nature research

Corresponding author(s): Dr. Jens H Gundlach

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Data was collected with custom data acquisition software written in LabView (National Instruments, version 2018). Custom software is available upon request.
Data analysis	Data analysis was conducted using custom software written in Matlab (the Mathworks, version 2018a). Custom software is available upon request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-calculate sample sizes in the constant- and variable-voltage sequencing verification experiments. Sample sizes were chosen as the net yield in reads of several days of nanopore experiments. We terminated data collection on both experiments once we had measured read accuracy on multiple pores (N =10 and N = 21 for constant and variable voltage, respectively) and over enough bases to accurately estimate the per-base error rates (N = 9368 bases and N = 17309 bases for constant and variable voltage, respectively). Likewise, no pre-calculation was used to determine sample sizes in the DNA stretching experiments. Enough samples were collected to generate a consensus signal for the target DNA sequence.
Data exclusions	In the sequencing verification experiment, reads failing certain basic quality requirements and reads of non-target DNA strands were excluded as follows. Reads exhibiting substantial long-duration mid-read gating (spontaneous drops in conductance not related to DNA sequence) were discarded prior to sequencing. Reads not containing a sufficient length of target DNA (>40 bp), either due to premature release by the motor enzyme or due to the capture of a small fragment from the restriction digest, were discarded from the final accuracy analysis. Adapter dimers, in which the enzyme-loading and pore-threading adapters used to facilitate capture and reading of the target DNA bind together without a payload of the target sequence in between (Supplemental Note 14) were discarded, as they did not contain the DNA sequence under study. We also observed and discarded a population of longer reads that did not have any significant alignment against the target pET28a reference sequence but did align to the E. coli reference genome. As the target pET28a DNA was grown up in an E. coli host, we concluded that these longer reads were reads of genomic E. coli DNA rather than the target sequence under study, and thus excluded them from our final accuracy analysis. The inclusion conditions of read length and lack of mid-event gating were established prior to data collection. The existence of non-pet28a reads due to the presence of E. coli DNA was not foreseen, and the exclusion of these reads was determined after data were collected.
Replication	Our sequencing verification experiments used data from multiple days and multiple pores, and evaluated sequencing accuracy over a range of sequence contexts (as contained in the pET28a reference sequence). We found that both constant- and variable-voltage sequencing accuracies were stable across different days and pores. As sequence context is known to affect nanopore sequencing accuracy, both methods were tested on the same target DNA sequence to provide a fair comparison of their relative performance. Different DNA sequence contexts yielding a higher (lower) accuracy for one of the two methods would similarly yield a higher (lower) accuracy for the other method. As this work is concerned with the relative performance improvement between the two methods, and the performance improvement was demonstrated to be stable across days, pores, and the presented sequence contexts, we conclude that a similar improvement would be observed on different target DNA sequences as well.
Randomization	Samples were not allocated into experimental groups as randomization was not relevant to our study.
Blinding	Investigators were not blinded as to the DNA sequence during the sequence verification experiments. Knowledge of the DNA sequence under study was necessary to the design of the experiment and preparation of the DNA for nanopore sequencing (choice of restriction enzymes and sequence of adapter strands; Supplemental Note 14). The data analysis and basecalling processes were put in place with fixed parameters prior to knowledge of the DNA sequence for the final verification experiment. Parameters were not tuned to generate best results on the target DNA.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging