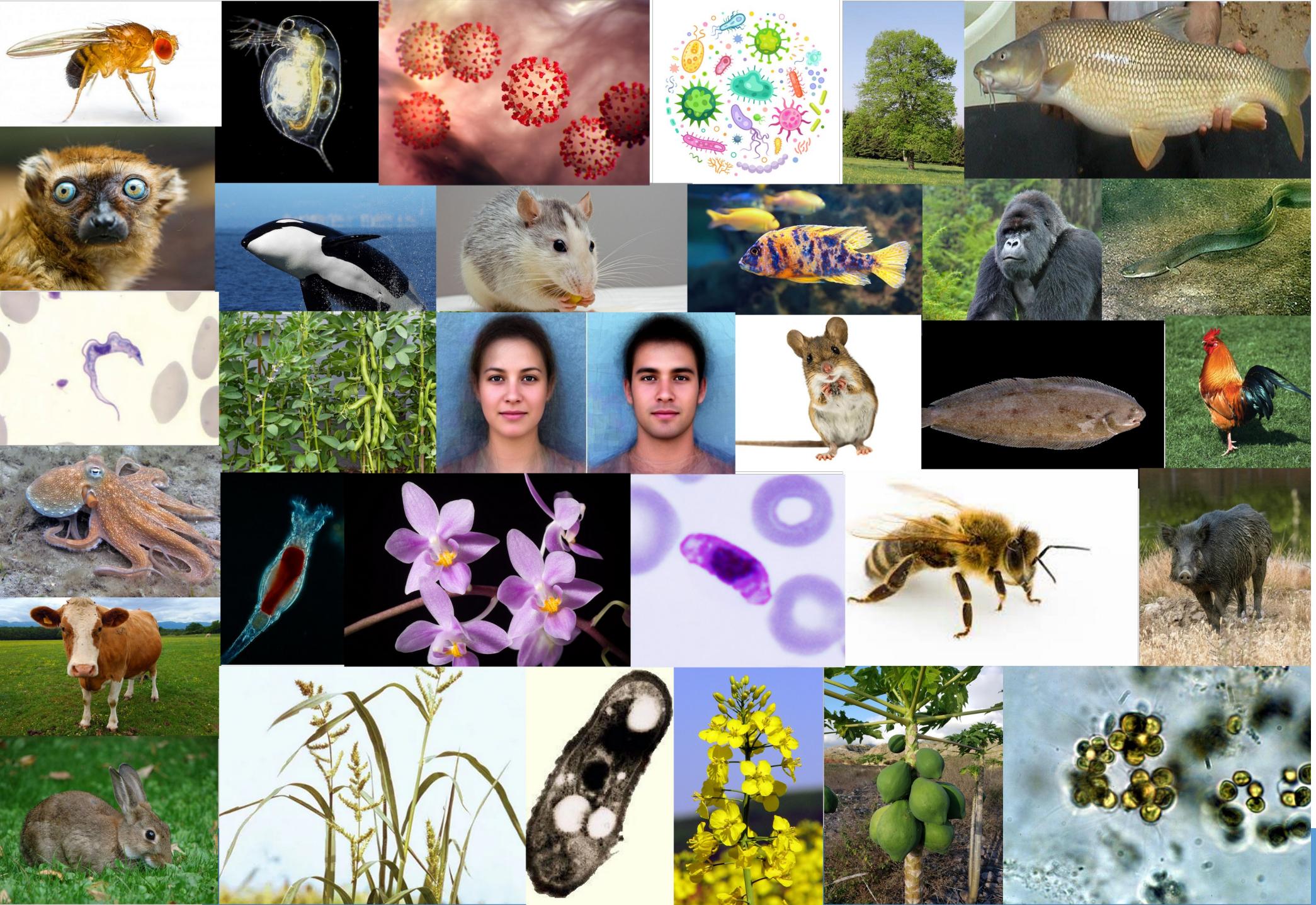


Long read sequencing

When to think about it and what to think about





Google Search

I'm Feeling Lucky

Google offered in: [Nederlands](#) [Français](#) [Deutsch](#)

Belgium

About Advertising Business

Gmail Images

<!DOCTYPE html>

```
html itemscope itemtype="http://schema.org/WebPage" lang="en-BE" coupert-item="9AF8D9A4E502F3784AD24272D81F0381" style="scroll-behavior: unset !important;">
<head>
  <meta charset="UTF-8">
  <meta content="origin" name="referrer">
  <meta content="Anm+hhtuh7NJguqSnXHEAIqqMaV+GXCKs8WYXHJKF716AeYMj+wO+f190dDqFnJTk9t0492Dykvxx4jpvFbxnA8AAABseyJvcmlnaW4i0iJc</meta content="images/branding/googleleg/1x/googleleg_standard_color_128dp.png" itemprop="image">
  <title>Google</title>
  <script src="https://apis.google.com/_scs/abc-static/_js/k=gapi.gapi.en.CzrNR.../sv=1/d=1/ed=1/rs=AHp0oo8xPbrtpW2bPUIcgU2ac</script>
<script nonce=">
  (function(){var _g={kEI:'CF1KZcOuBoGghbIP1pG3mAM',kEXPI:'31',u:'ea888a19',kBL:'1zo0',kOPI:89978449};(function(){var a;(nu
  var hthis||self;function l(){return void 0!==window.google&&void 0!==window.google.kOPI&&0==window.google.kOPI?window.g
  function t(a,b,c,d,k){var e="";-1==b.search("&ei=")&&(e="&ei="+p(d),-1==b.search("&lei=")&&(d=q(d))&&(e=="&lei="+d));d=
  document.documentElement.addEventListener("submit",function(b){var a;if(a.b.target){var c=a.getAttribute("data-submitref")
  var h=this||self;var k=window.performance;function m(a,b,d,c){a:{c:void 0==c?11:c;for(var e=a;e&&e!=#t
  function n(a){return"none"==a.style.display?!0:document.defaultView&&document.defaultView.getComputedStyle?(a=document.c
  function p(a,b,d,c,e){var f=e(a),l=f.left+(d?0>window.pageXOffset),q=f.top+(d?0>window.pageYOffset),r=f.width,t=f.height,
  var g=this||self;function h(a){try{a()}catch(b){google.ml(b,11)}}google.caft=function(a,b){null==google.aftQh(a):(googl
  function G(a,b){var c=google.timers[b||"load"];b=c.m;if(!b||!b.pr){var d=m()?:F("qsubits");0<d&&(b=F("fbts"),0<b&&c.t.s
  function ja(a){var b=void 0!;b&&a&&(b=j(a,!1));b&&a&&(google.c.gscp=j(a,!1));!b&&ca&&(b=j(a,!0));var c=a.parentElement
  function ra(){var a=Q==P,b=0==N;a=a&b:a==yM==ma:a;S&&a&&google.c.u("il",V)}
  function Y(){if(!R){var a=Q==P,b=0==N,c=x&&oa==na;a&&(google.c.e(V,"aft","1"),go...
  </script>
<script nonce=">
  (function(){google.xjs={ck:'xjs.hd.SRgRPd06aEw.L.W.O',combam:'CAAAAAAAAAAAAAAAAEEAAAAIgaCicAnoAAgQAQAAAAAgYIUUQjAAQMB
  </script>
<script nonce=">
  (function(){google.kEXPI='0,3300114,18,13189,9432,7443,32625,2272,1527,1037,2112,254,427,82,200,46,22,116,144,329236,653,
  (function(){window.google.xjsu='/xjs/_js/k/x3dxjs.hd.en.fG1V7-UH_5g.0/am\x3dCAAAAAAAAAAAAAAAAEEAAAAIgaCicAnoAAgQAQAAA
  0715Yw/ee\x3dcE90b:ws9T1c;qddgKe:x4FYXe,d7YSfd;yxTchf:KUM7Z;dt10hd:1LQWFe;eHDf1:ofjVkb;qaS3gd:yiLg6e;nAF3:NTMZac,s3954;
  </script>
<script defer src="/xjs/_js/k=xjs.hd.en.fG1V7-UH_5g.0/am=CAAAAAAAAAAAAAAAAEEAAAAIgaCic..s\vmc/m=cdos,hsm,jsa,mb4ZUb,d,csi
<script nonce=">
  (function(){window.google.erd={jsr:1,bv:1896,sd:true,de:true}});(function(){var sdo=false;var mei=10;
  var hthis||self;var k,l=null;!(k=hh.mei)?k:1,n,p=null;!(n=hh.sdo)?n!:0,q=0,r,t=google.erd,v=t.jsr;google.ml=function(a,b,c
  b(t.bv);var f=a.lineNumber;void 0!=f&&(c+="&line="+f);var g=a.fileName;g&&(0<g.indexOf("-extension:")&&(e=3),c+="&script
  [[null,null,"https://www.gstatic.com/og/_js/k=og.qtm.en_US.I43EUmH7Doc.2019.0/rt=j/m=qabrqql,q_dnp,qcwid,qbg,qbd,a
  try{
    _-F_toggles_initialize=function(a){("undefined"!=typeof globalThis?globalThis:"undefined"!=typeof self?self:this).-_F_t
  /*

    Copyright The Closure Library Authors.
    SPDX-License-Identifier: Apache-2.0
  */
  var fa,la,oa,pa,xa,ya,za,Aa,Ba,Da,Ea,Fa,Ia,Xa,Wa,$a,bb,ab,cb,db,lb;_.aa=function(a,b){if(Error.captureStackTrace)Error.ca
  _.ha=function(){return da?!!ea&&0<ea.brands.length:!1};_.ia=function(){return _.ha()?!1:_t("Opera")};_.ja=function(){ret
  la=function(){return _.ha()?fa("Chromium"):_t("Chrome")||_t("CriOS")}&&(!_.ha()?:_t("Edge"))||_t("Silk")};_.na=fuc
  _.ua=function(a){let b="",c=0;const d=a.length-10240;for(;c<d;b+=String.fromCharCode.apply(null,a.subarray(c,c+10240));
  Aa=function(a,b){b[_.u]=(a[34]&-14557);Ba=function(a){a>a>14&1023;return 0==a?536870912:a};Da=function(a){return !(a|!
  Ia=function(a,b){b=.Ha?bf(.Hal:void 0)&&(a!.Hal=.wa(b));Ka=function(){const a>Error():Ja(a."incident"): .ba(a)};
```

Sources of genetic information

DNA

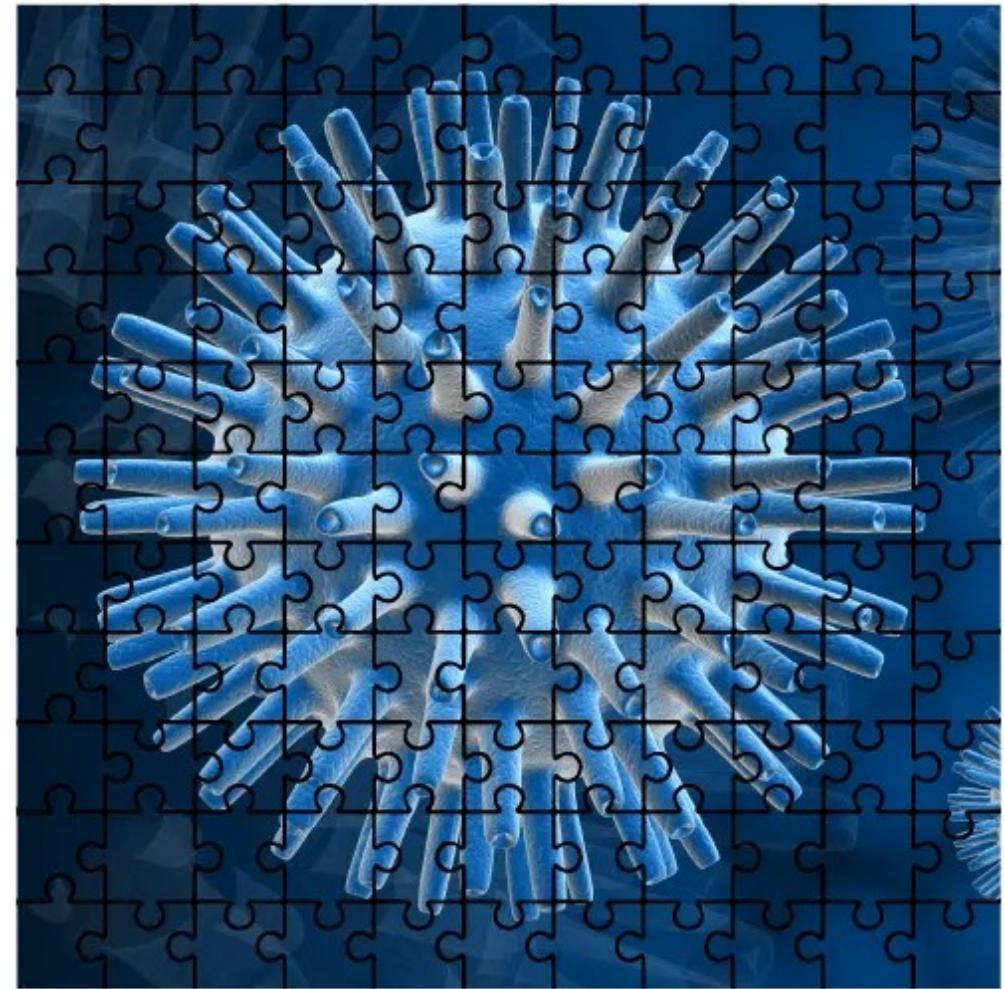
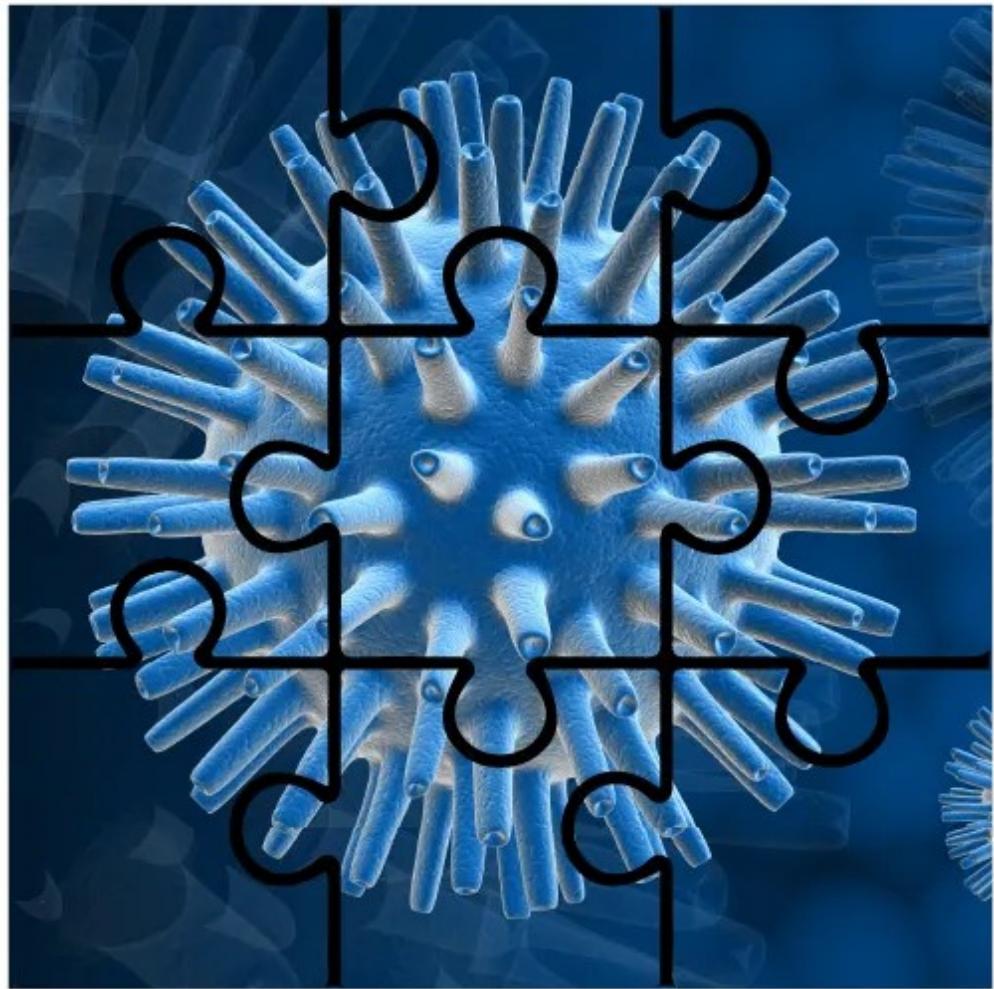
- SNVs and indels
- Tandem repeats
- Structural variation
- Segemental duplicatons
- Pseudogenes
- Methylation et al.
- Epigenetic marks
- phasing

RNA

- Expression levels
- Different isoforms
- Different RNA forms
- Full length transcripts

single cell genomic

- Heterogeneity of DNA mutations
- Cell type specific isoforms



What is long read?

- Sequencing but not short reads
 - 1kb or way longer in general
- Linked reads, long reads, optical genome mapping
- There is no ‘best’ technology
- Main technologies
 - Nanopore
 - Pacbio
 - Bionano



No ‘one ring’ to rule them all

Nanopore sequencing

- MinION
- GridION
- PromethION
- SmidgION



Oxford Nanopore sequencing

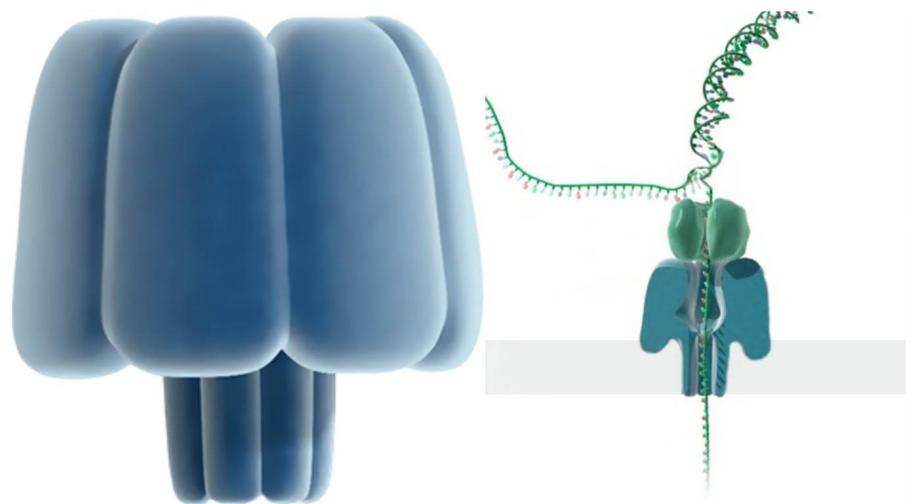


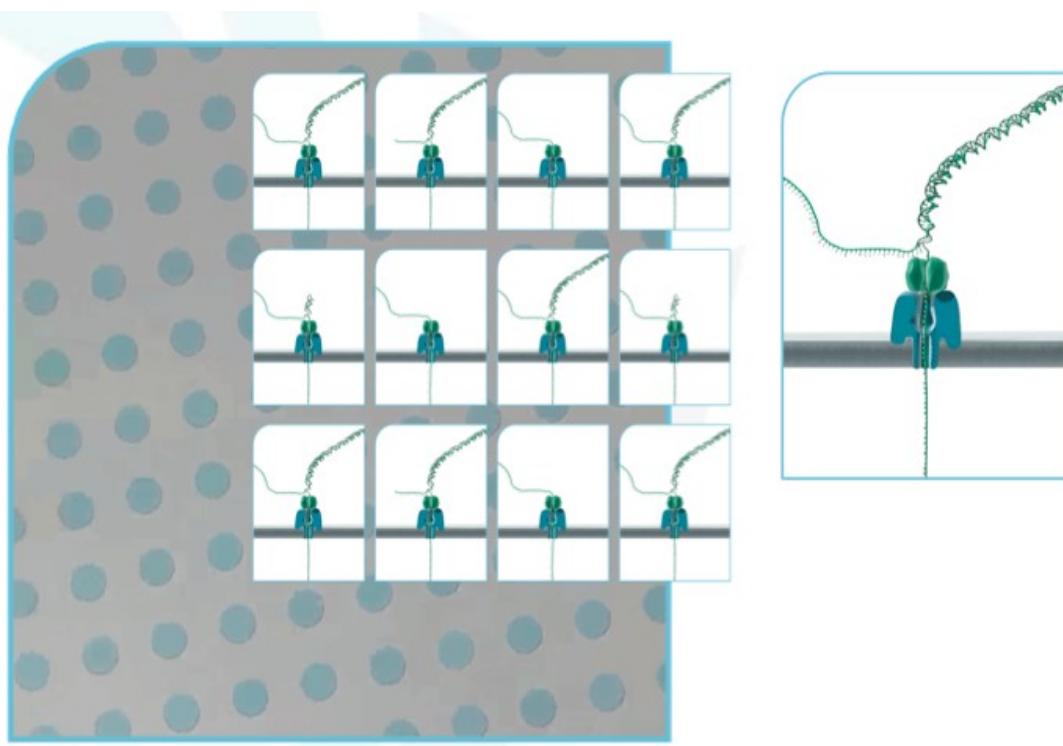
Tether keeps DNA fragment on the membrane leading to a ~20K fold higher DNA concentration close to the pore.

Motor protein unwinds DNA and ratchets it though the pore.

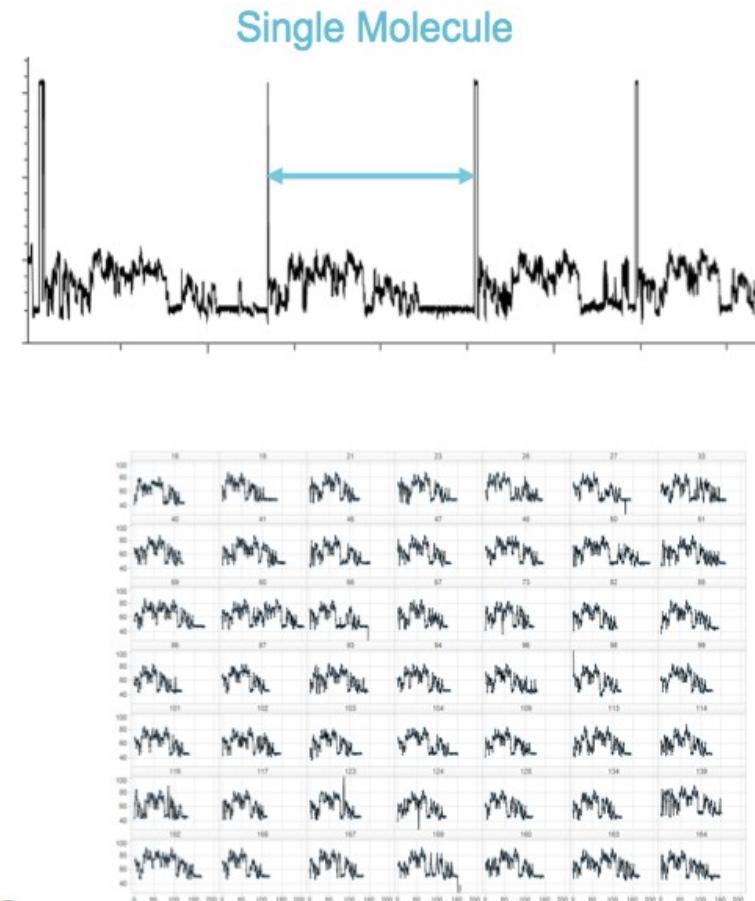
Abasic nucleotides in the hairpin are a recognition point.

Brake protein prevents the motor protein from zipping through the complement strand.





- Data acquired as full length reads – real time
- Data throughput = No. pores x average speed/pore



From squiggles to sequencing

- New basecaller: extract more correct information from squiggles
- Training of base caller for methylation data

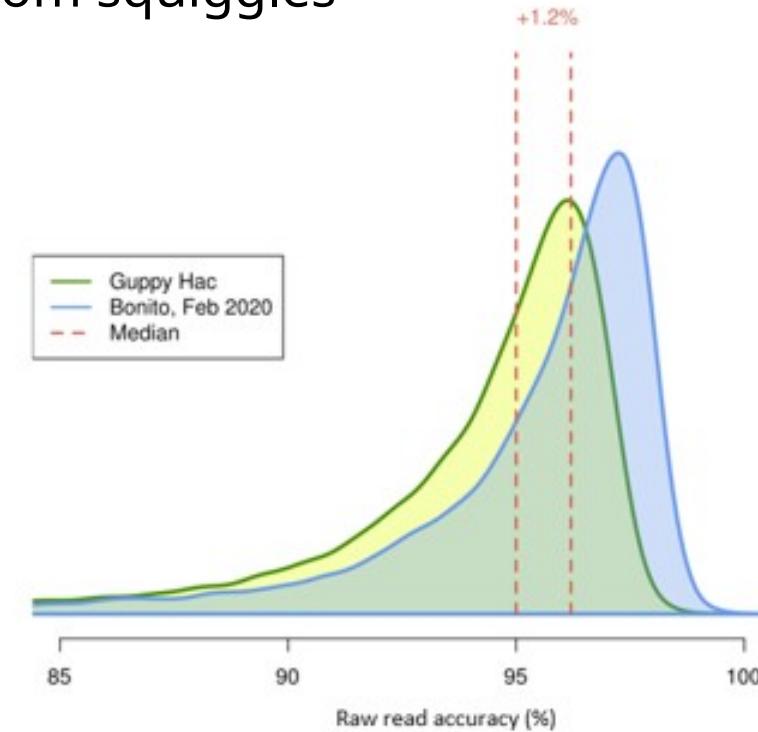
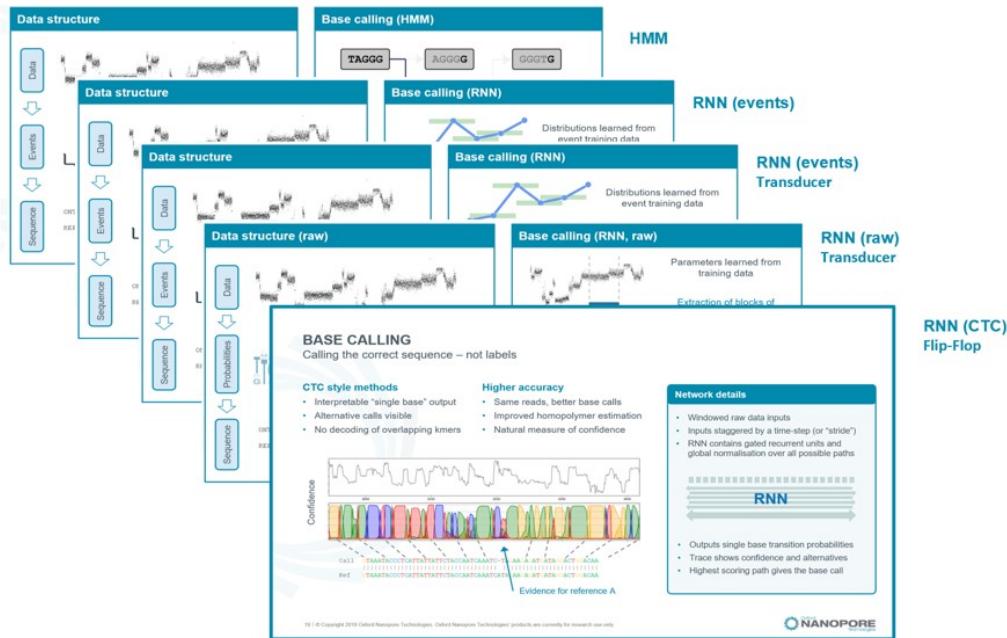
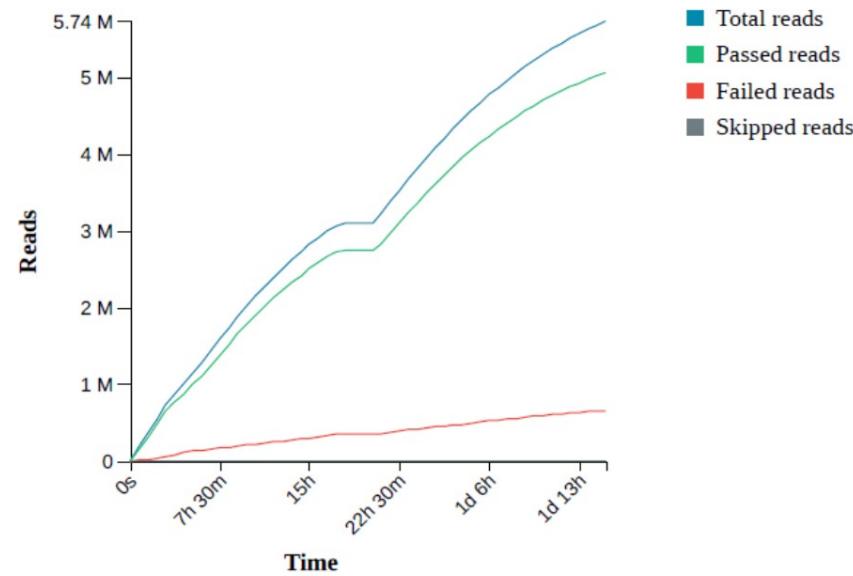


Figure 1: Raw read accuracy of Bonito basecaller on the human reference genome NA12878 against high-accuracy Guppy, currently integrated into MinKNOW onboard nanopore devices.

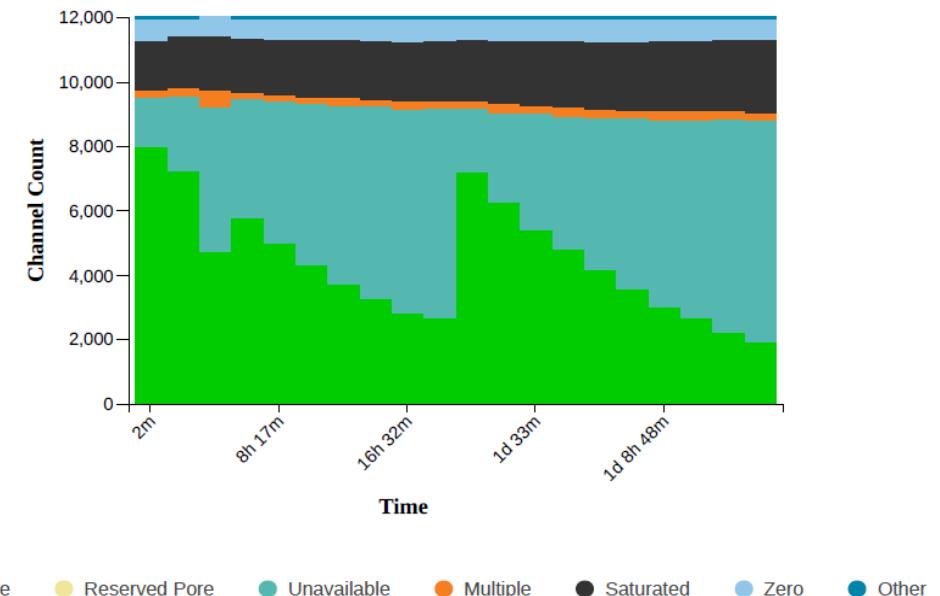
Run QC on Nanopore

- Every flowcell is different
- Nuclease wash (and refueling) increases output

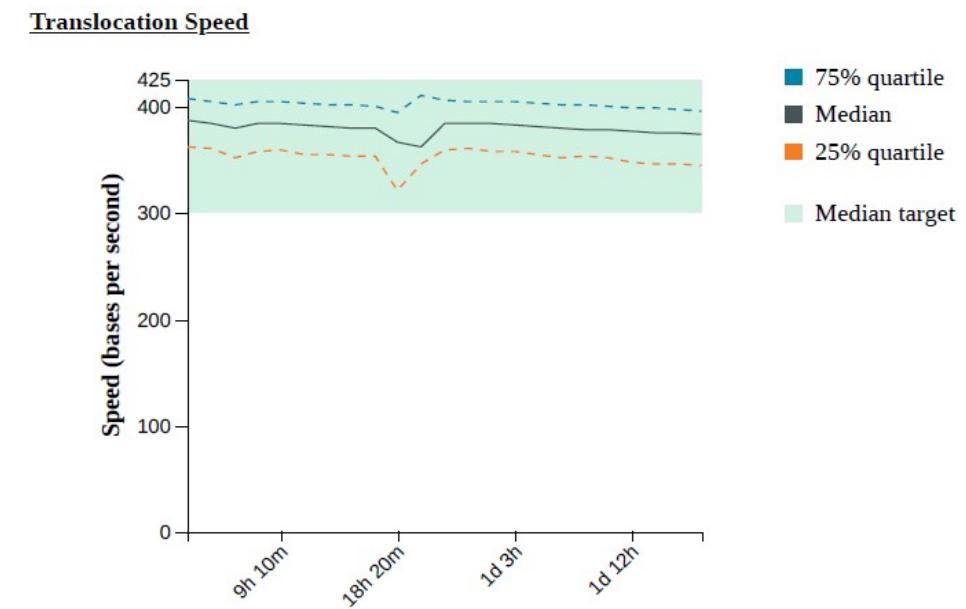
Cumulative Output Reads



Mux Scan Categorised

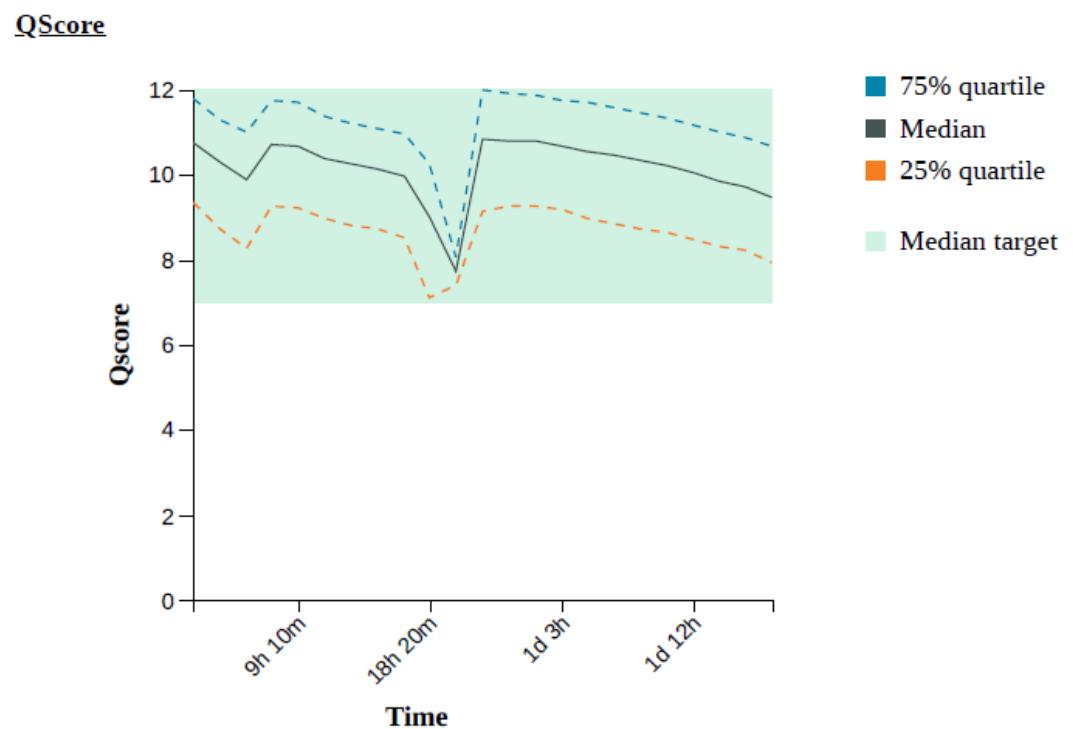


Translocation speed and pore status



Run QC on Nanopore

- Q-score need to be stable
- New flowcell chemistry improves Q-score
- Barcode selection ☾ select high quality door



Pacbio Revio

100M

ZMW / run

360 Gb

HiFi yield per run

24 hr

Sequencing time

15-18 kb

Read length

5mC

DNA methylation

90% \geq Q30

Base quality

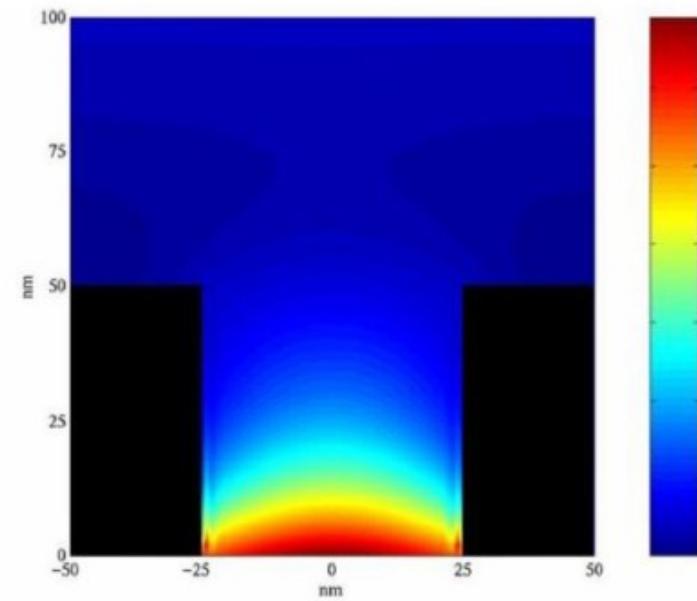
Revio™ system



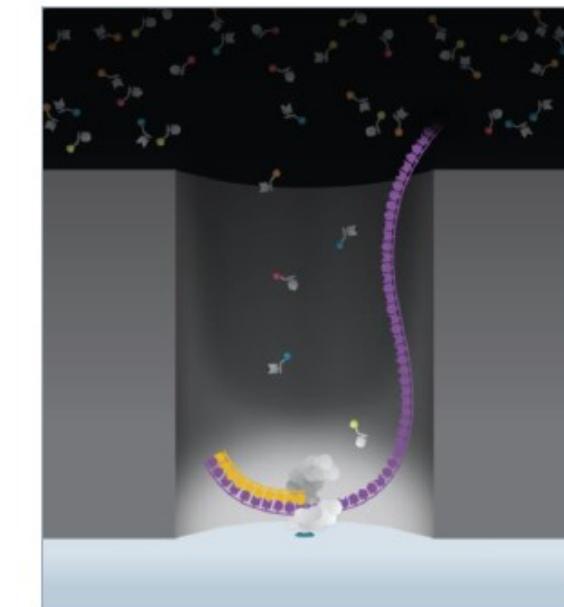
ZERO-MODE WAVEGUIDES (ZMWS) ENABLE HIGHLY SENSITIVE OPTICAL-BASED DETECTION OF SINGLE MOLECULES



A. Each ZMW is illuminated from below and works *via* the same principle as a metallic microwave oven door screen



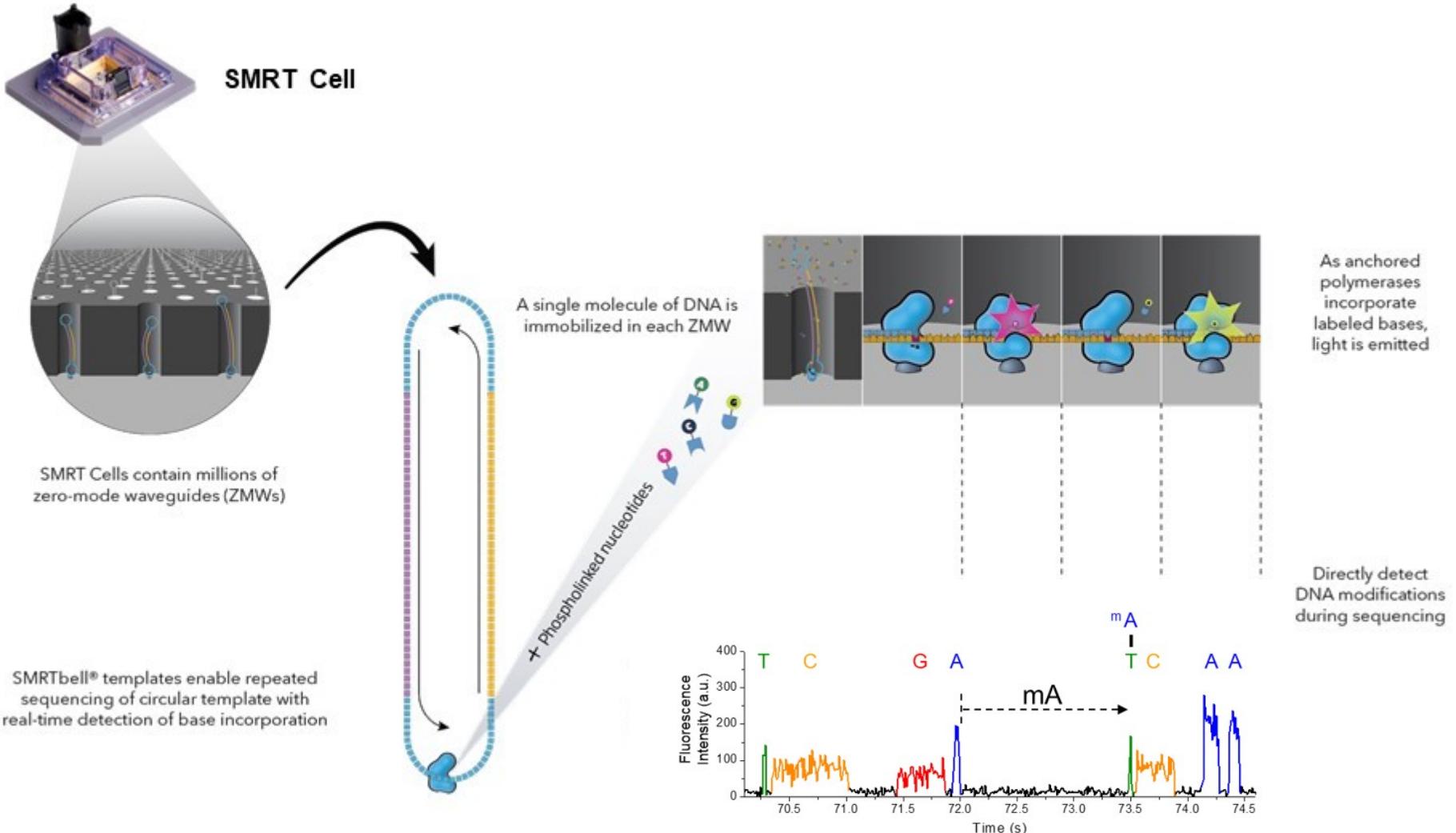
B. Because the illumination wavelength is much greater than the ZMW diameter, light penetrates only nanometers into the ZMW, providing a vanishingly small illumination volume



C. A DNA polymerase-template complex is immobilized onto the bottom surface of the ZMW

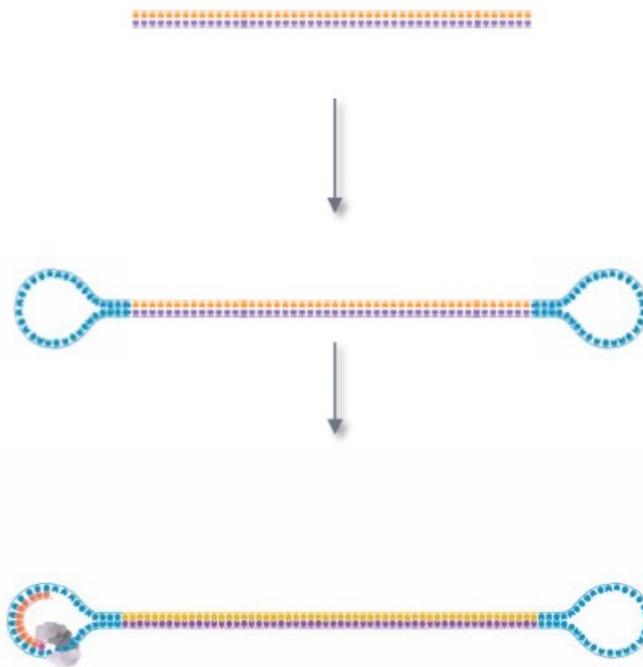
The illuminated volume is small enough to enable real-time observation of DNA synthesis at the single-molecule level

SINGLE MOLECULE, REAL-TIME (SMRT) SEQUENCING

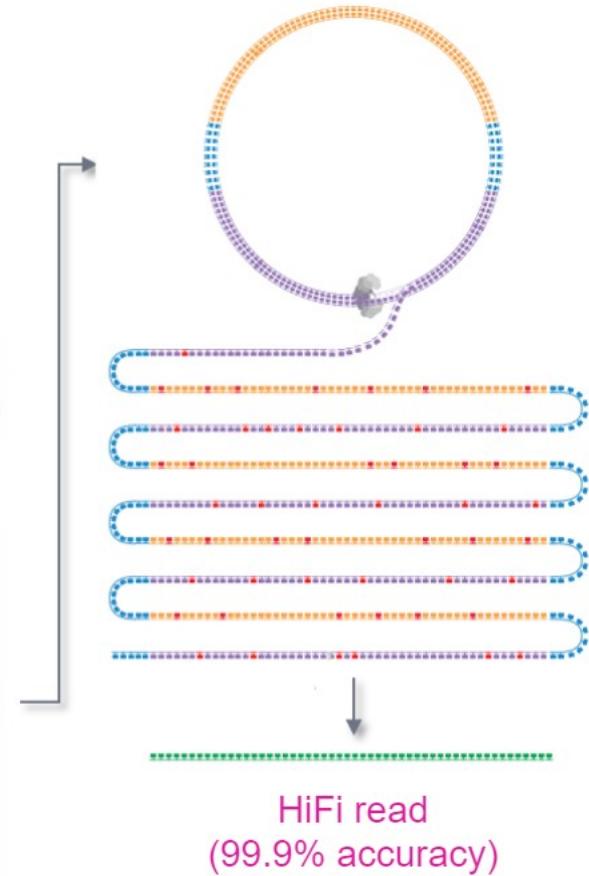


Circular consensus sequencing generates HiFi reads

Start with high-quality dsDNA



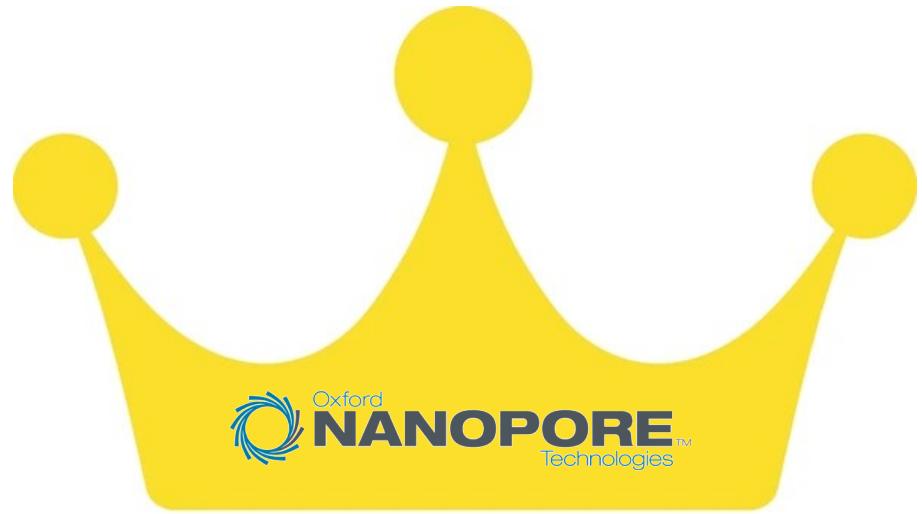
Anneal primers and bind DNA polymerase



Benefits of HiFi reads

- Long read lengths up to 25 kb
- High read accuracy ≥99.9%
- Easy library preparation
- Low coverage requirements
- Small file sizes to minimize compute time
- A single technology solution for a range of applications

The main differences



length



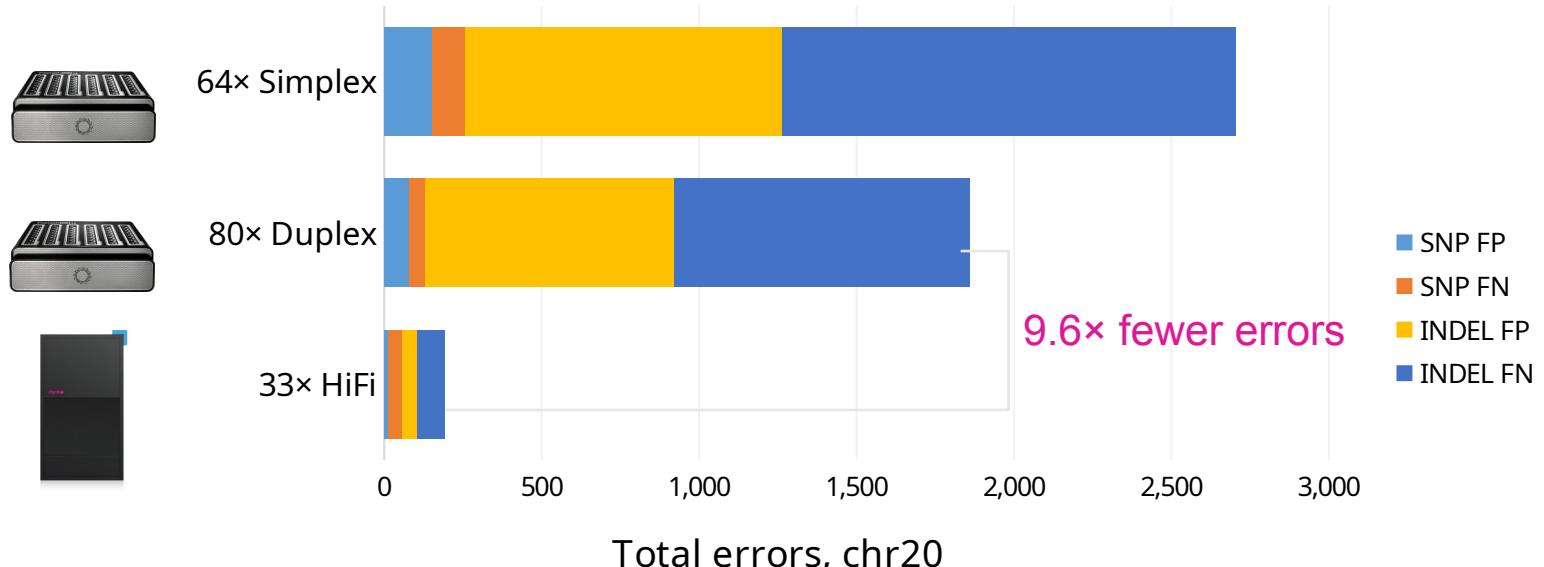
(indel) accuracy

Accuracy



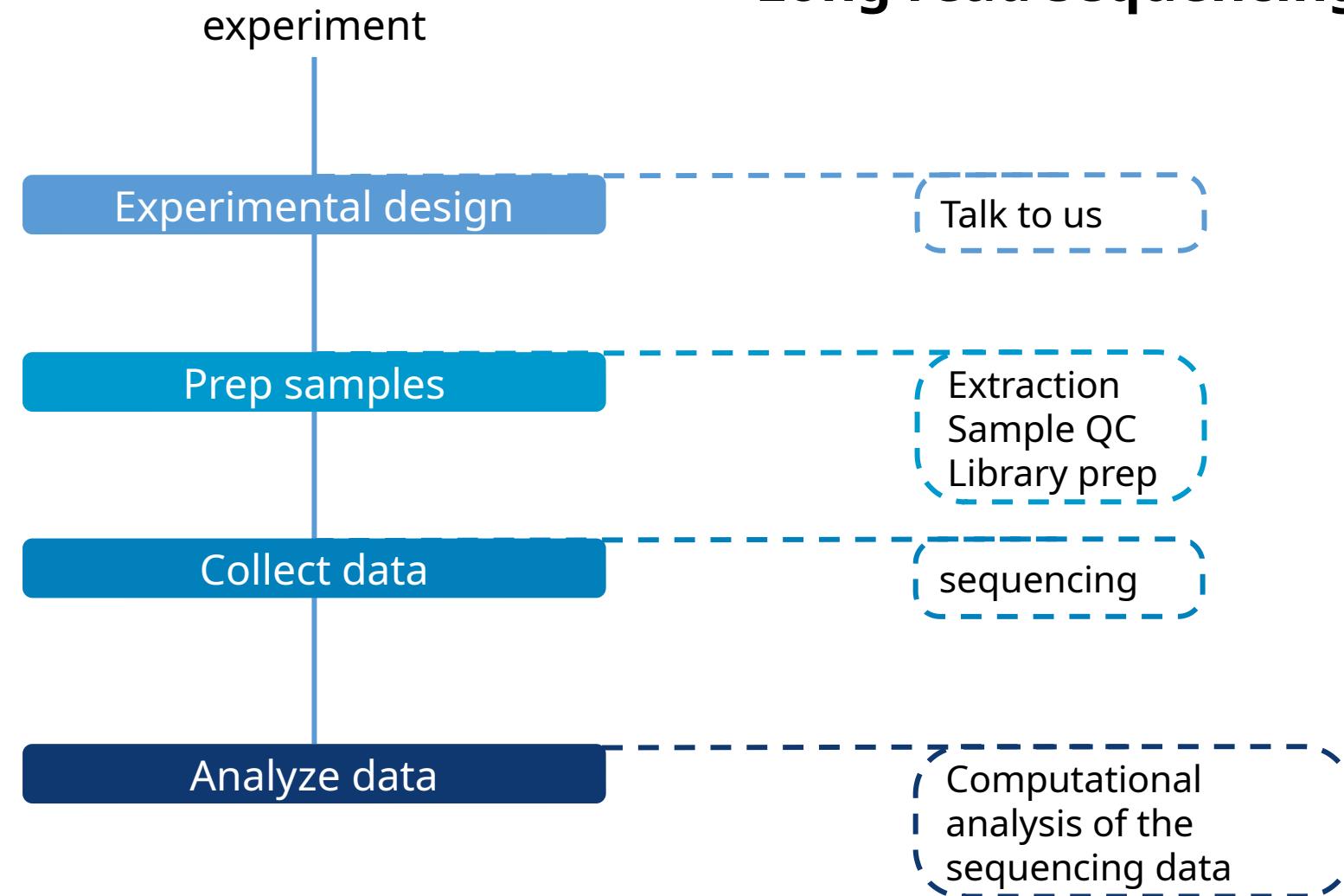
HG002 GIAB
v4.2.1¹

 DeepVariant
v1.5, February 2023



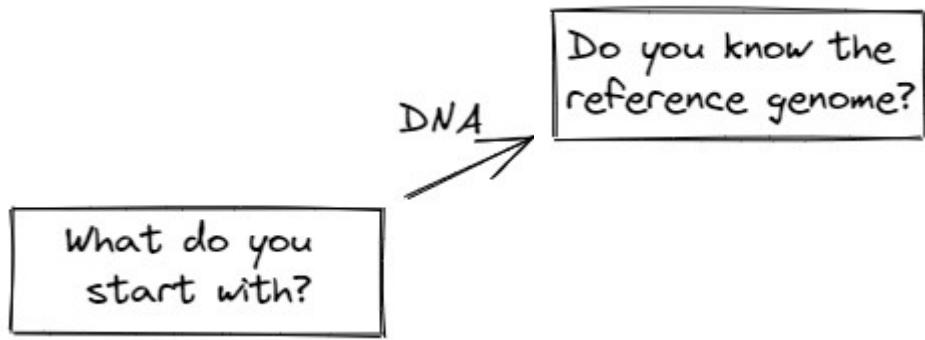
Sequencing			SNV, chr20				INDEL, chr20			
Platform	Reads	Coverage	F1	TP	FP	FN	F1	TP	FP	FN
ONT PromethION	Simplex ²	64x	99.82%	70,061	152	105	88.40%	9,189	1,009	1,439
ONT PromethION	Duplex ³	80x	99.91%	71,280	79	53	92.39%	10,316	788	940
PacBio Revio	HiFi ⁴	33x	99.94%	71,288	14	45	99.21%	11,494	48	86

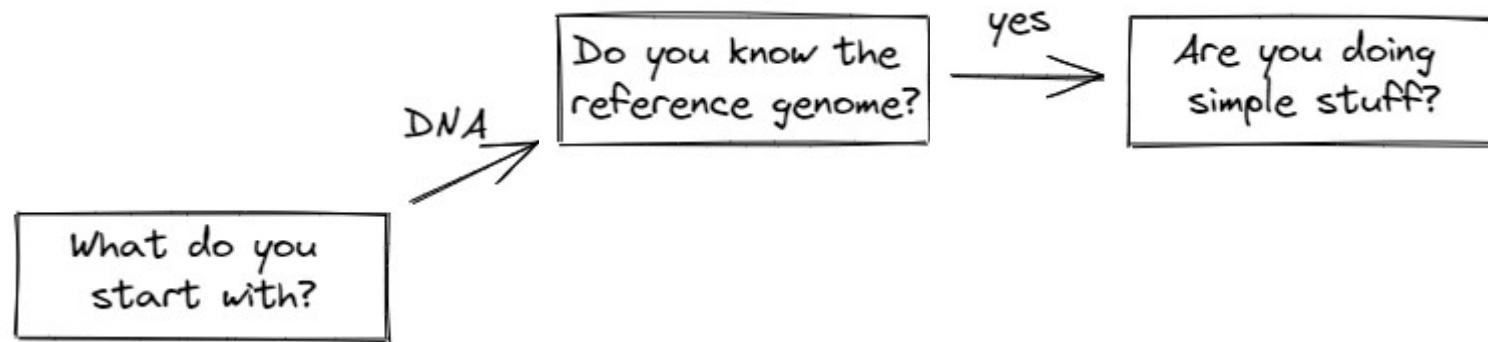
Long read sequencing

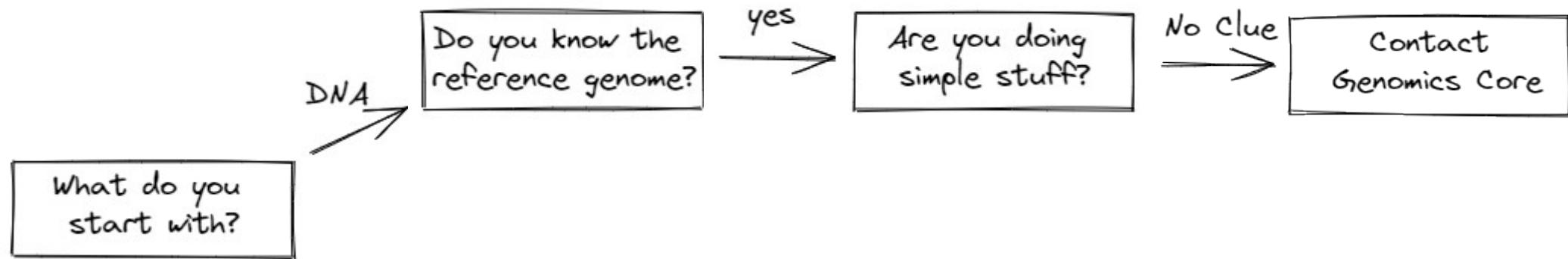


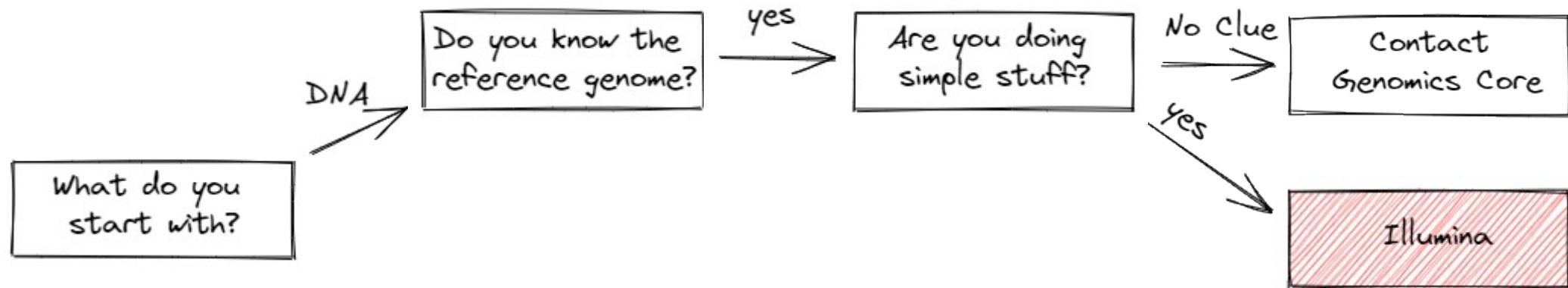
Experimental design

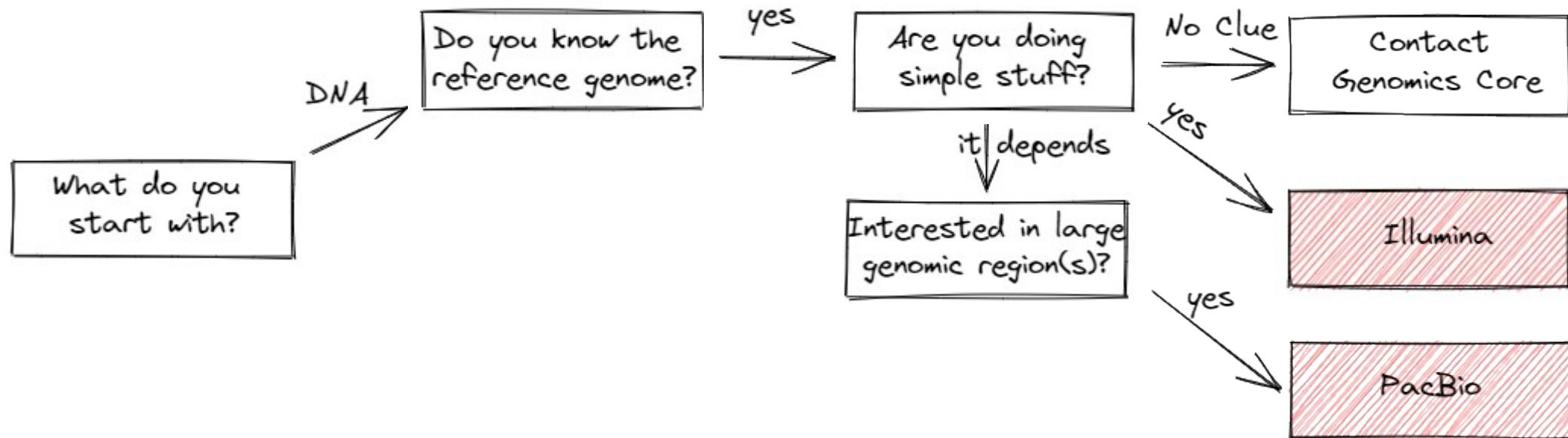


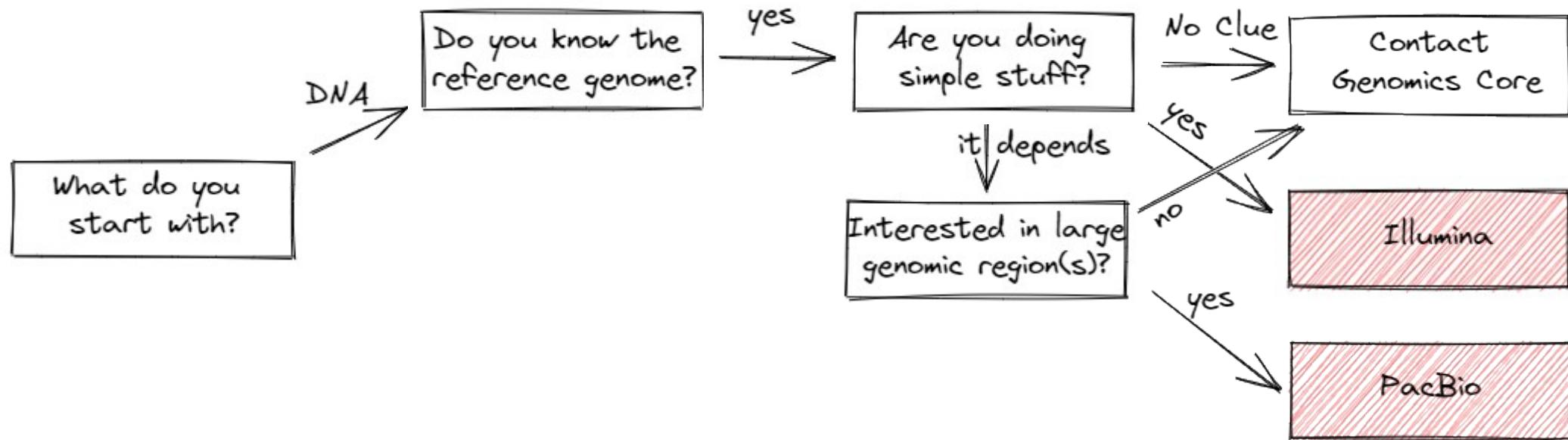


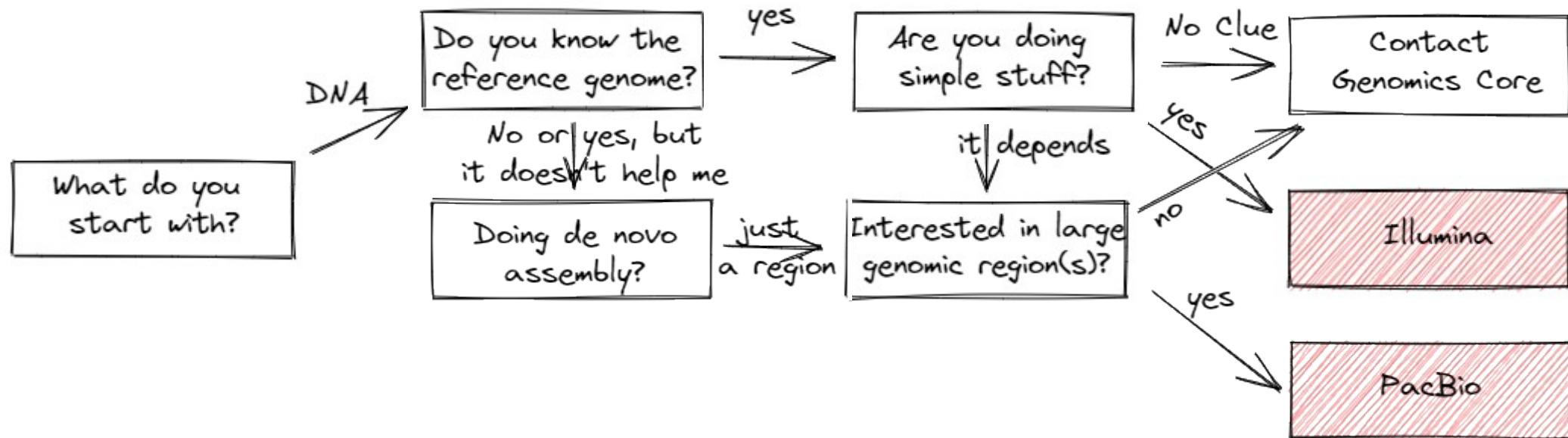


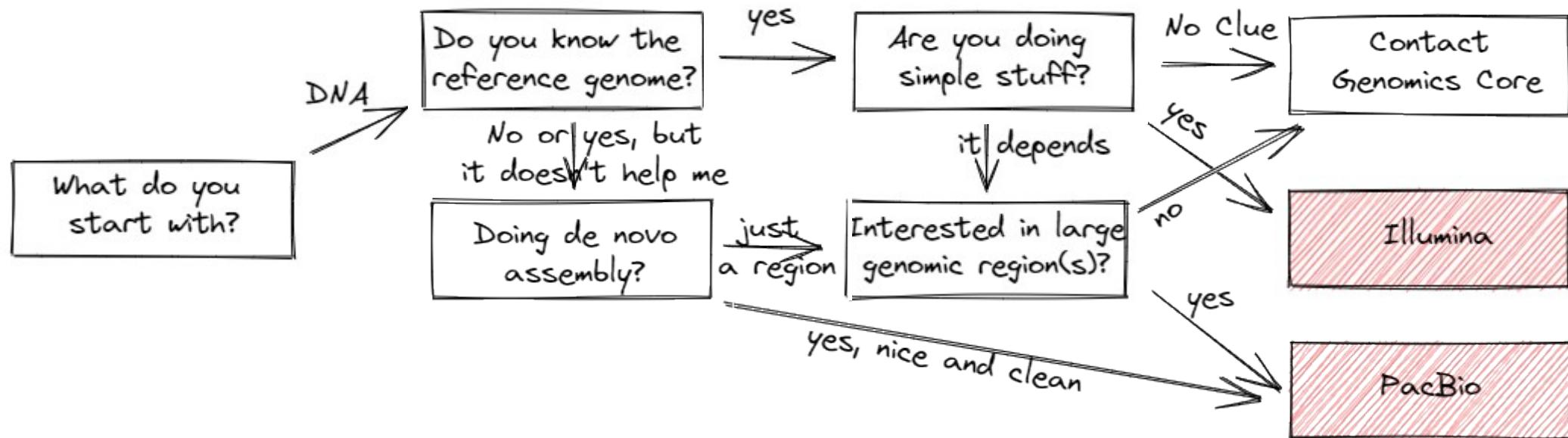


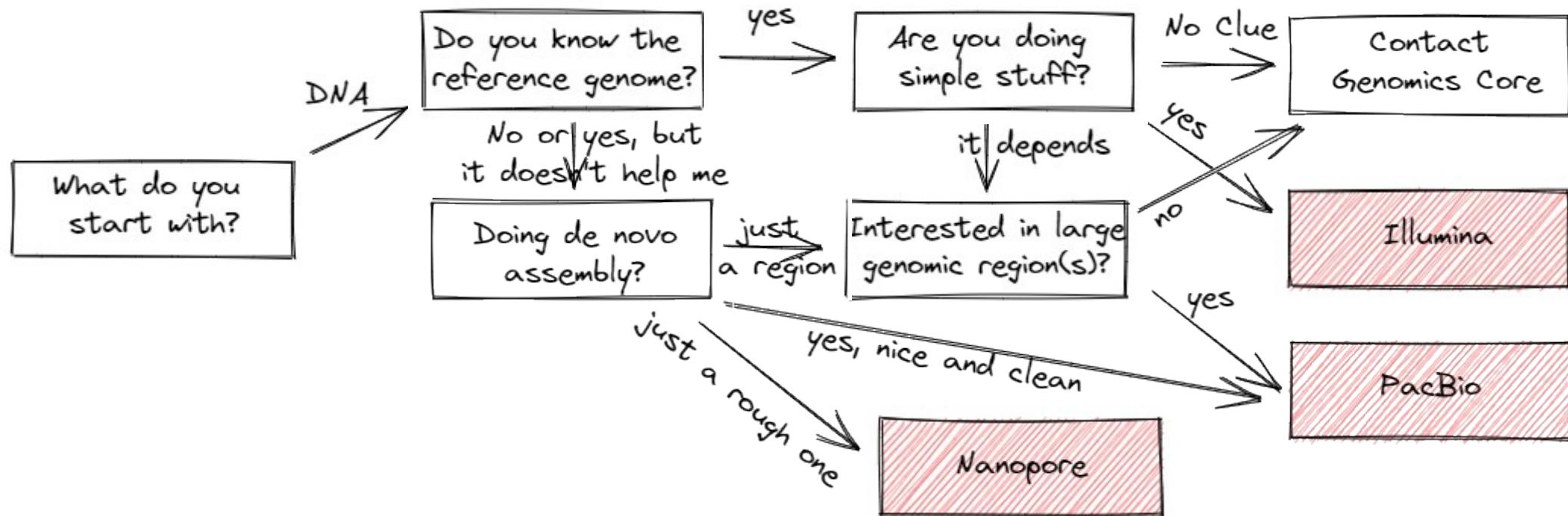


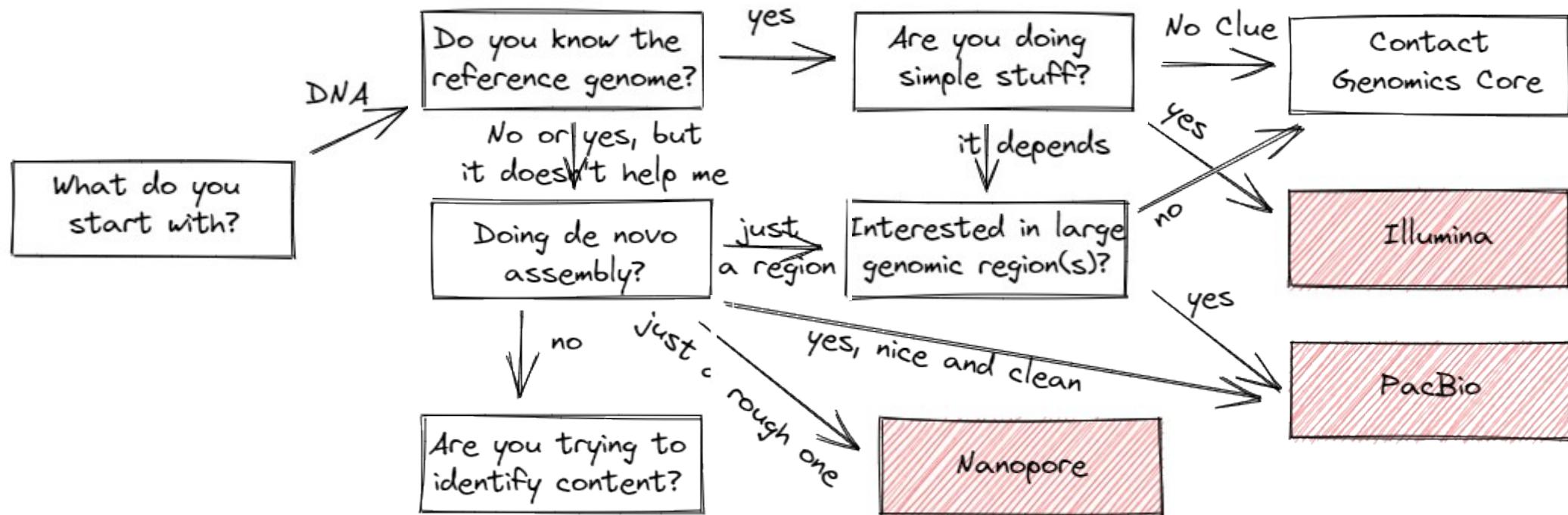


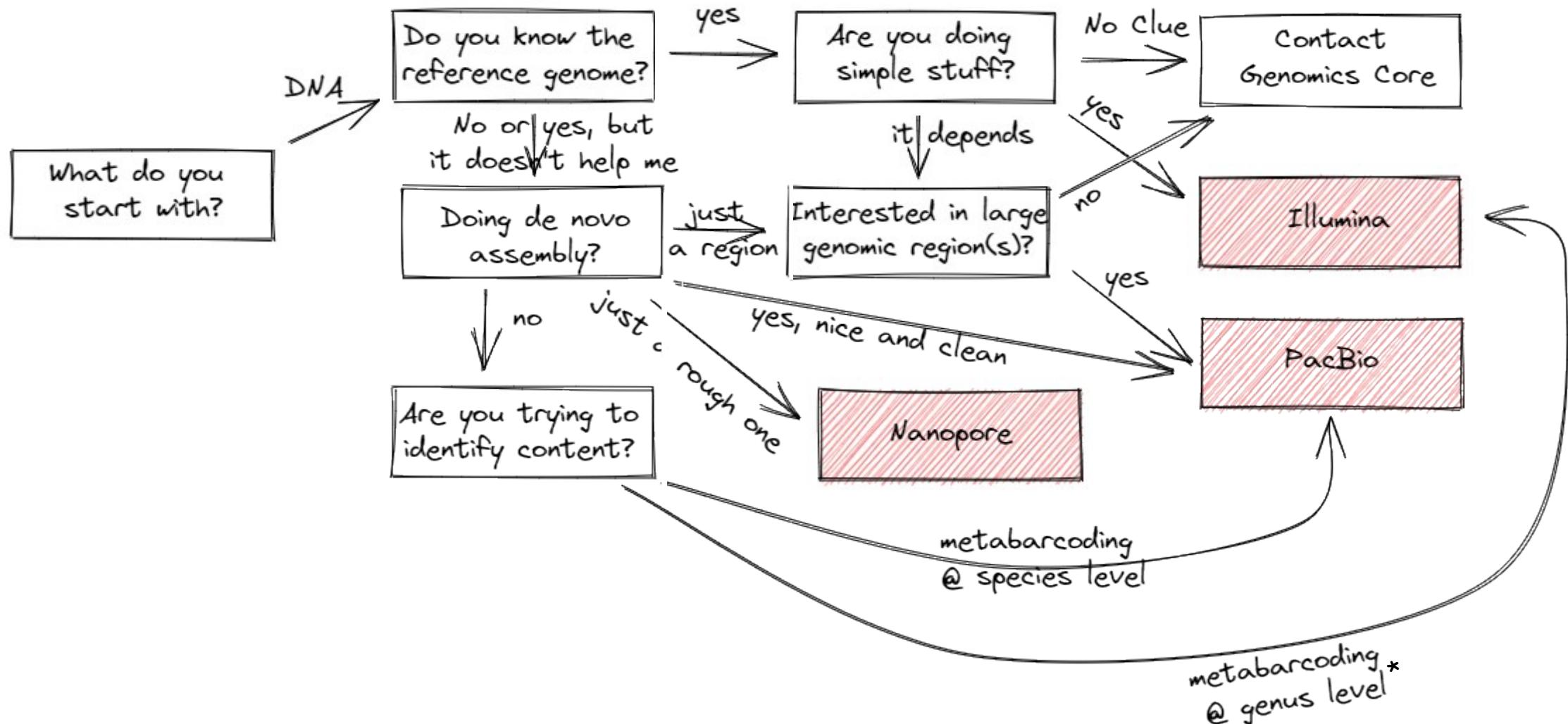


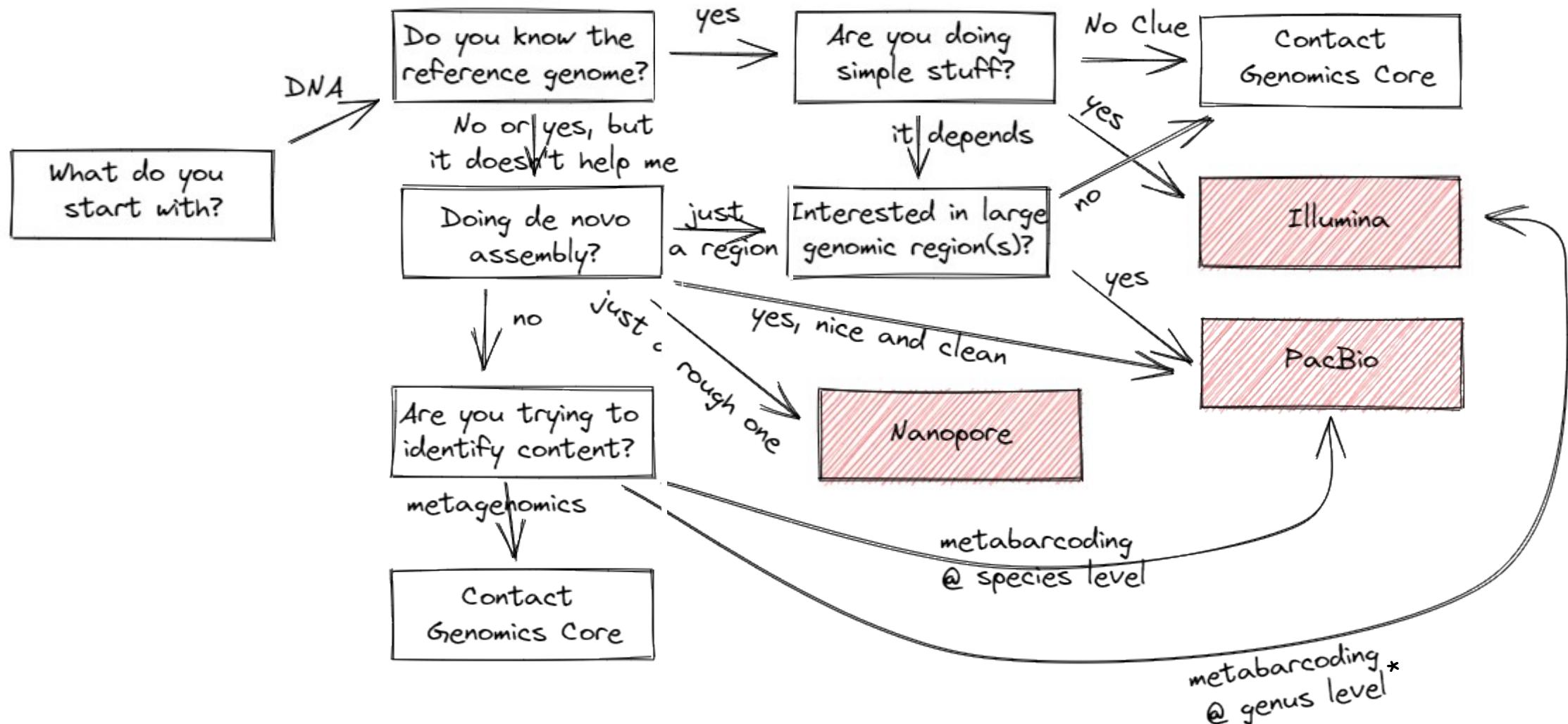






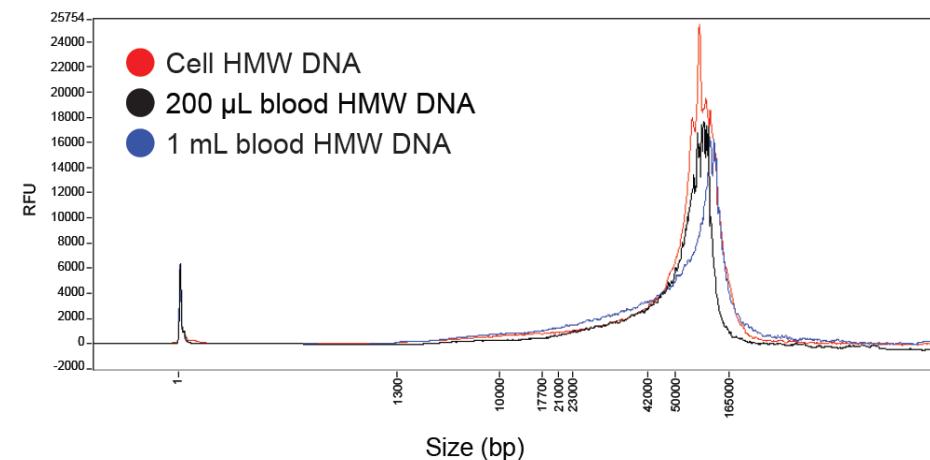






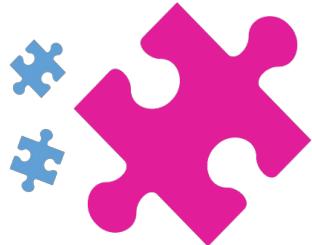
DNA extraction

- If it's broken, we can't fix it
- Nanobind is frequently the best option
- Protocol for
 - blood
 - Animal tissue
 - Plant tissue
 - Bacteria
 - Cell culture

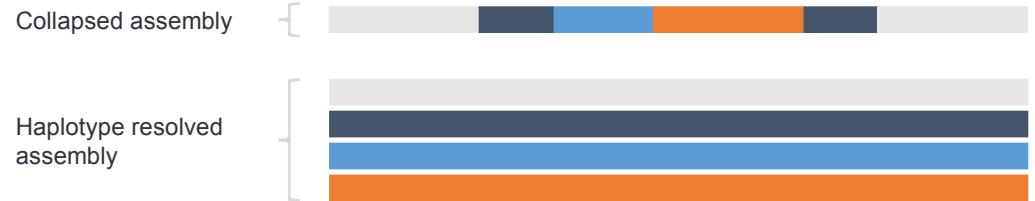


DNA analysis methods

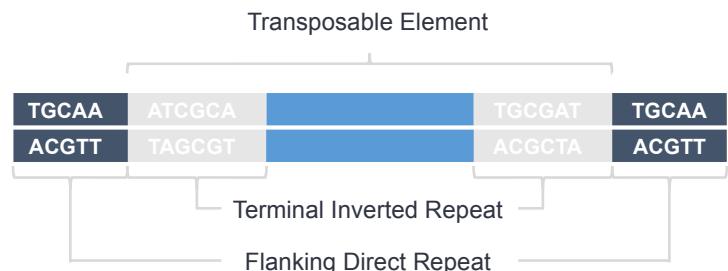
Why long reads?



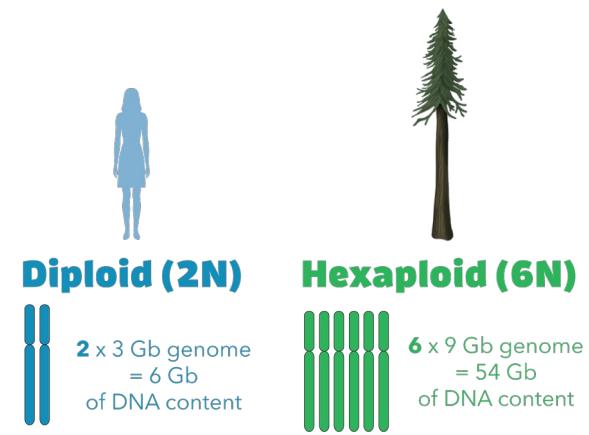
Bigger is better – when it comes to easy assembly



Accuracy helps phase haplotypes in outbred diploids and polyploids



Long reads span short repeats and accuracy allows differentiation between copies of long repeats



HiFi reads are better for higher ploidy genomes

High-quality genome assembly – the four Cs

Assemblies using long reads check all the boxes of a high-quality genome assembly



CONTIGUITY



High-contig N50



COMPLETENESS



No missing bases or
fragmented genes



CORRECTNESS

AGTCCGT**T**CAATGT
GCAATAGACAGTC
TACAGTTGGACAT
GCAGATACAGATA

High base accuracy
+ phased alleles

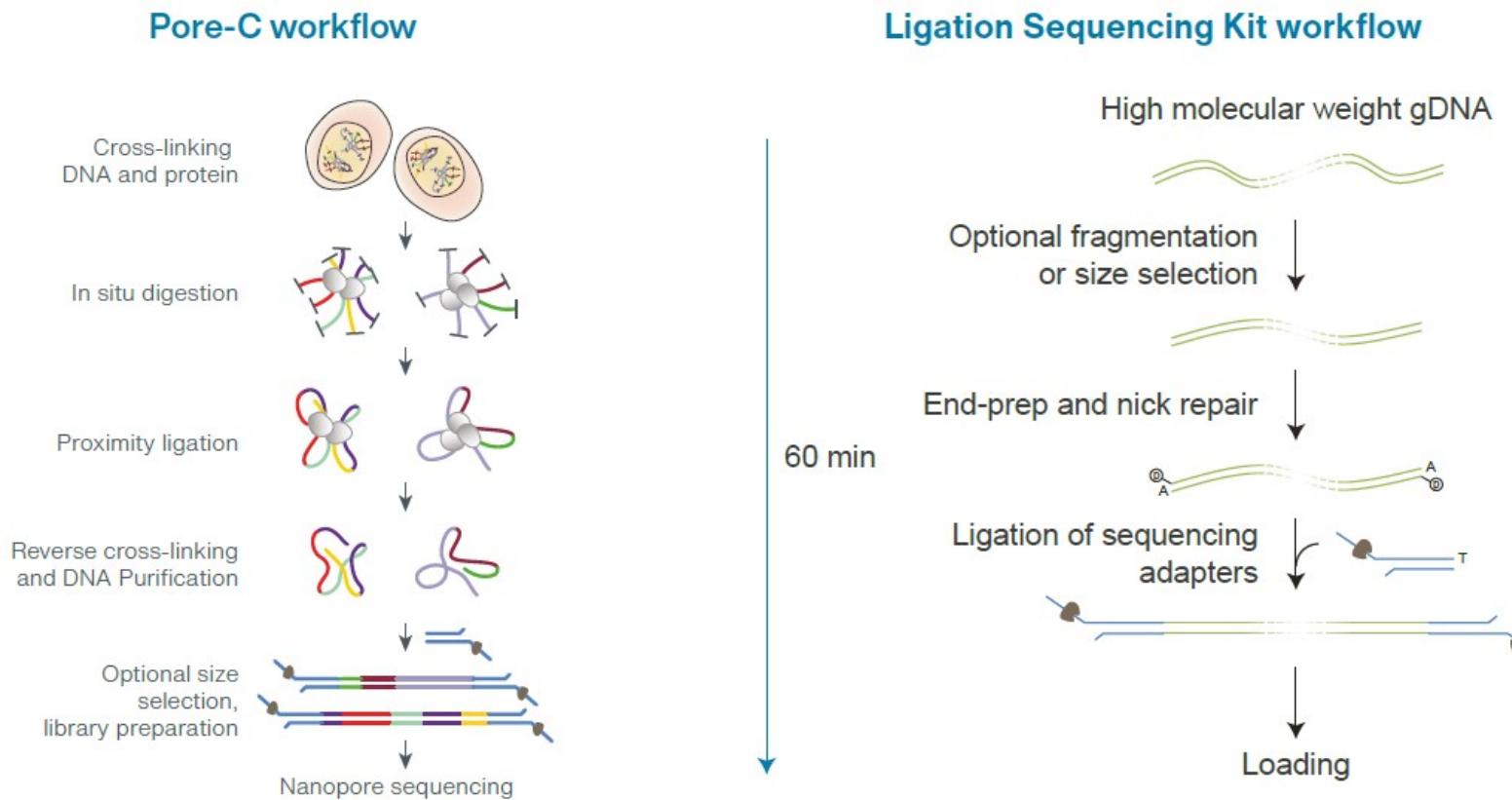


COMPUTE



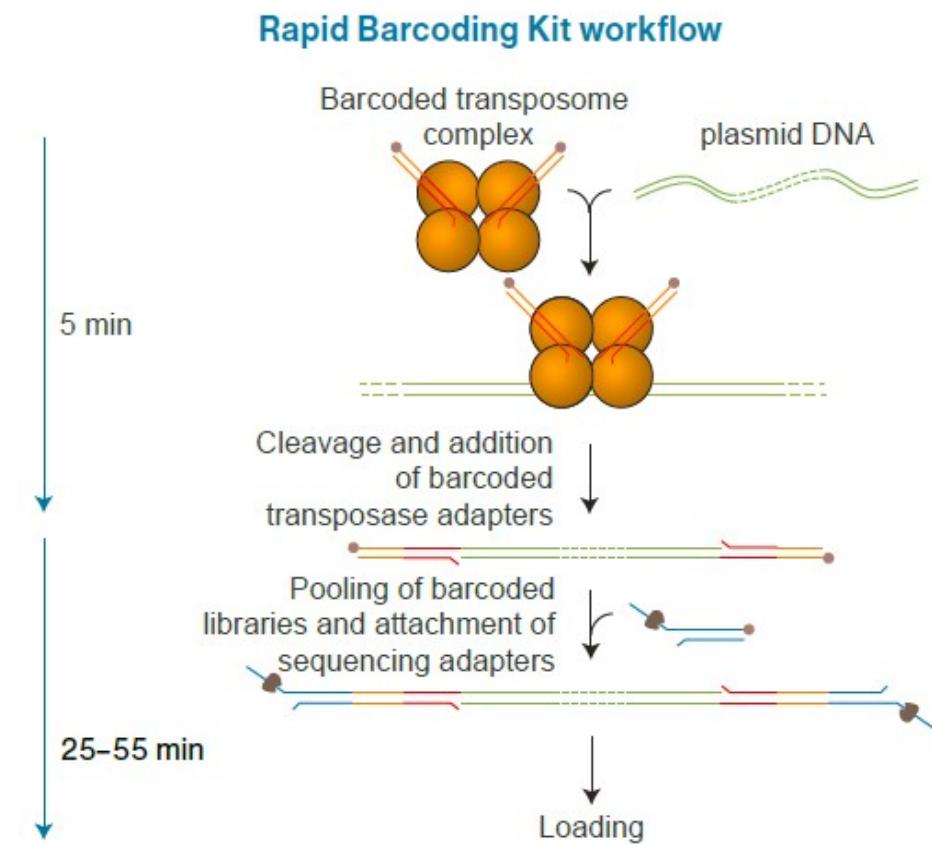
Small file sizes +
fast analysis time

Scaffolding

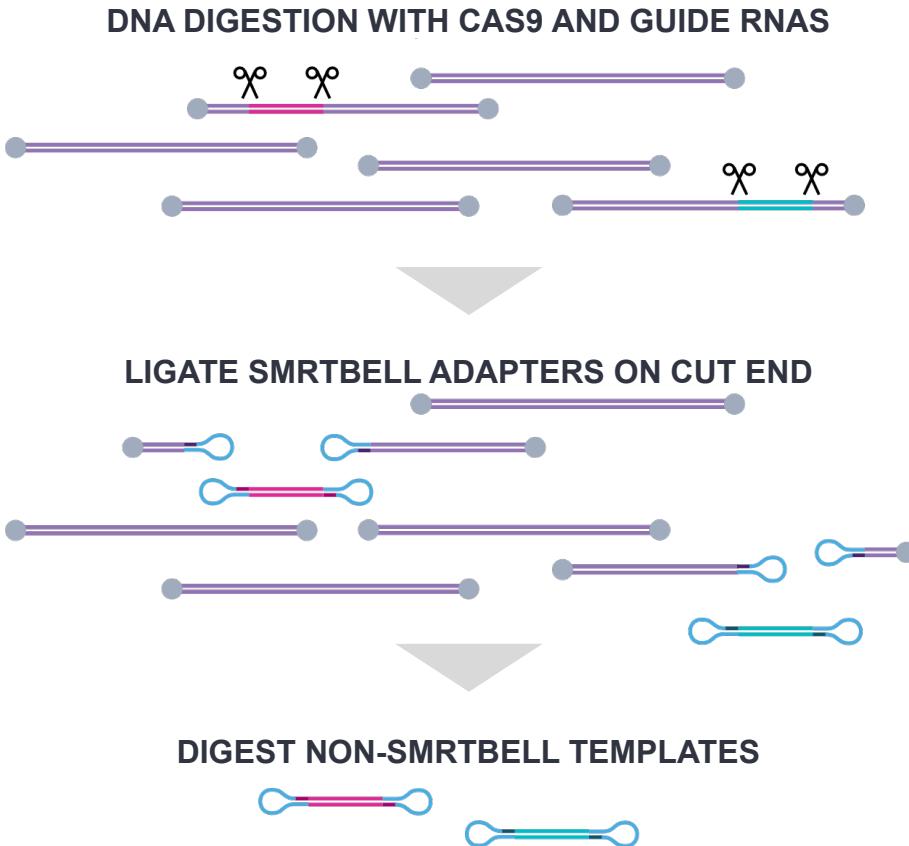


Nanopore full plasmid sequencing

- Multiplex up to 96 plasmids
- Full plasmid sequence: no assumptions
- EPI2Me analysis pipeline

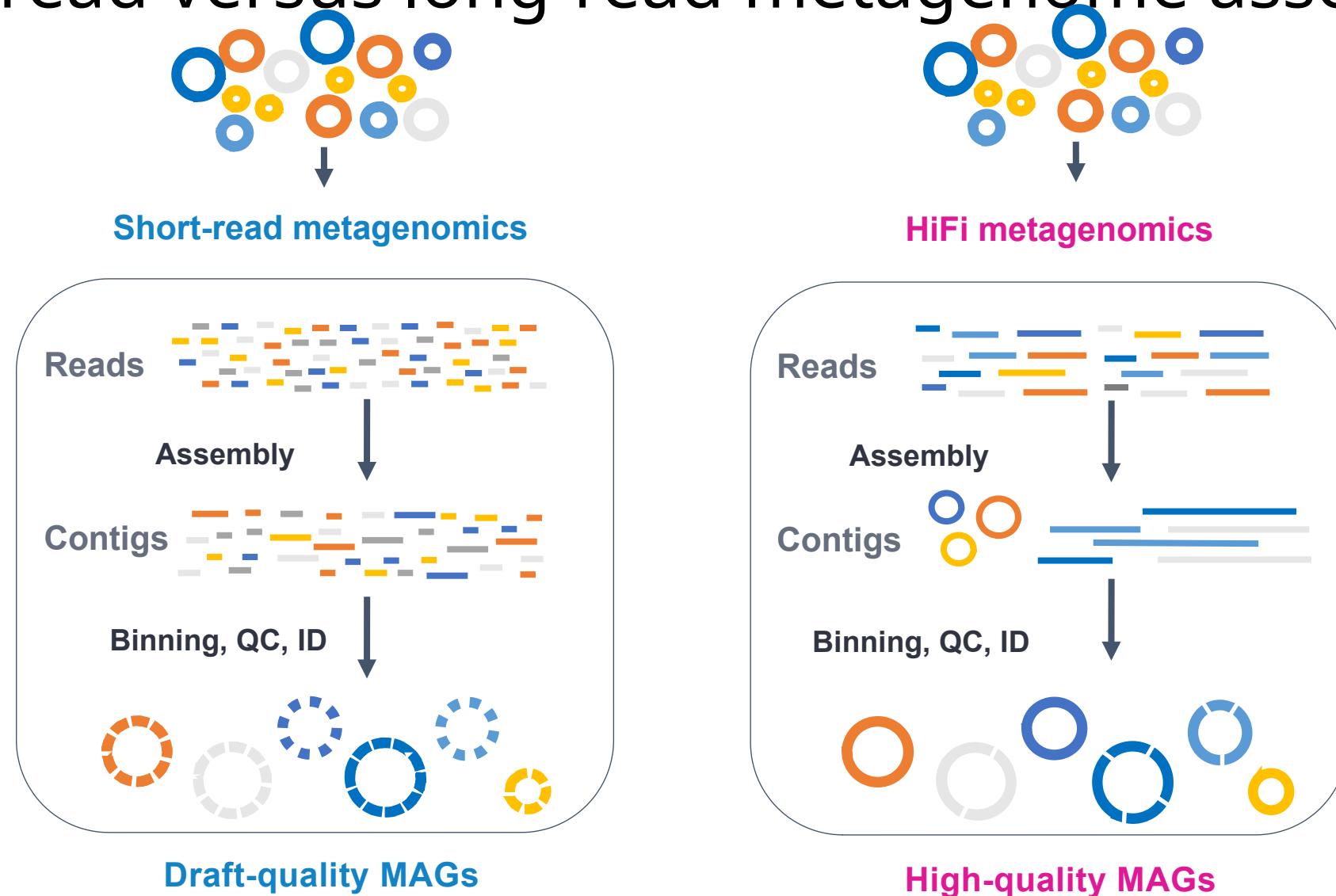


PureTarget enrichment with CRISPR-Cas9 is more accurate than PCR or hybrid capture

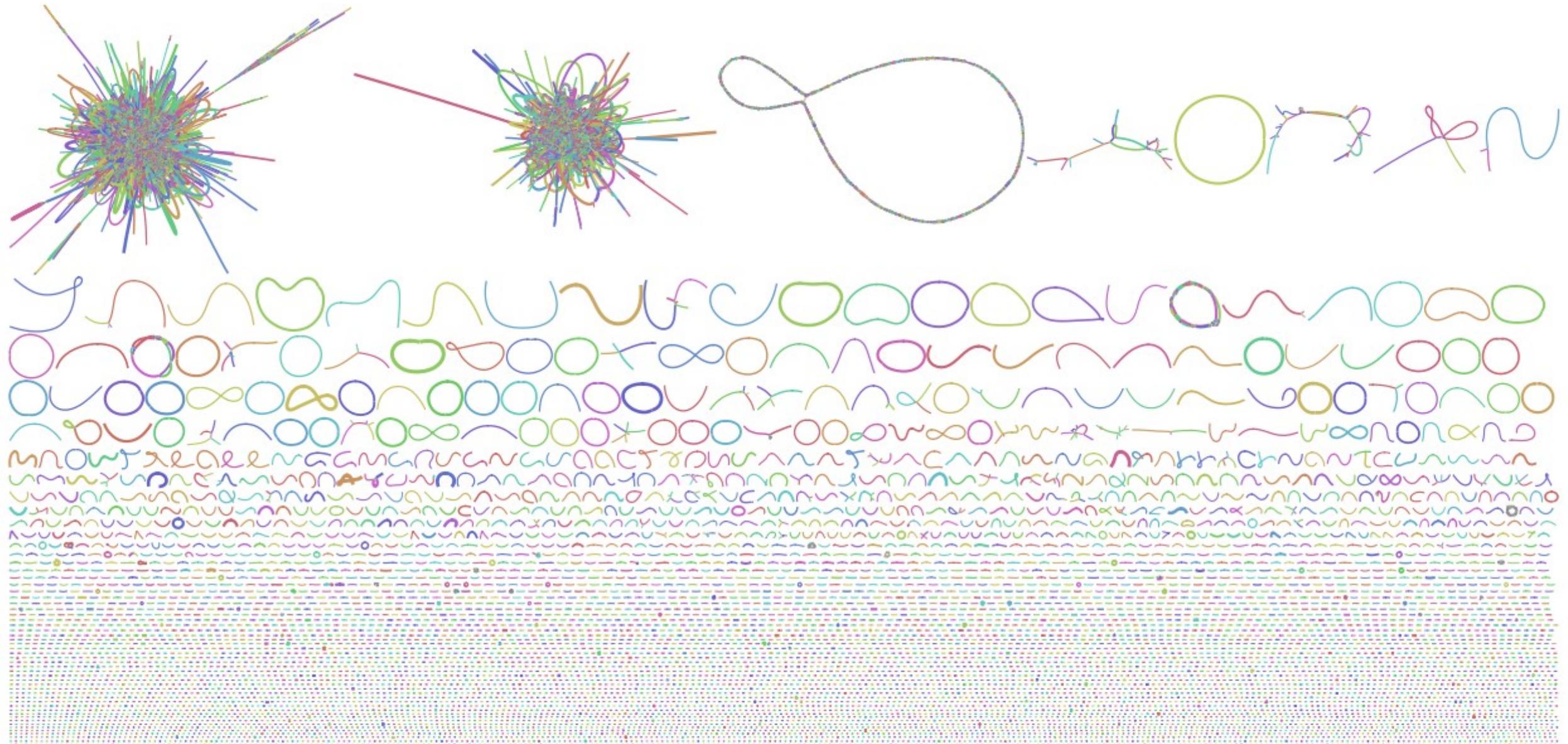


	PureTarget (no-amp)	PCR amplicons	Long-read hybrid capture (Twist)
Uses PCR	No	Yes	Yes
Retain methylation	Yes	No	No
GC dropout	No	Yes	Yes
Size bias	Low	High	High
Replication errors	No	Yes	Yes

Short-read versus long-read metagenome assembly

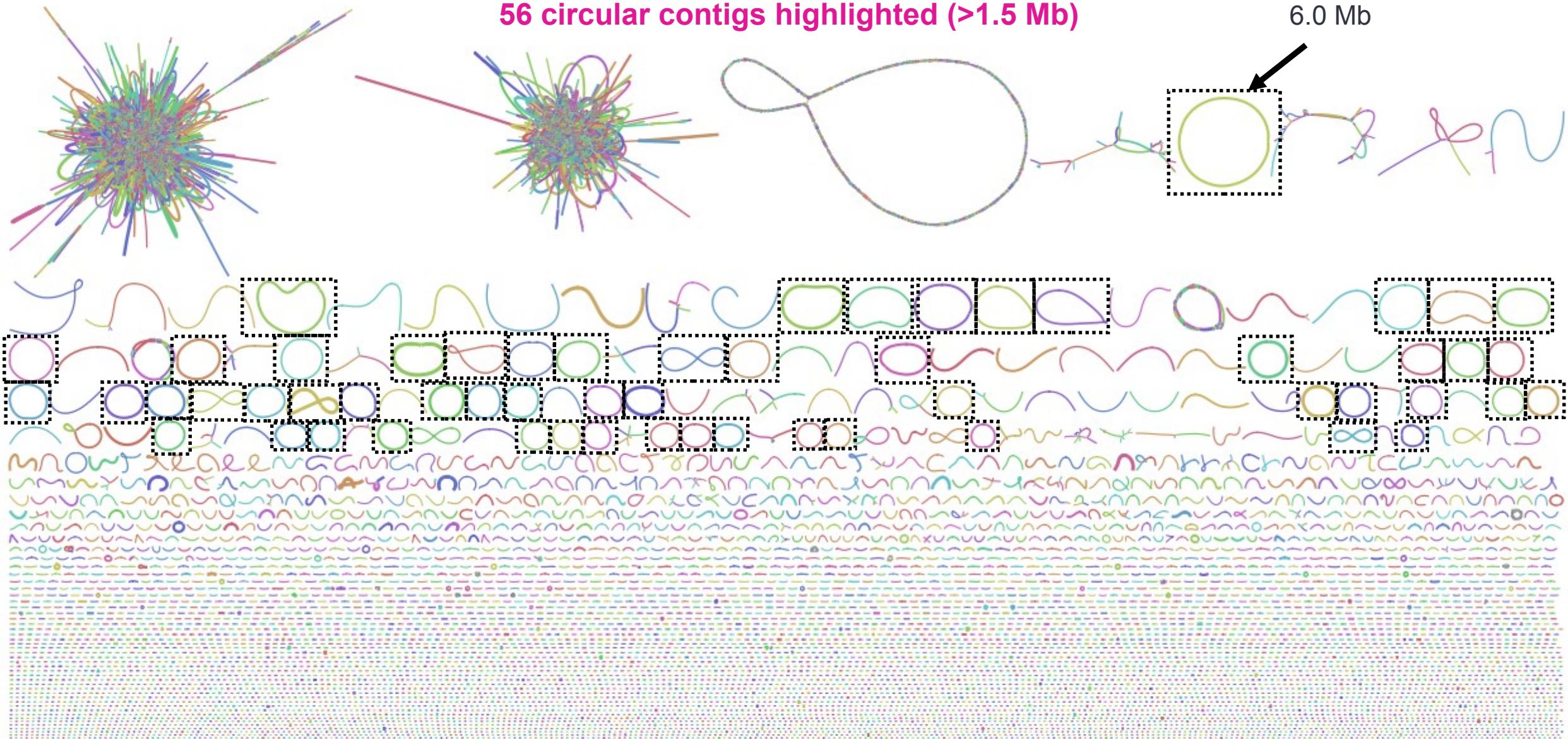


hifiasm-meta assembly graph



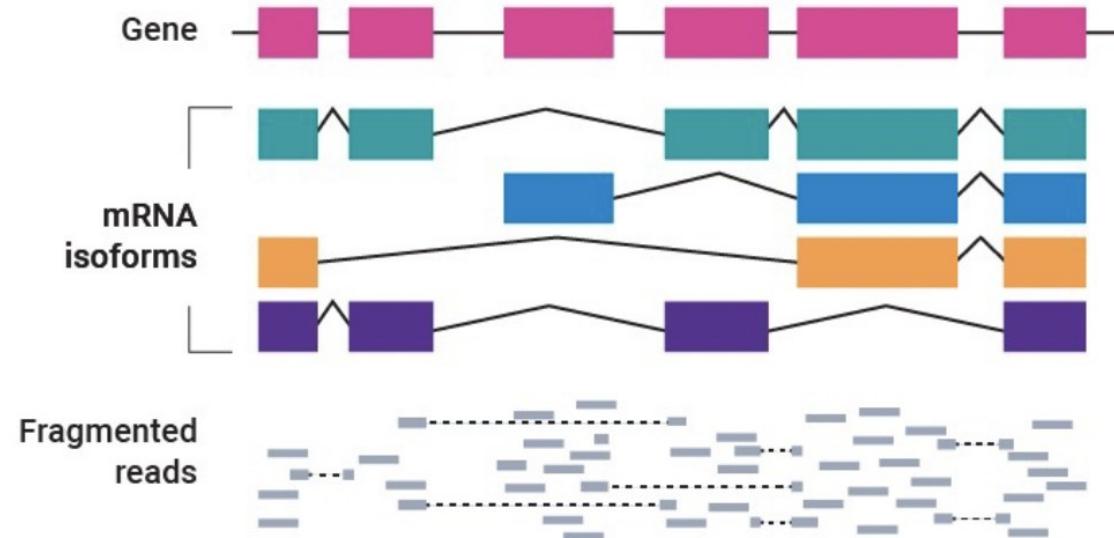
hifiasm-meta assembly graph

56 circular contigs highlighted (>1.5 Mb)



Iso-Seq method delivers full-length transcripts

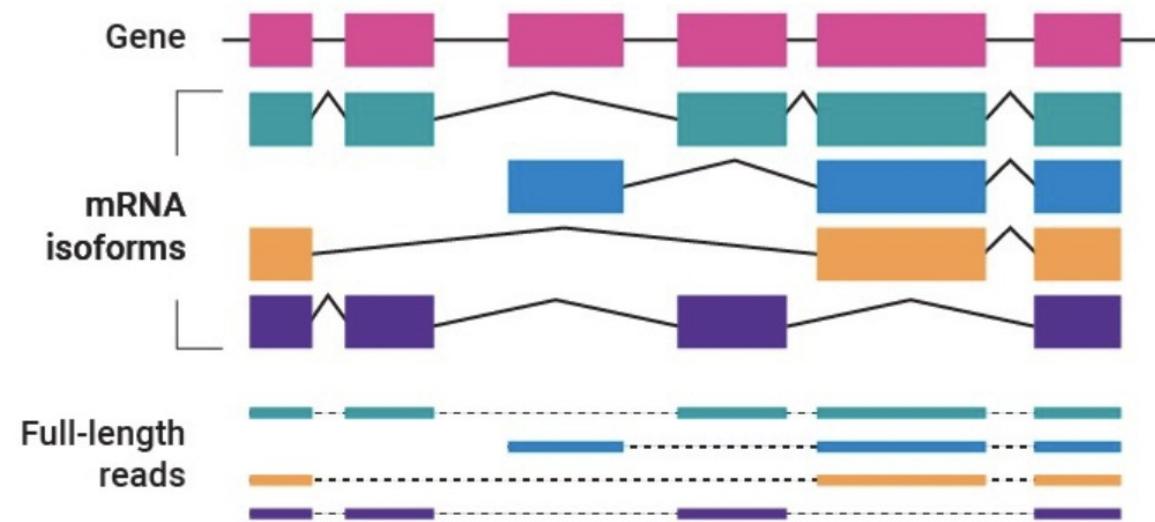
Short read sequencing



Short-read sequencing can only assemble ~20 to 40% of human transcriptomes

PARTIAL view of isoforms

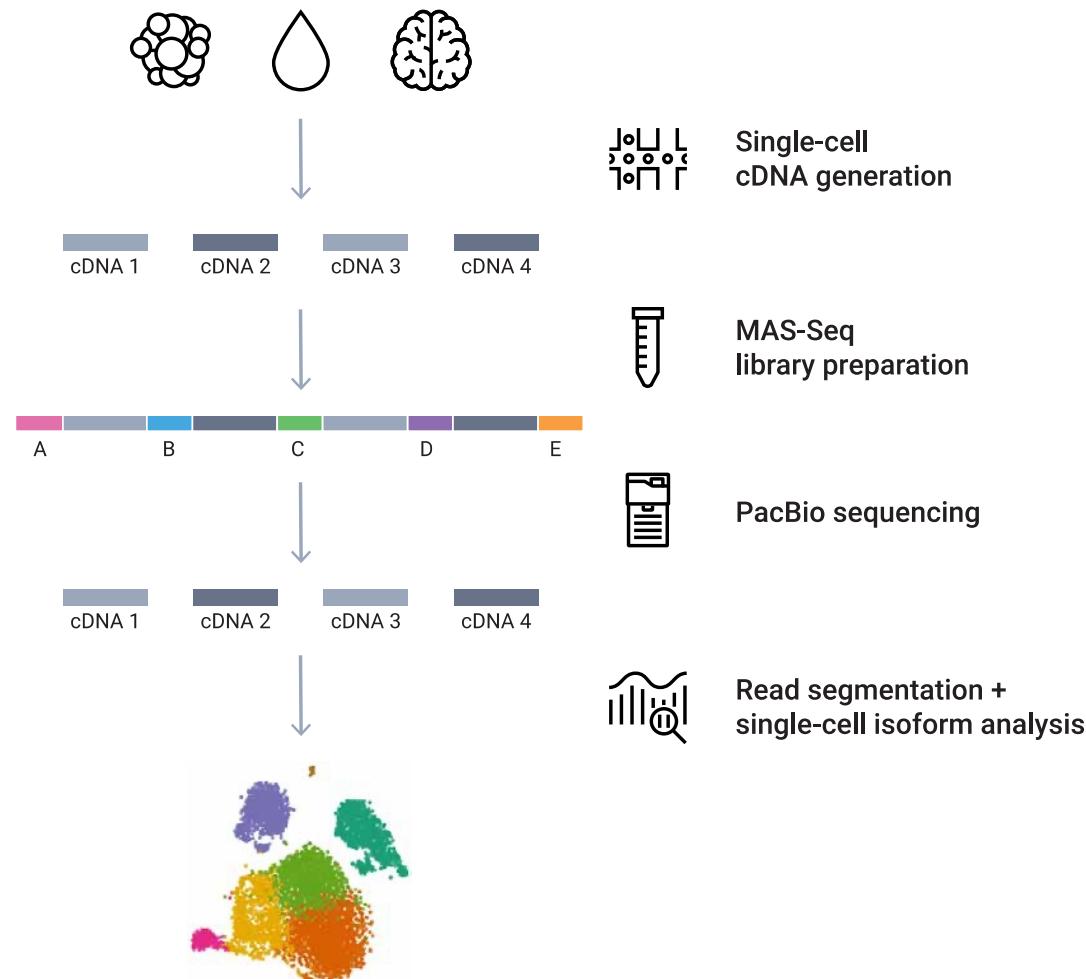
Long read sequencing



PacBio's long-read sequencing offers superior **isoform discovery power**

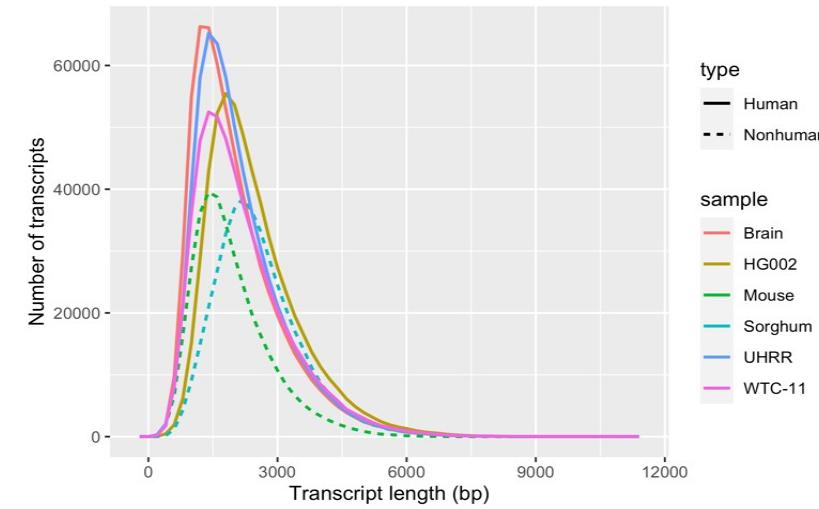
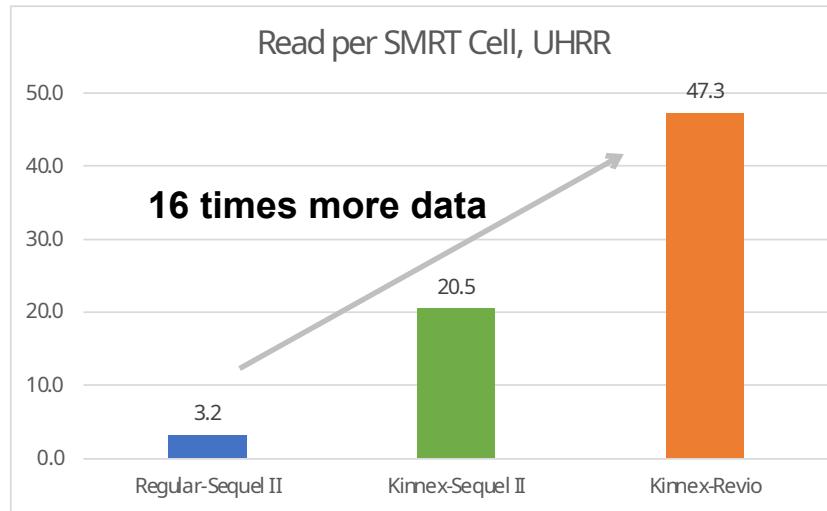
COMPLETE view of isoforms

Increasing throughput with concatenation: Kinnex kits

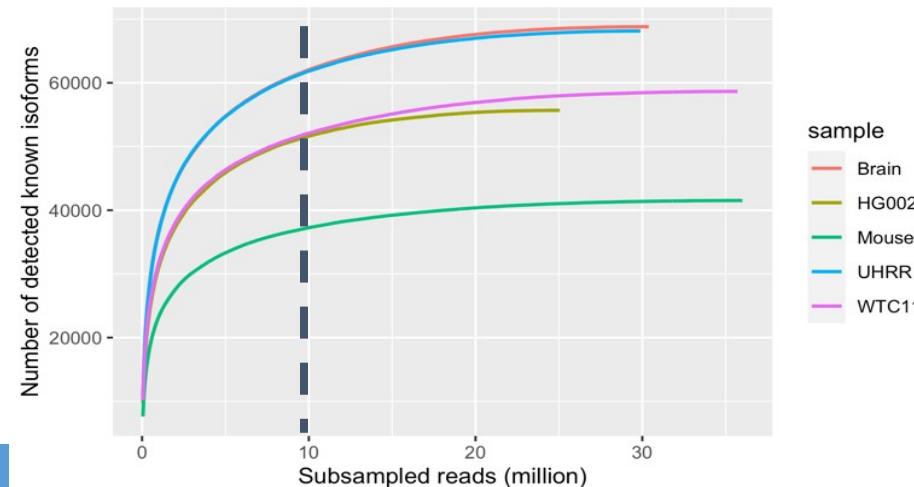


Kinnex full-length RNA kit: exceptional quality with more reads in less time

Get more reads in less time



Capture long transcripts in all variety of species and tissues



At 10M reads, 80% of known isoforms are detected

3 varieties of Kinnex kits

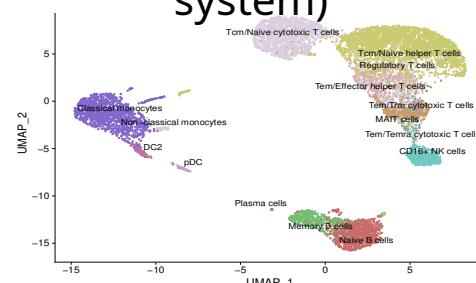


Kinnex single-cell RNA kit

Upgrade to *MAS-Seq* for
10x Single Cell 3' kit

Support 10x 3' and 5'; up to 4-plex

40M reads (Sequel II and IIe systems), 80M reads (Revio system)

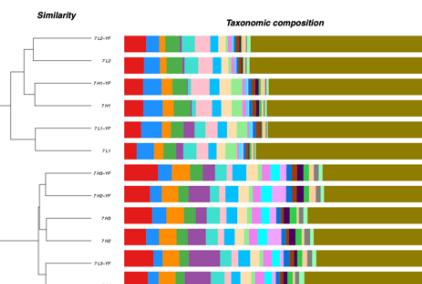


Kinnex 16S rRNA kit

Full-length 16S rRNA
for species identification

Up to 1,536-plex

25M reads (Sequel II and IIe systems), 60M (Revio system)



Kinnex full-length RNA kit

Full-length
RNA sequencing

Up to 48-plex

15M reads (Sequel II and IIe systems), 40M reads (Revio system)

