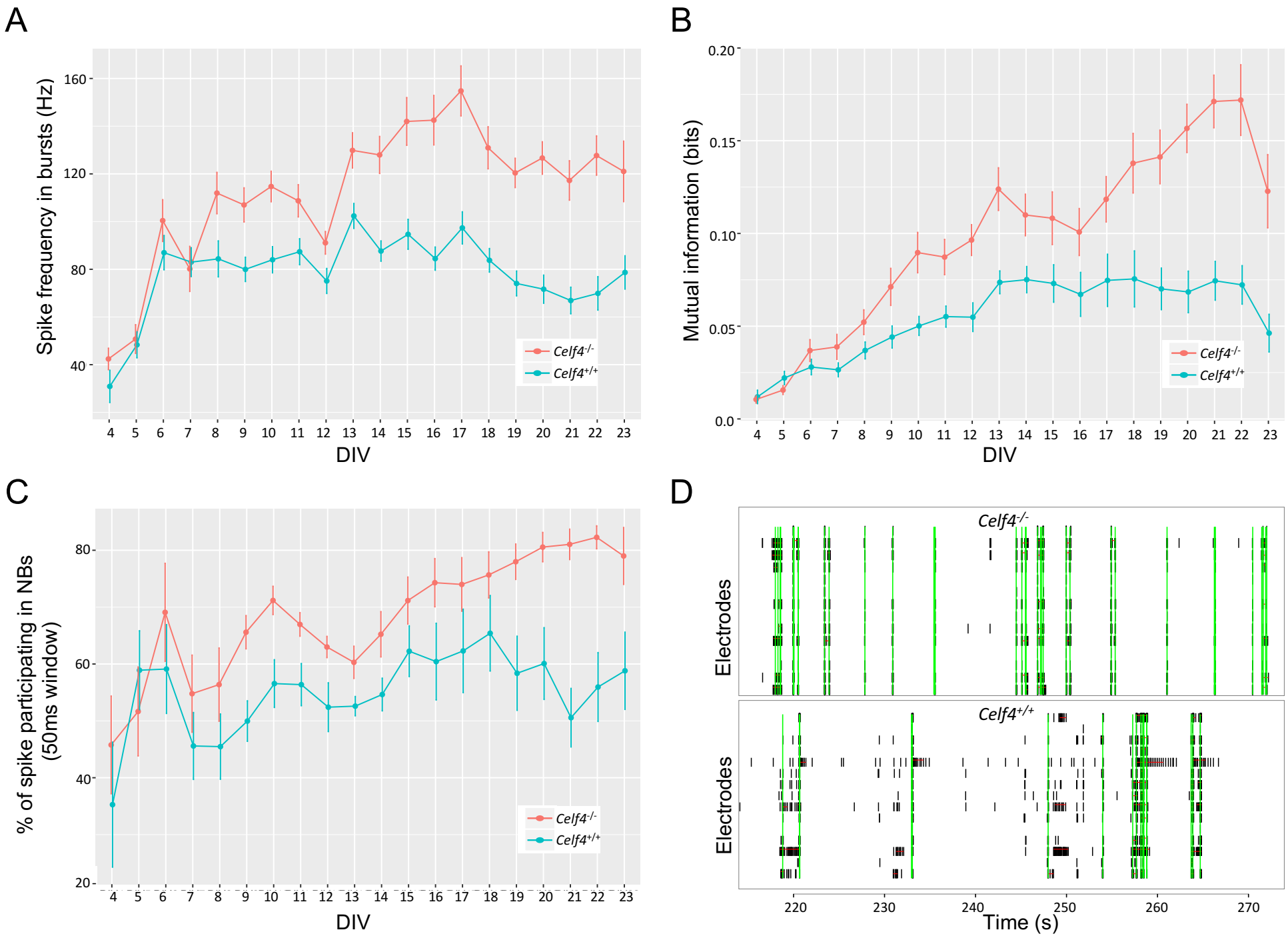

A
Spike frequency in bursts (Hz)
160
120
80
40
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
DIV
Celf4-/-
Celf4+/+
B
Mutual information (bits)
0.20
0.15
0.10
0.05
0.0
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
DIV
Celf4-/-
Celf4+/+
C
% of spike participating in NBs (50ms window)
80
60
40
20
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
DIV
Celf4-/-
Celf4+/+
D
Celf4-/-
Celf4+/+
Electrodes
Electrodes
220 230 240 250 260 270
Time (s)

Figure 5

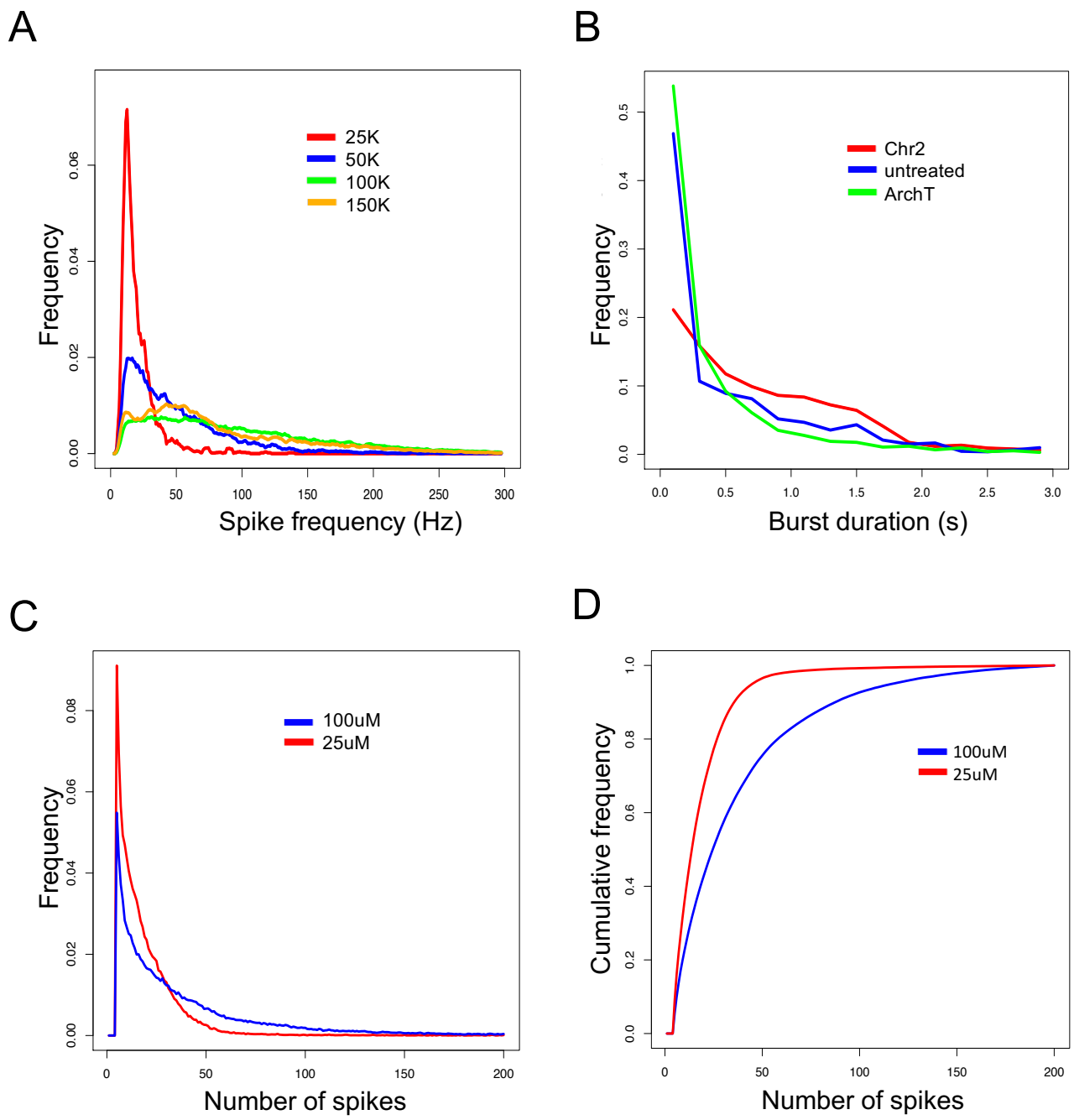

A
25K
50K
100K
150K
Frequency
0.00
0.02
0.04
0.06
0
50
100
150
200
250
300
Spike frequency (Hz)
B
Chr2
untreated
ArchT
Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Burst duration (s)
C
100uM
25uM
Frequency
0.00
0.02
0.04
0.06
0.08
0
50
100
150
200
Number of spikes
D
100uM
25uM
Cumulative frequency
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
Number of spikes

Figure 6