Creating complete sample-to-analysis workflows with the interactive protocol selector

The selector is designed to simplify experimental design by helping users navigate through the many extraction, library preparation and analysis options, creating bespoke end-to-end workflows

Contact: **publications@nanoporetech.com** More information at: **www.nanoporetech.com** and **publications.nanoporetech.com**

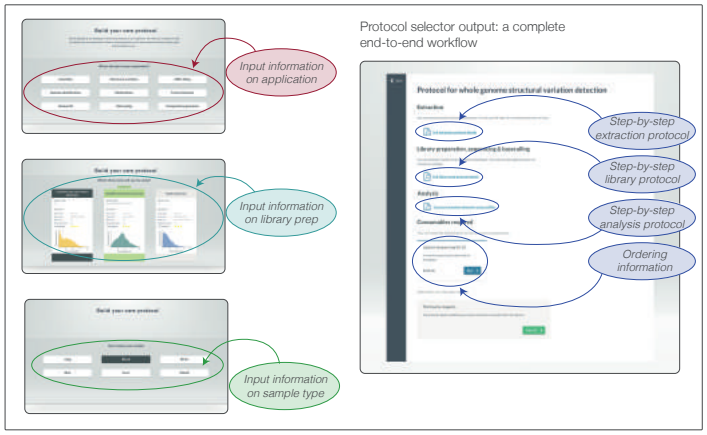
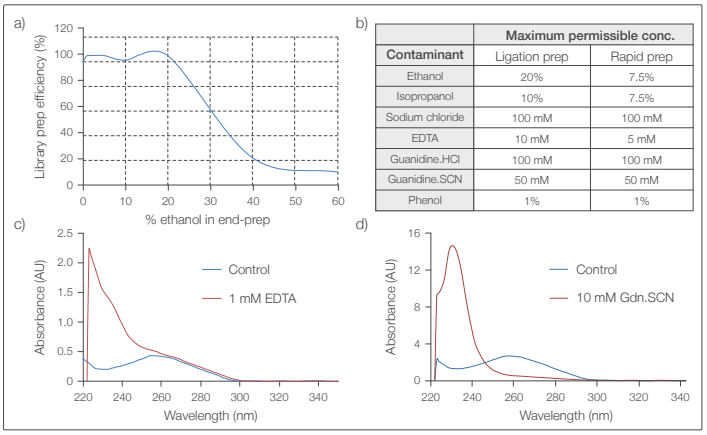


Fig. 1 End-to-end workflows created using the protocol selector



	Maximum permissible conc.	
Contaminant	Ligation prep	Rapid prep
Ethanol	20%	7.5%
Isopropanol	10%	7.5%
Sodium chloride	100 mM	100 mM
EDTA	10 mM	5 mM
Guanidine.HCl	100 mM	100 mM
Guanidine.SCN	50 mM	50 mM
Phenol	1%	1%

Fig. 2 Contaminants a) ethanol titration b) maximum concentrations c) and d) Nanodrop traces

A user-friendly, interactive protocol selector for bespoke workflow creation

We have built a protocol selector tool to help users select the best combinations of extraction and library preparation protocols with bioinformatics analysis workflows, for a wide range of sample types (Fig. 1). This selector is designed to help users obtain optimal results from a nanopore sequencing experiment, allowing for the requirements of the analysis and those of the user. The output is a complete and easy-to-follow workflow tailored to the specific experiment and desired analysis. This selector also presents users with QC data for expected DNA / RNA fragment length and purity to assist with potential troubleshooting, and explains the probable consequences of selecting alternative extraction and library prep protocols.

Cataloguing the signs and effects of contaminants introduced during extraction

A wide variety of chemicals is used for extraction of nucleic acids and traces of these substances can contaminate the final extract. This could have a deleterious effect on the yield after library prep, and on the sequencing chemistry. Titration experiments of all the common extraction chemicals have revealed that our library preps and flowcells are robust to fairly high concentrations of these chemicals, and these are listed on our website (Figs. 2a and 2b). Contaminants can be detected and quantified approximately using a Nanodrop (Figs. 2c and 2d). If the concentration of a contaminant in a sample exceeds the maximum permissible concentration, further clean-up steps can be performed.

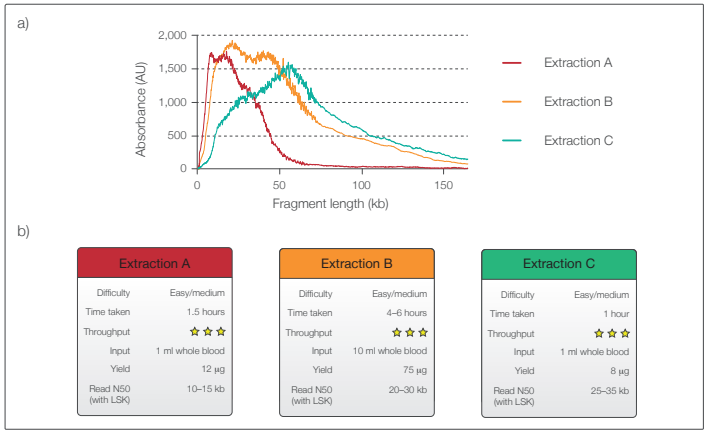
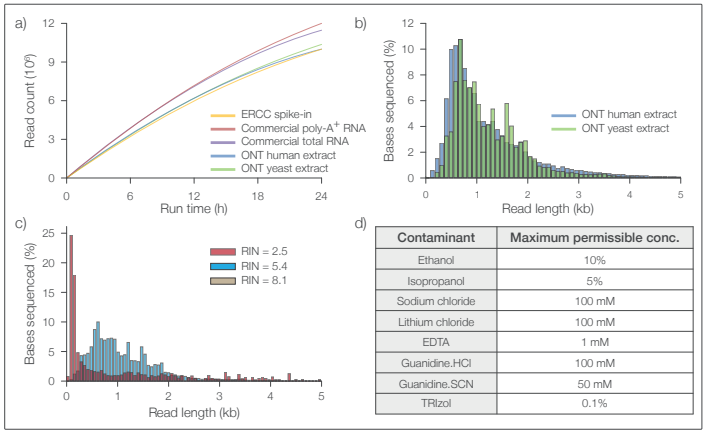


Fig. 3 Comparison of rabbit blood DNA extraction a) read length b) performance summary



Contaminant	Maximum permissible conc.
Ethanol	10%
Isopropanol	5%
Sodium chloride	100 mM
Lithium chloride	100 mM
EDTA	1 mM
Guanidine.HCl	100 mM
Guanidine.SCN	50 mM
TRIzol	0.1%

Fig. 4 RNA extraction a) throughput b) read length c) RNA integrity d) contaminants

Extraction methods differ substantially, in terms of performance and user experience

There are many extraction kits and protocols available, and these can be used with a wide variety of sample types. The skills required to get good results can differ substantially between kits, as can the purity, integrity and yield of the extracted nucleic acids. As a consequence, some extraction protocols are more suitable for a given downstream analysis than others. For instance, an extraction which yielded shorter fragments of DNA would not be recommended for *de novo* assembly of a large genome. For each sample type, the selector presents the user with typical fragment-length distributions for all tested extraction methods (Fig. 3a), and one extraction is recommended for the chosen analysis. The user is also presented with performance metrics for the other methods, allowing them to override the recommended choice, if desired.

Robust extraction methods and quality-control recommendations for RNA

As with DNA, we are investigating RNA extraction methods for a variety of samples. Following extraction, total RNA is prepared for sequencing using the PCR–cDNA Sequencing Kit (SQK-PCS109) and run on MinION, using commercial RNA samples as controls (Fig. 4a). We are releasing extraction protocols that can give high throughput and good read-length distributions (Fig. 4b). It is important to determine the RNA Integrity Number (RIN) before library preparation. The lower the RIN, the more degraded the sample, leading to an abundance of short reads in the sequencing run (Fig. 4c). Contaminants from the extraction can co-purify with the RNA and can perturb the efficiency of the library preparation process. We have determined the maximum permissible concentration of common contaminants in PCR–cDNA library preparation (Fig. 4d).