Nanopore sequencing for Illumina library control and pooling

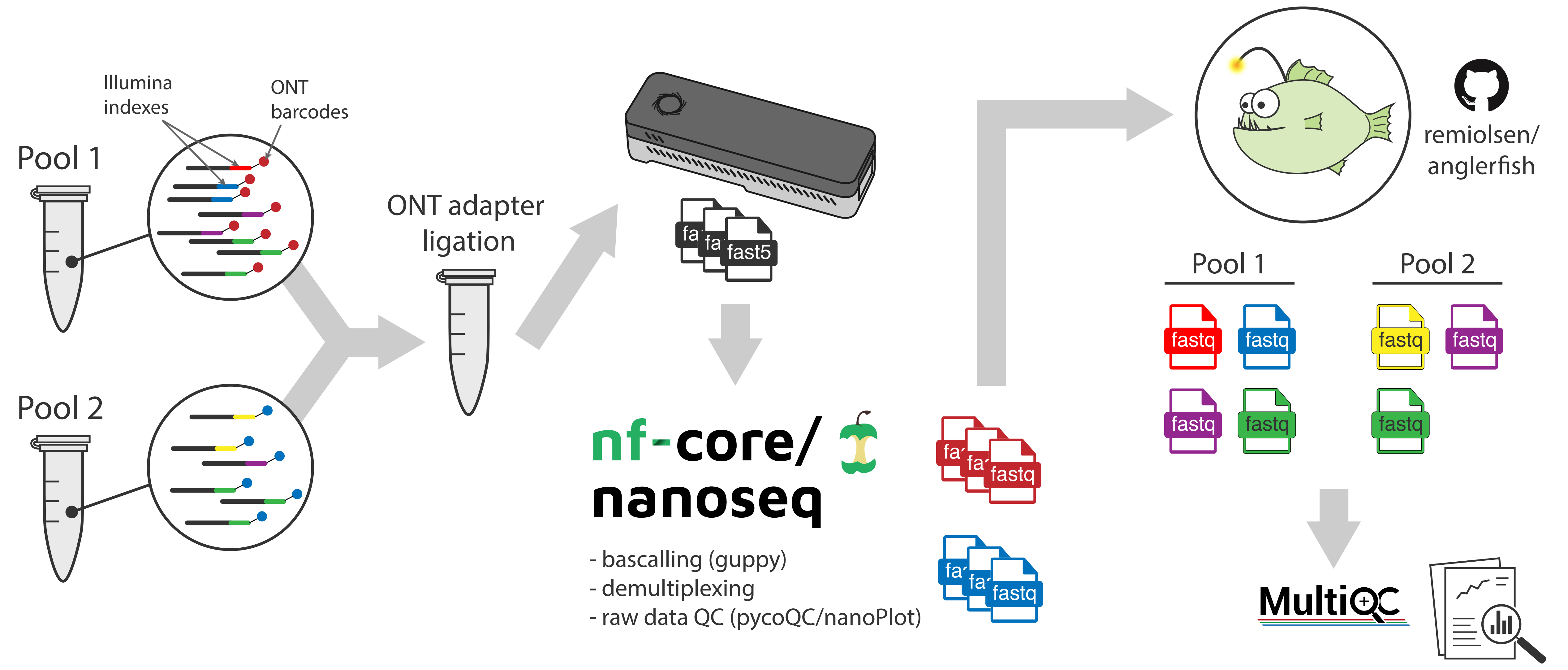
Remi-Andre Olsen[1], Franziska Bonath[1], Fanny Taborsak-Lines[2], Sara Sjunnebo[1], Carl Rubin[2]

[1]Science for Life Laboratory, National Genomic Infrastructure, Department of Biochemistry and Biophysics, Stockholm University
[2] Science for Life Laboratory, National Genomic Infrastructure, School of Engineering Sciences in Chemistry, KTH Royal Institute of Technology

Abstract

The increased capacity of Illumina sequencers allow pooling of large amounts of samples without compromising on sequencing depth. Flowcells for each run are very expensive, making pooling errors more costly. The increased number of samples also adds risk of clashing or misclassified indexes.

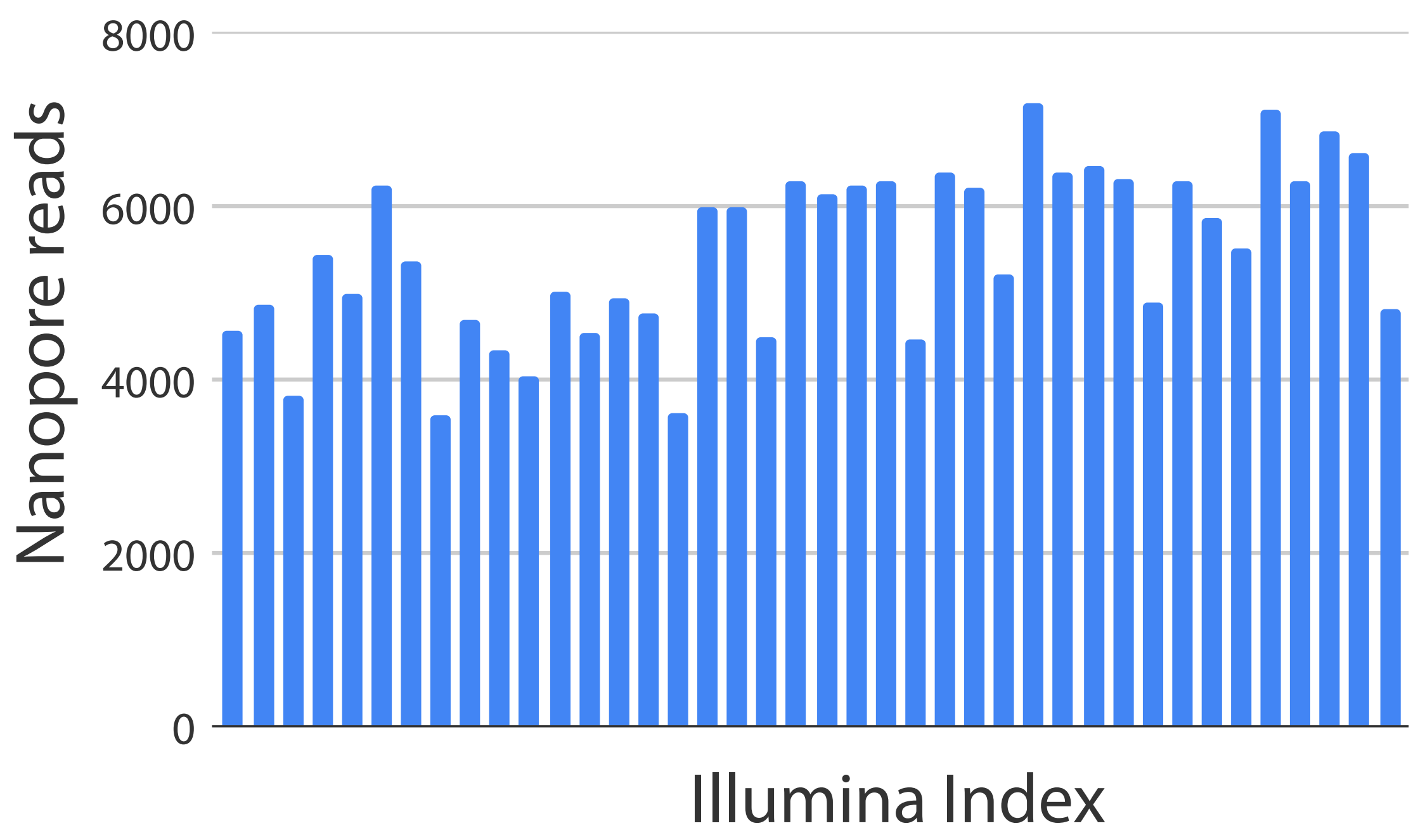
We propose Nanopore sequencing using Flongles as a cheap method to identify pooling errors and to gain additional metrics about the libraries before Illumina sequencing. The nanopore sequencing results can also be used as a low depth preview of the data to be expected in the full sequencing run on expensive Illumina flow cells.



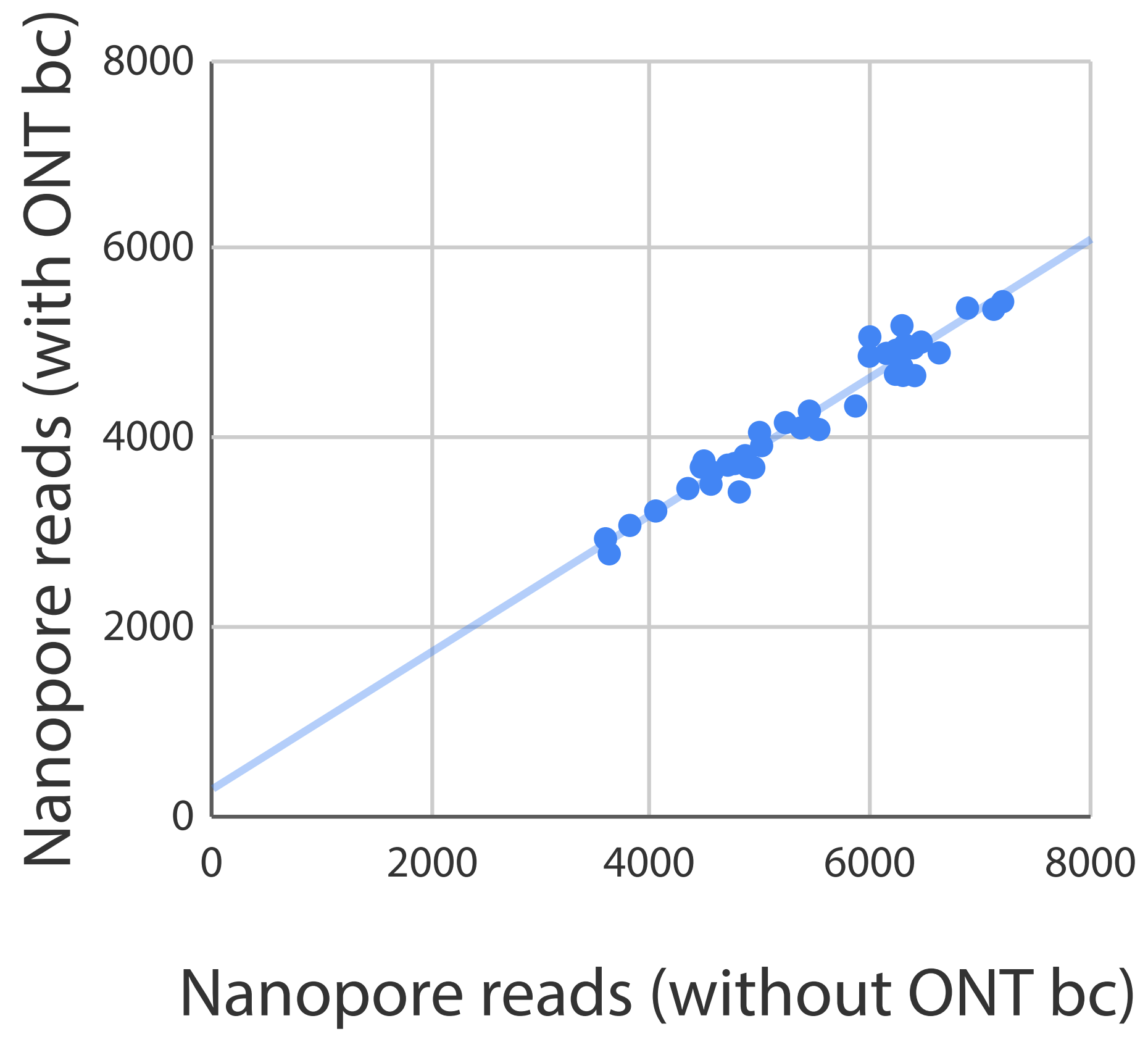
- bascalling (guppy)
- demultiplexing
- raw data QC (pycoQC/nanoPlot)

Detection of Illumina indexes using Anglerfish

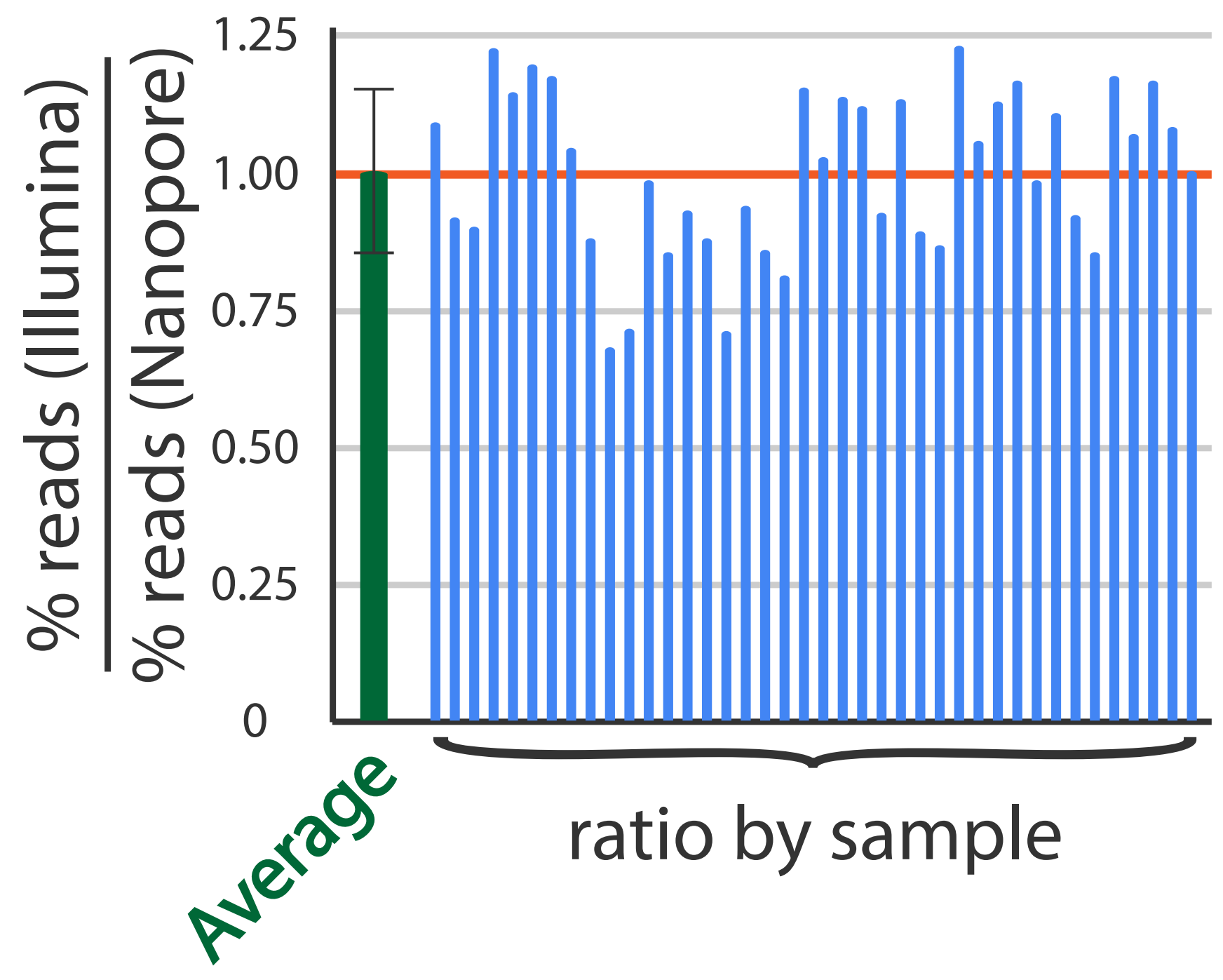
Illumina Pool:	Truseq, Stranded total RNA
Number of Samples:	40
total reads (nanopore):	676,292 (100%)
reads passing raw data QC:	536,968 (79%)
reads assigned to Ill. indexes:	219,061 (32.4%)



ONT barcoding does not affect Illumina index detection



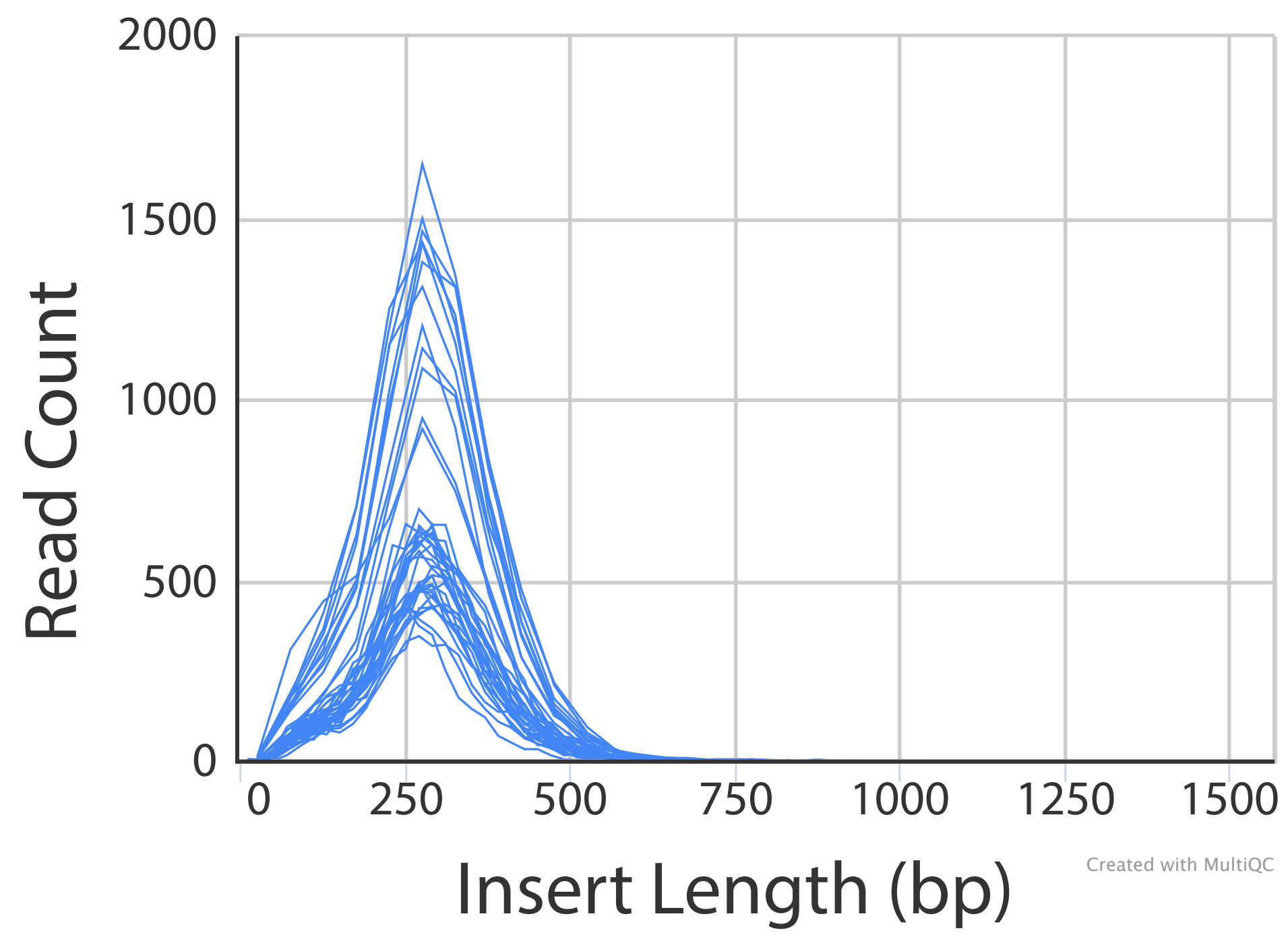
Comparing Nanopore and Illumina read abundance



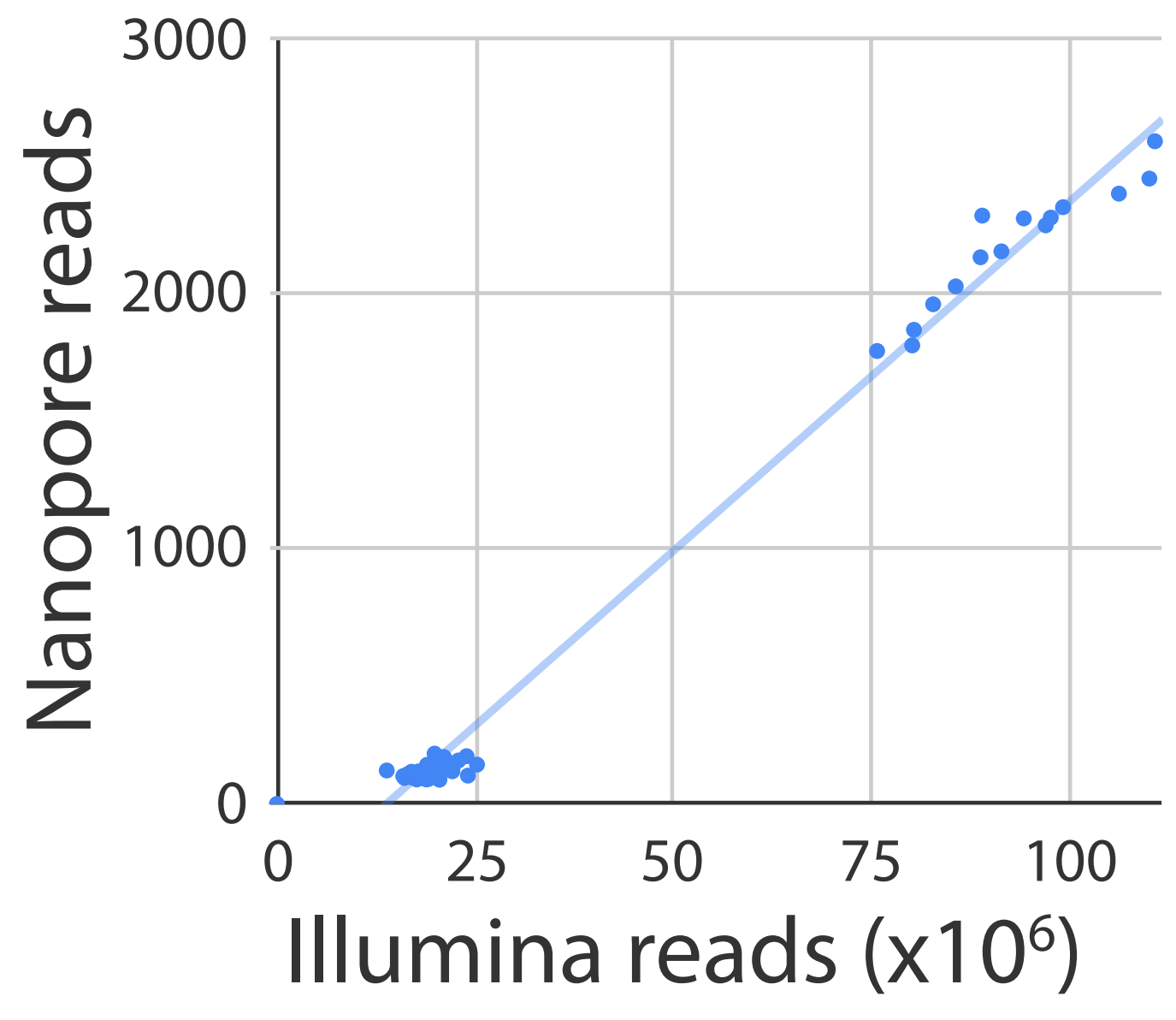
*Error bar represents the standard deviation

Example downstream uses

Insert length analysis

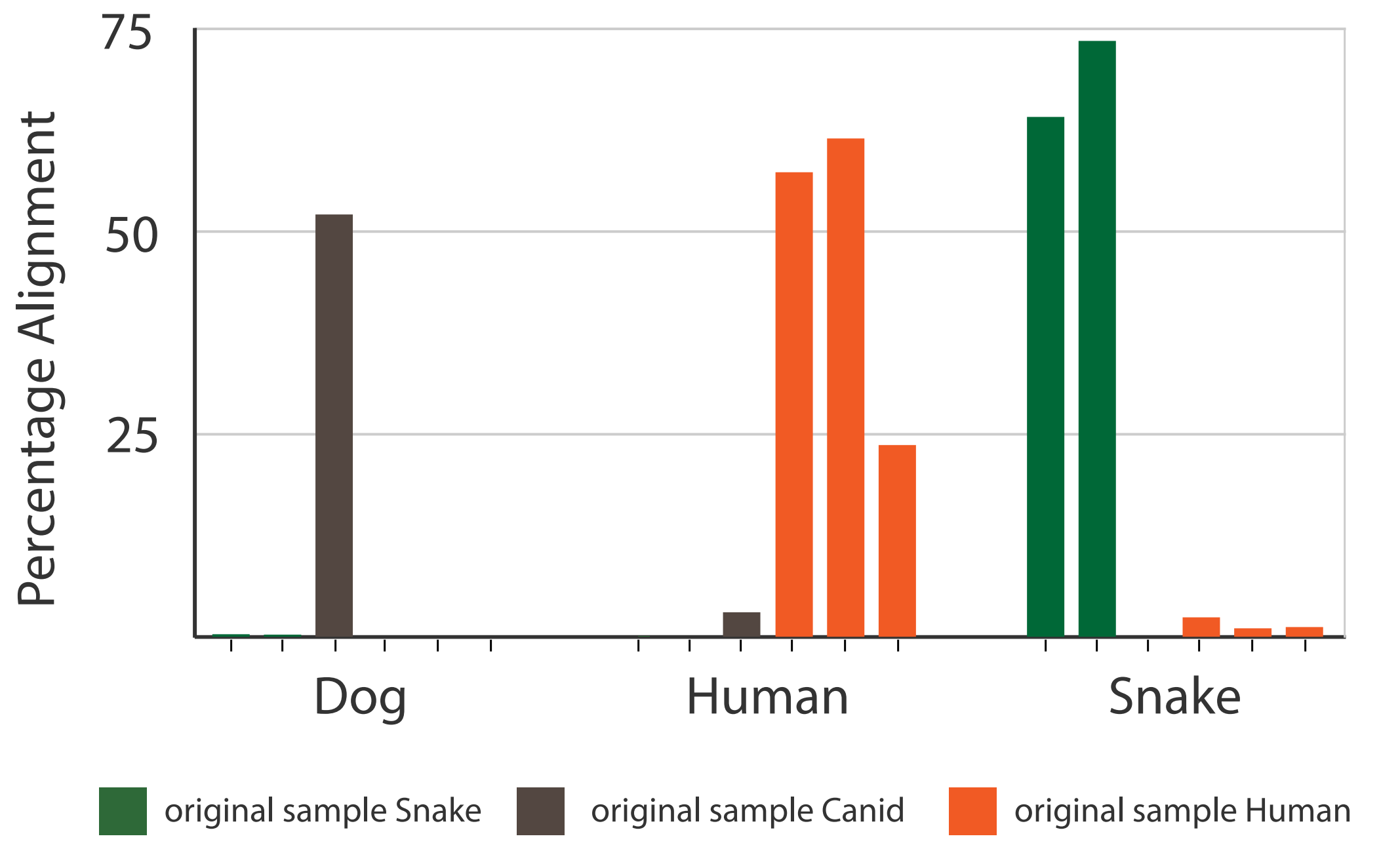


Indentification of pooling errors



Sample identity confirmation by mapping to target genomes

Hits on target genome only



https://nf-co.re
@nf_core
@ngisweden