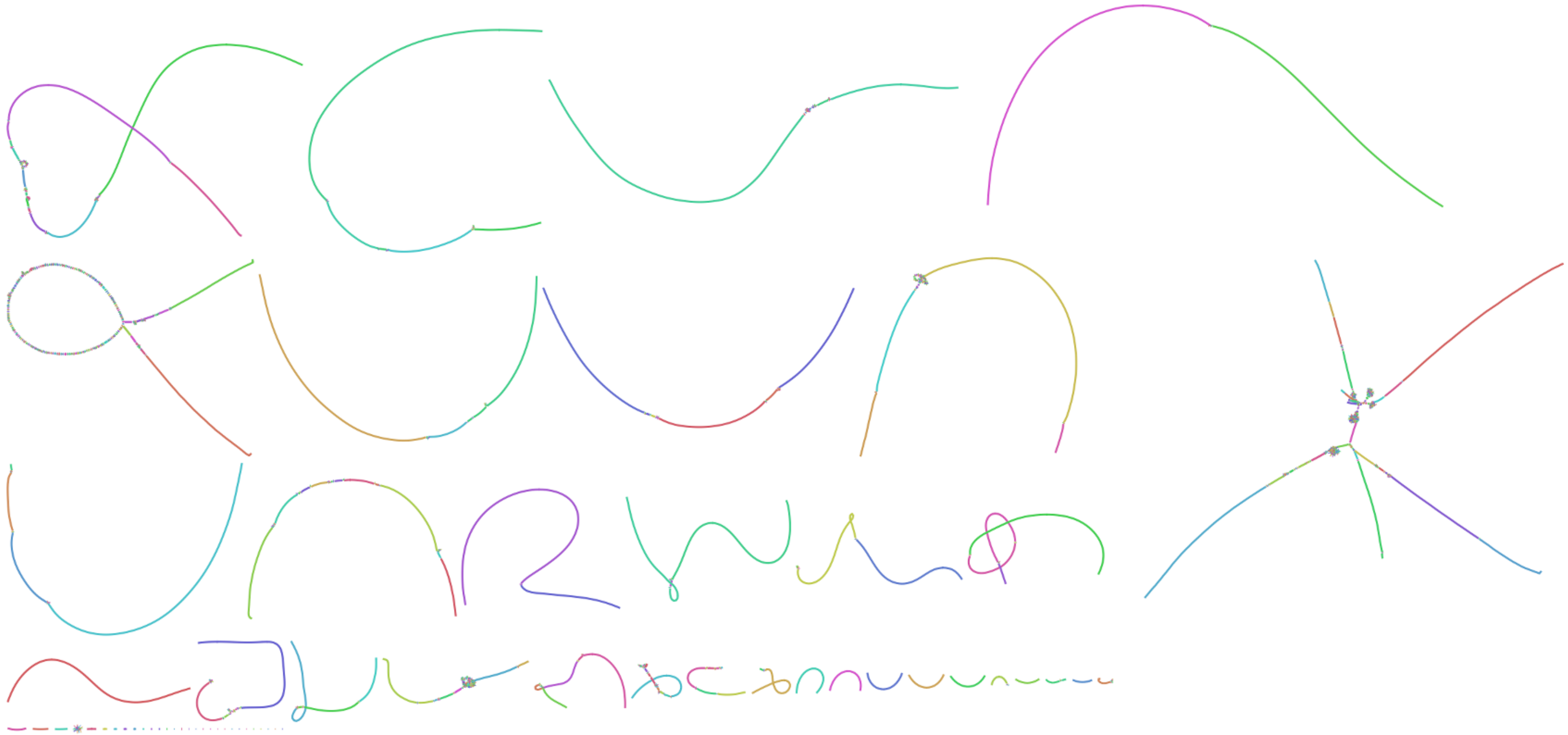
THE DNA INNOVATION ENGINE DRIVING CROP IMPROVEMENTS

Dick Roelofs

START PRESENTATION

FROM TELOMERE TO TELOMERE

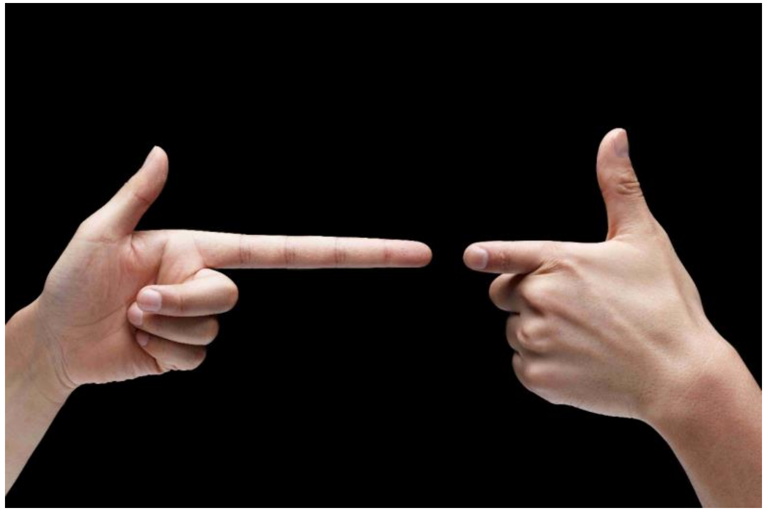


Source: Adam Phillippie https://genomeinformatics.github.io/CHM13v1/

FROM TELOMERE TO TELOMERE

THE LONG AND SHORT OF SEQUENCING: TECHNOLOGY

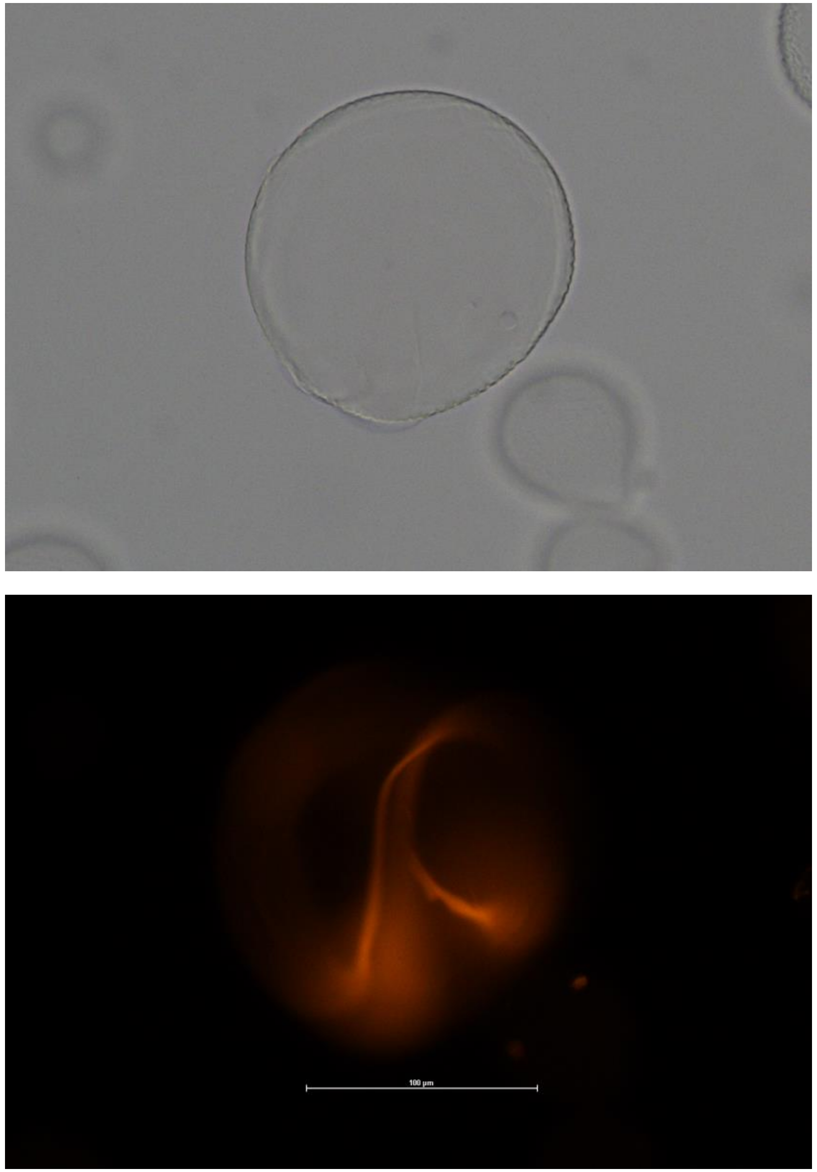
- From long to ultra long read sequencing
- Innovations in raw read accuracy
- PacBio® High Fidelity (HiFi)
- Oxford Nanopore Technologies (ONT)

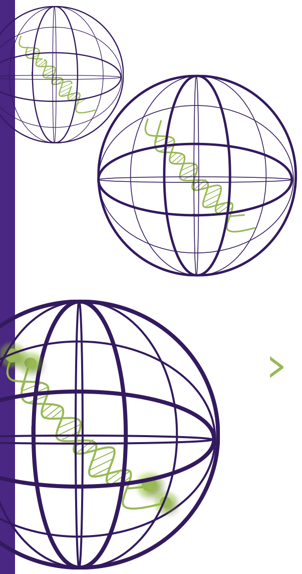


THE LONG AND SHORT OF SEQUENCING: DNA ISOLATION

- KeyGene has a longstanding track record in DNA isolation improvements
- Current High Molecular Weight (HMW) DNA isolation protocol
  - In solution
  - Nuclei as starting point
  - Without pipettes
  - Flexible
- Ultra HMW DNA isolation: encapsulation

N50 > 100 Kb !





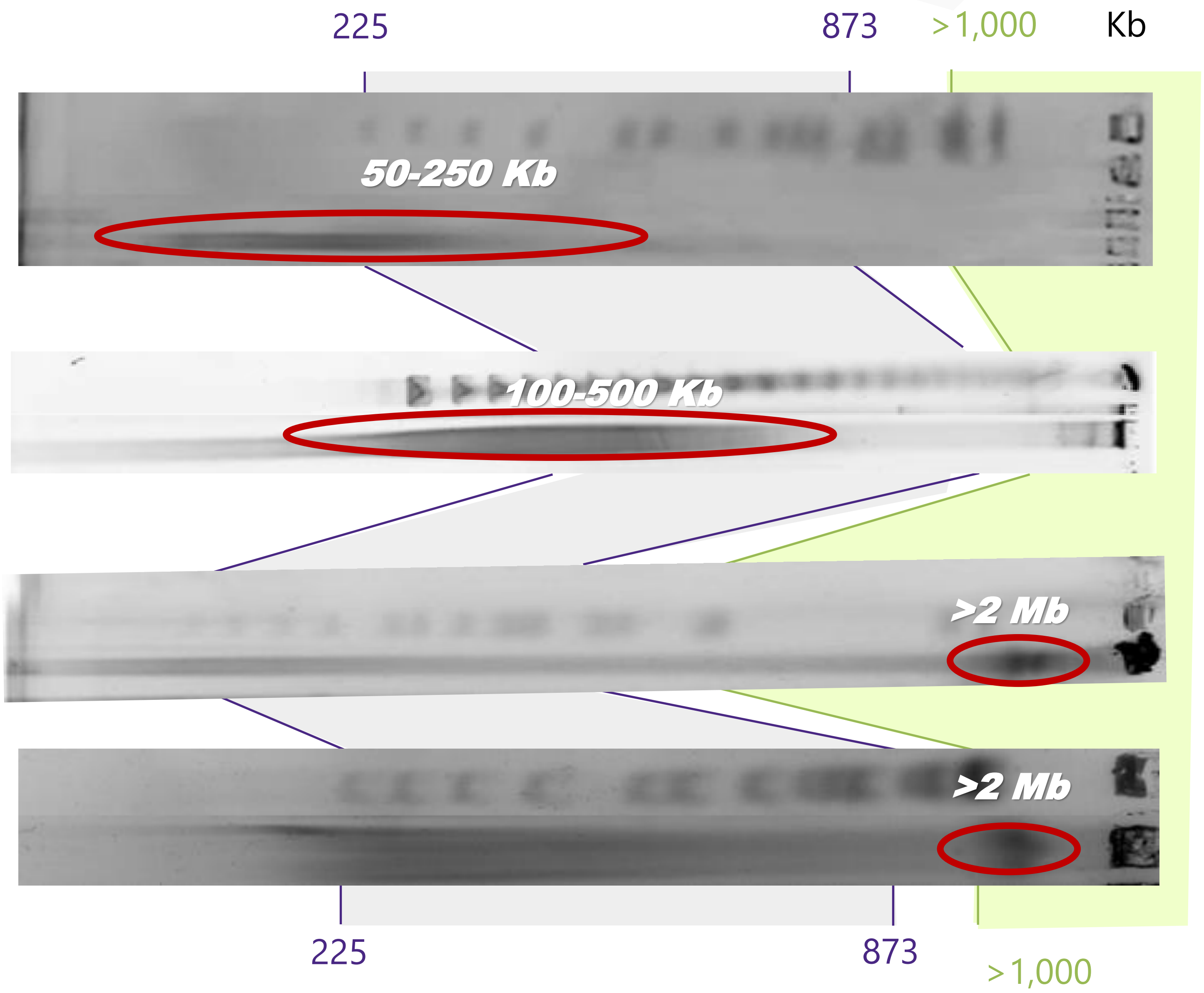
FROM HIGH TO ULTRA HIGH MOLECULAR WEIGHT

Solution-based gDNA isolation

circulomics

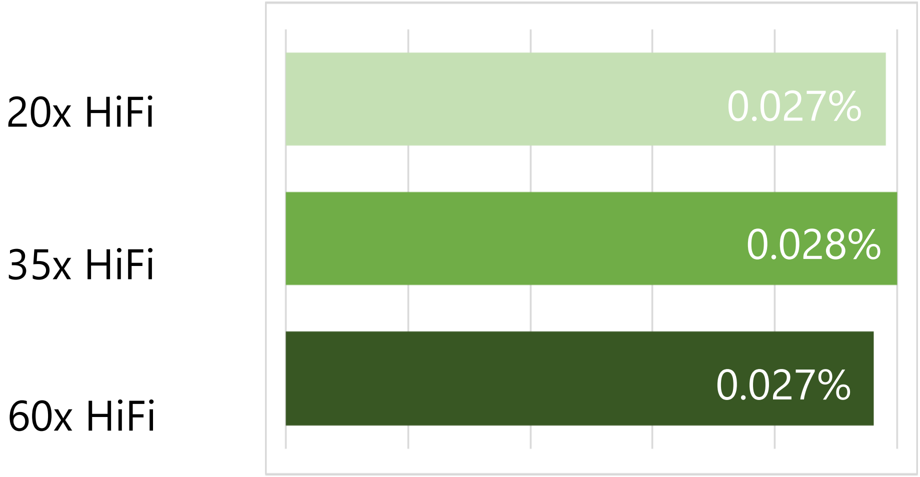
Agarose-embedded gDNA

KeyGene's 2021 uHMW innovation

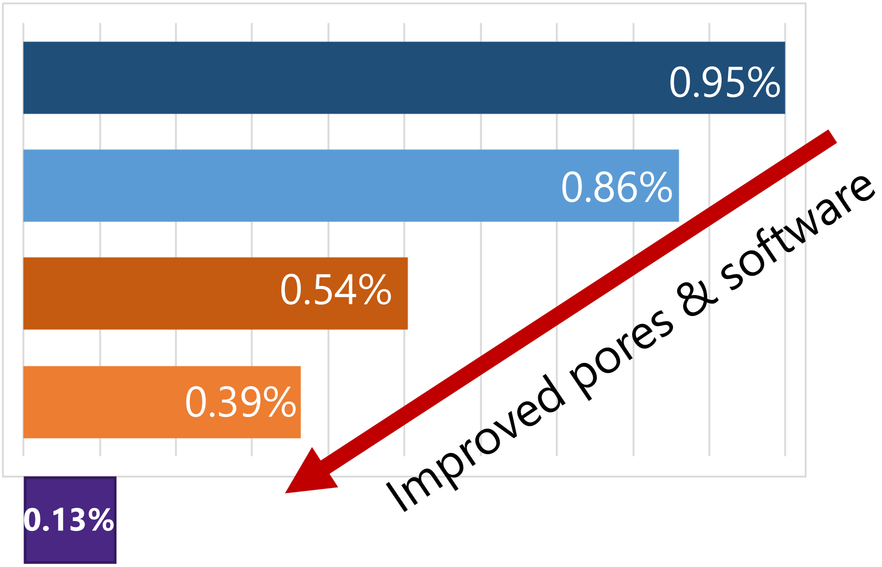


ACCURACY OF DE NOVO GENOME ASSEMBLY MELON

Percent <u>differences</u> to reference



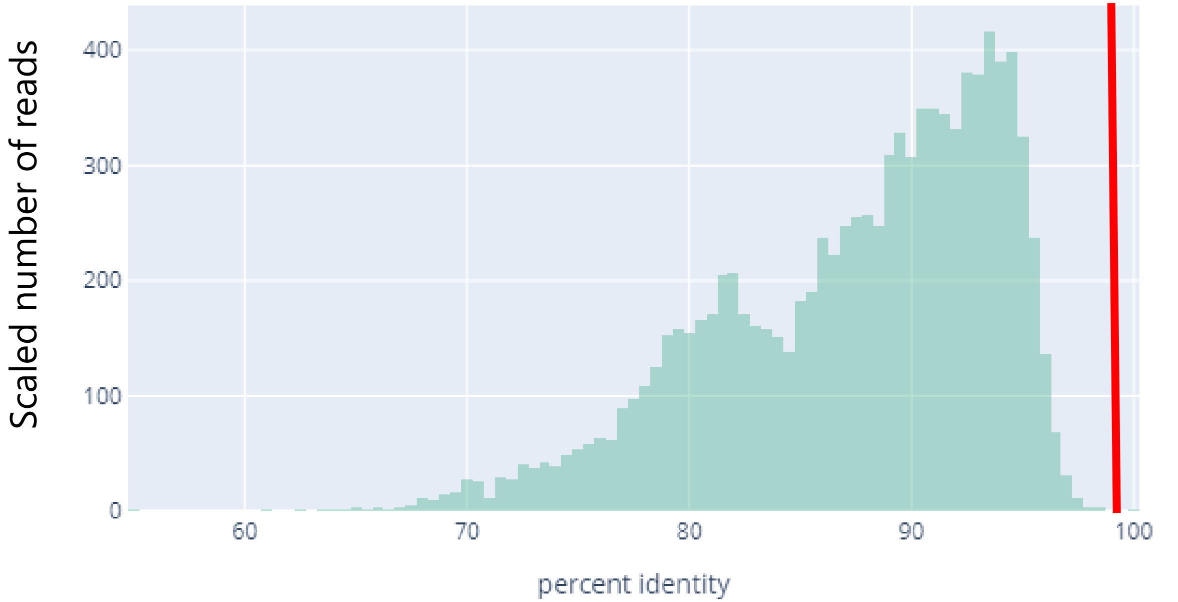
Percent <u>differences</u> to reference



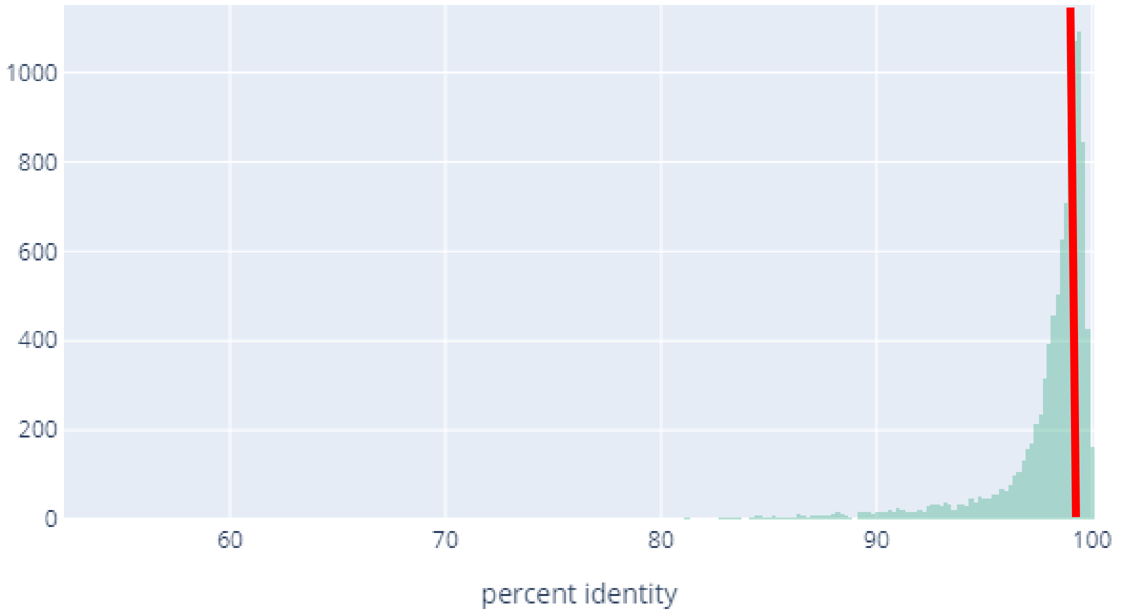
Combine data for optimal contiguity

ONT RAW READ ACCURACY IMPROVEMENTS

2020

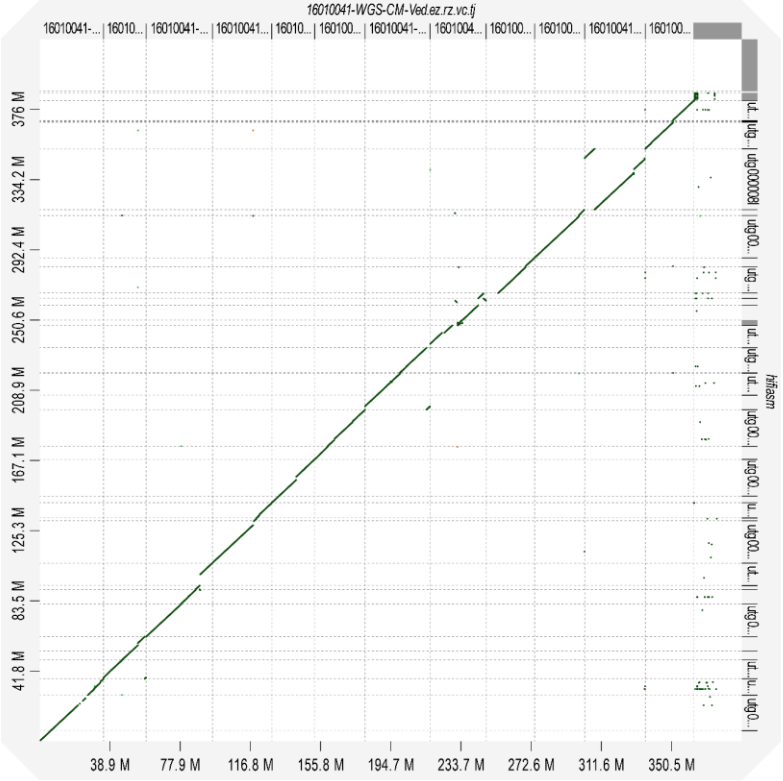


Last week's result

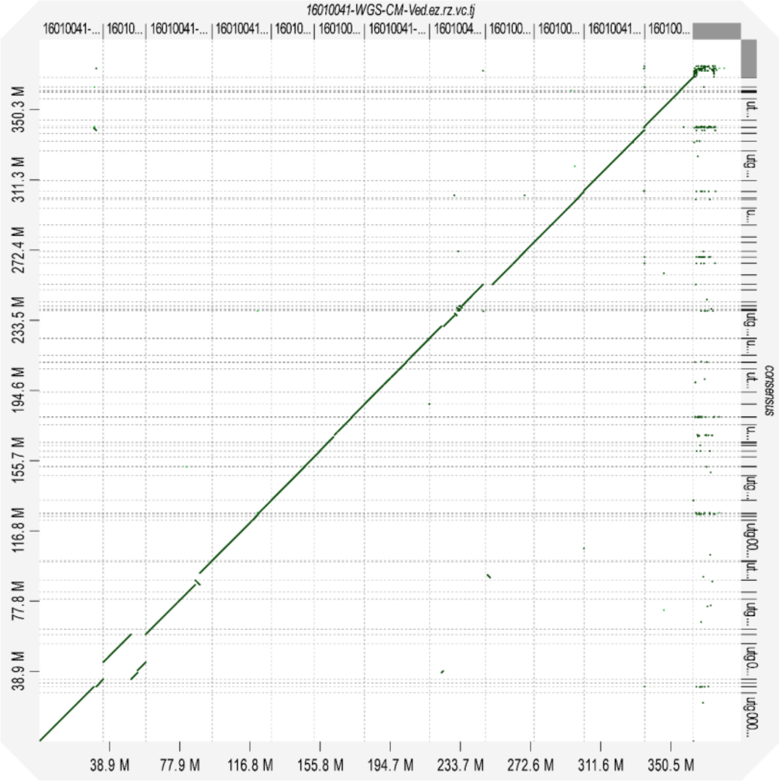


COMPARISONS DE NOVO GENOME ASSEMBLIES

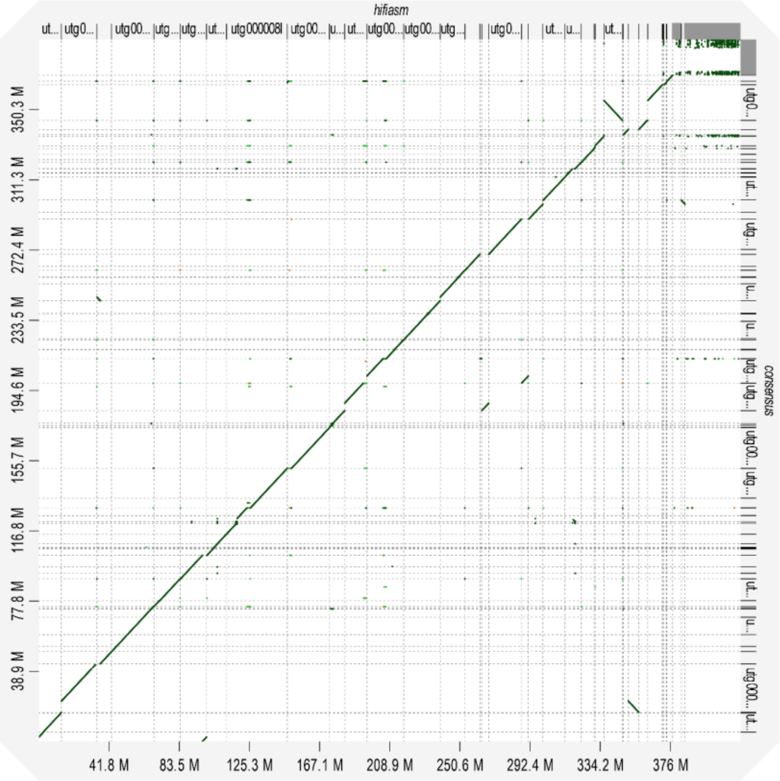
- Structural level



PacBio vs reference



ONT vs reference

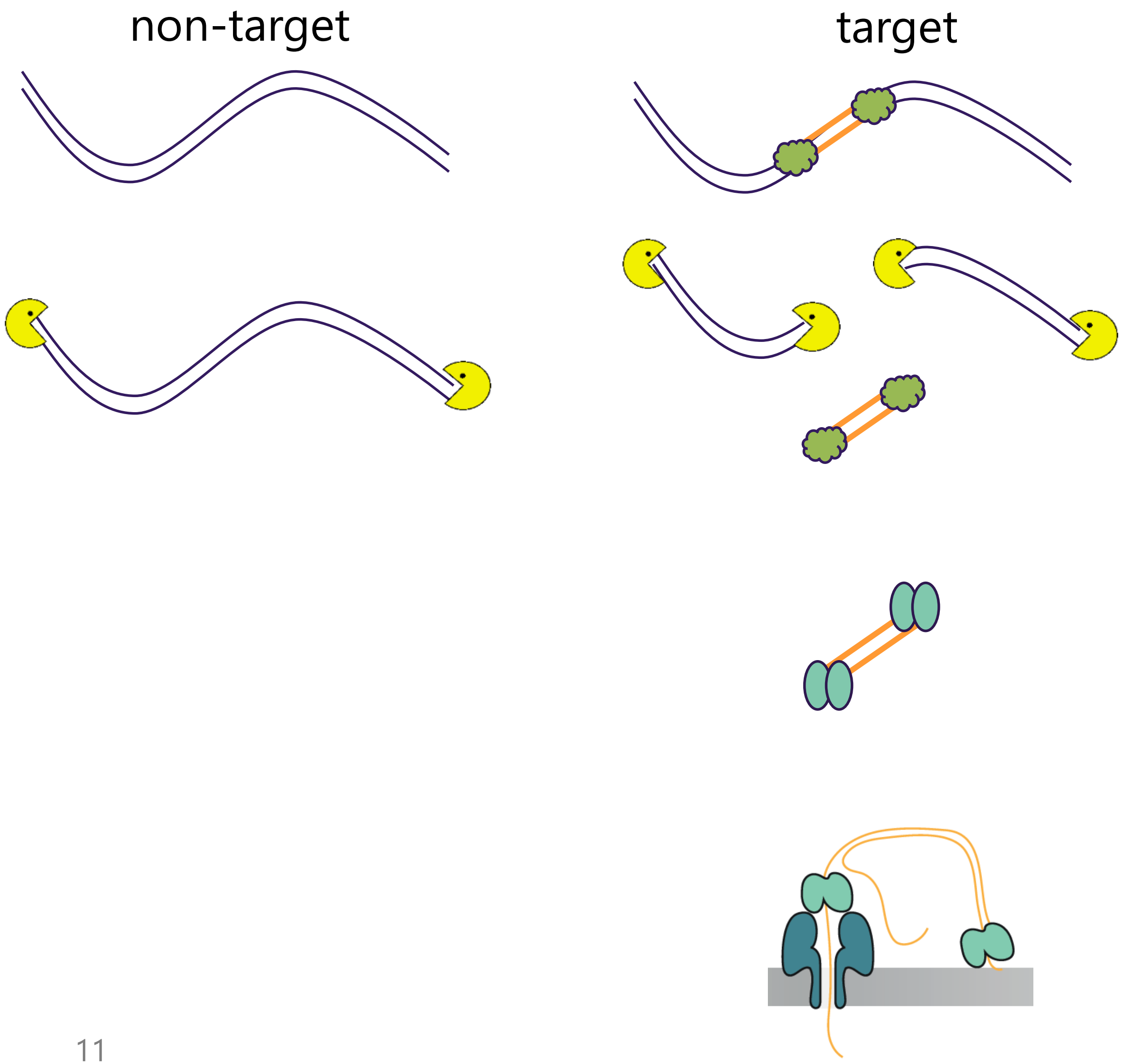


ONT vs PacBio

TARGETED SEQUENCING: A CUTTING EDGE TECHNOLOGY

- Resolve a long standing need in breeding research to screen for causative mutations/variation linked to traits of interest
- On native DNA -> no PCR amplification
- Long read sequencing technologies
- Flexibility

WORKFLOW KEYGENE'S TARSEQ TECHNOLOGY



Protocol:
1. (HMW) DNA isolation

2. Targeting loci: CRISPR enzyme-based (●)

3. Non-target DNA removal: exonuclease (ᗧ)

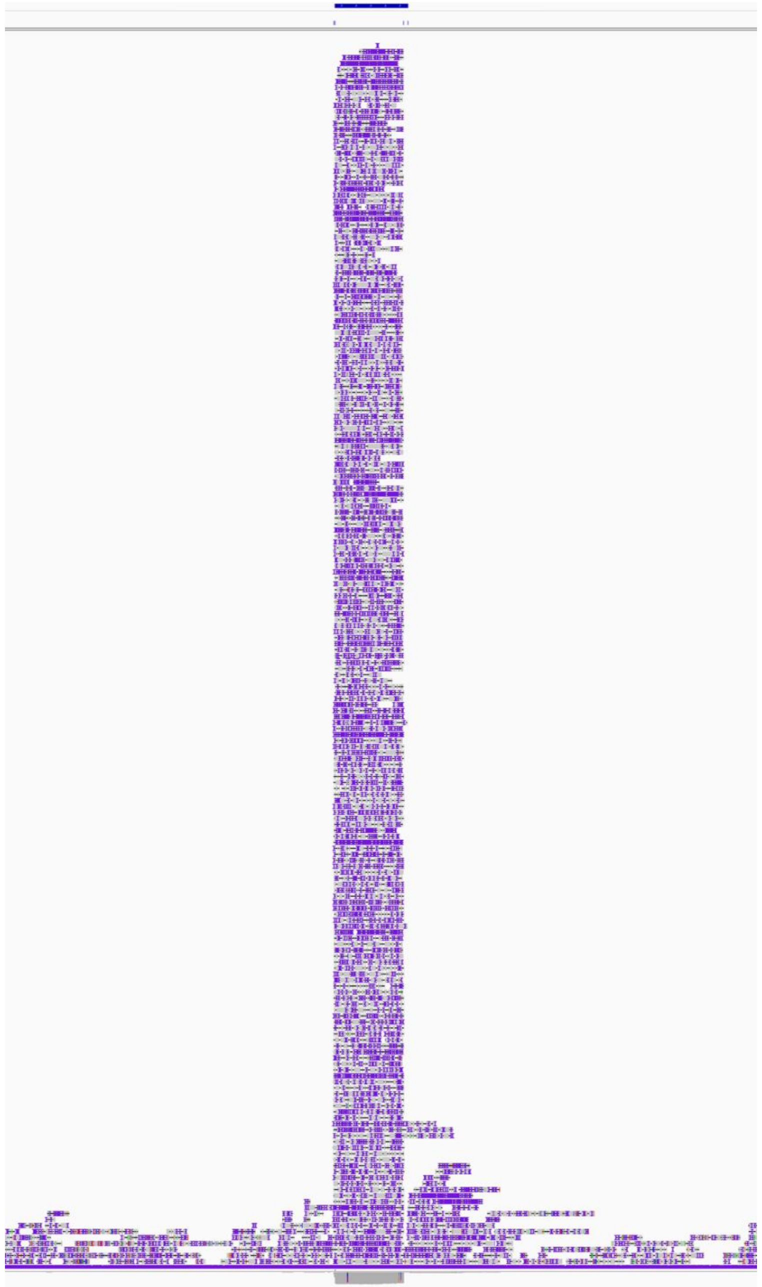
4. Clean up

5. ONT Library prep (motor protein ⬮⬮)

6. ONT Sequencing

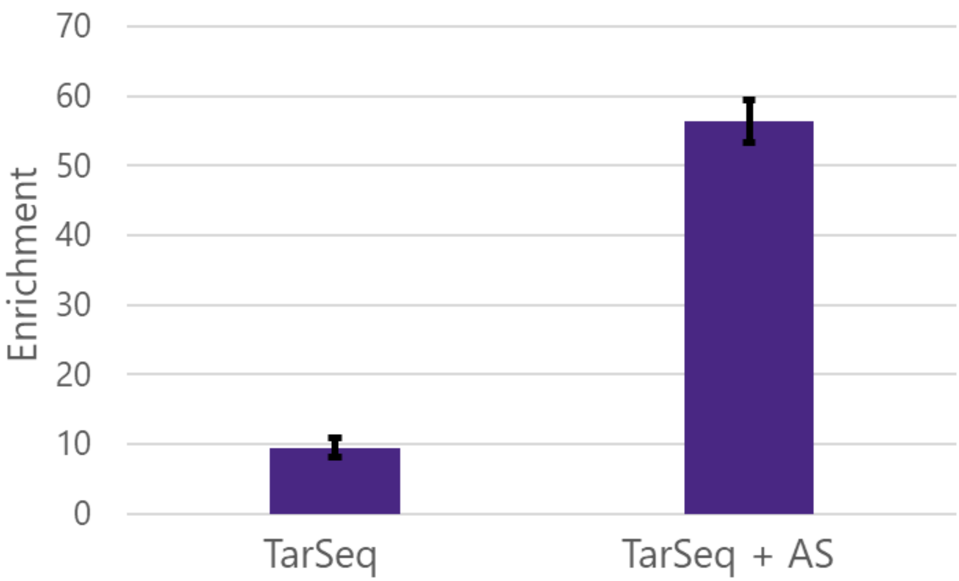
A SHOWCASE IN MELON

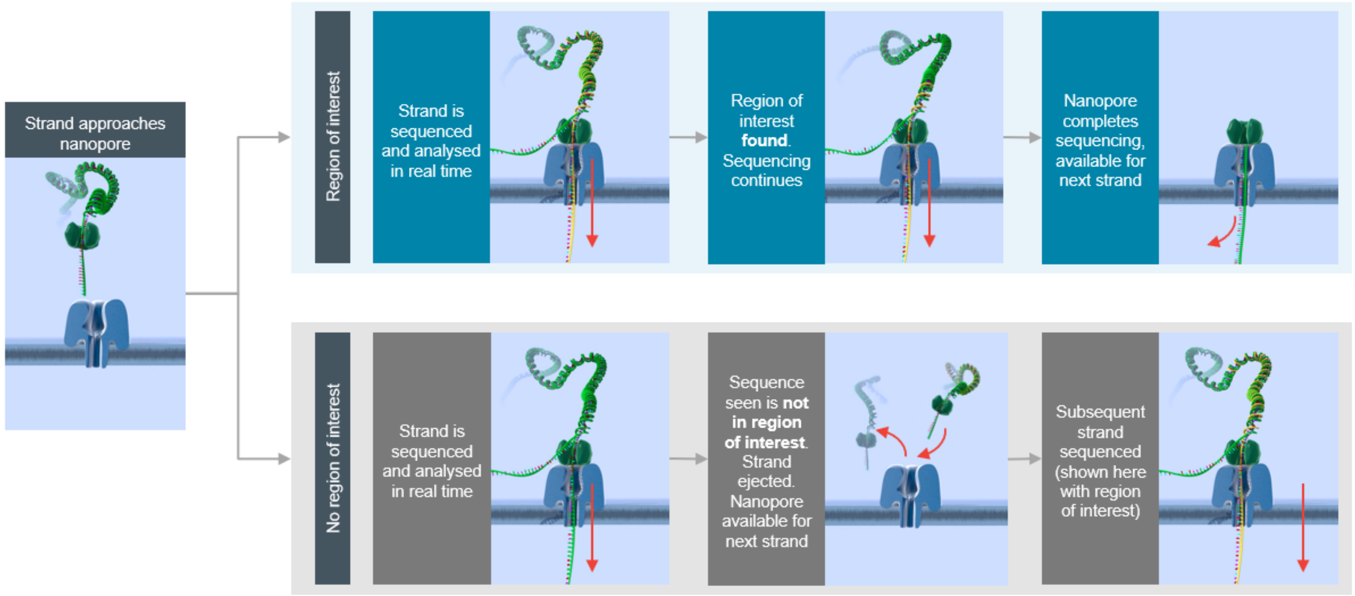
- 814 loci were targeted in 19 melon samples
- Sequencing on ONT PromethION platform
- 80% of target loci >20X coverage
- Variation detection: SNPs, structural variants & methylation



BOOSTING TARGET ENRICHMENT

- Adaptive sampling (AS)
- Combine with TarSeq
- Combined approach boosts enrichment up to ~55X!




Strand approaches nanopore

Region of interest

Strand is sequenced and analysed in real time

Region of interest **found**. Sequencing continues

Nanopore completes sequencing, available for next strand

No region of interest

Strand is sequenced and analysed in real time

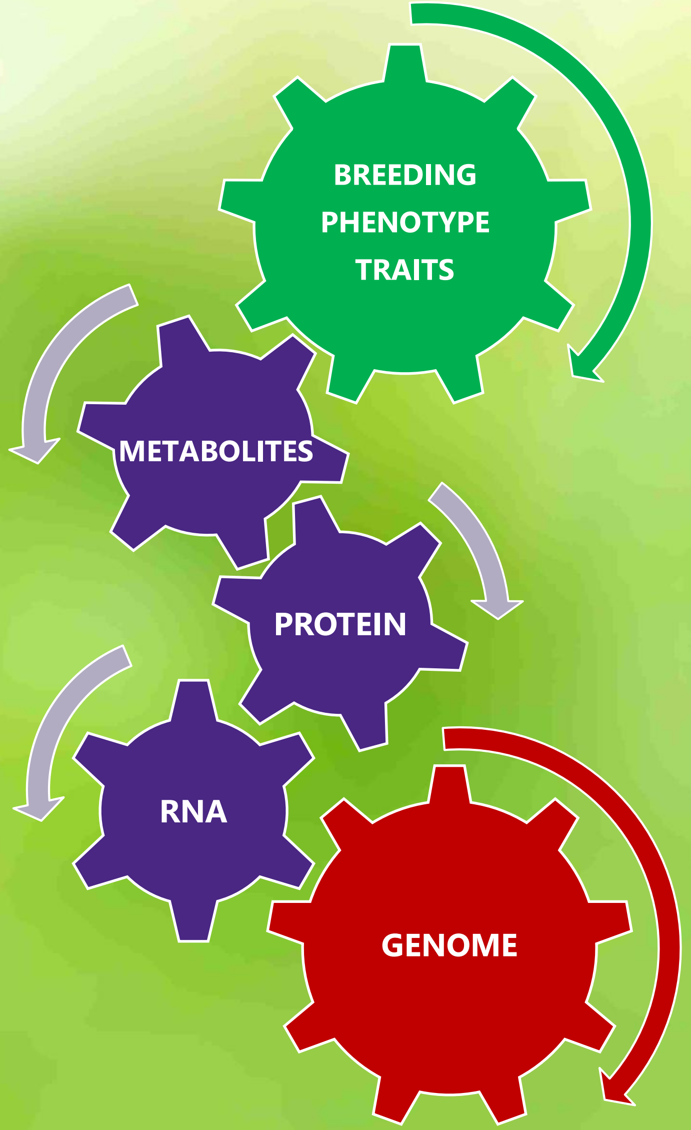
Sequence seen is **not in region of interest**. Strand ejected. Nanopore available for next strand

Subsequent strand sequenced (shown here with region of interest)

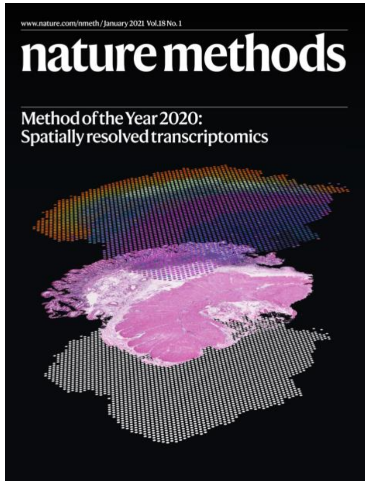
THE INNOVATIONS ENGINE: FUTURE PERSPECTIVE

- A systems biology approach:
  - integrate -omics information
  - Smooth interplay wetlab & bioinformatics

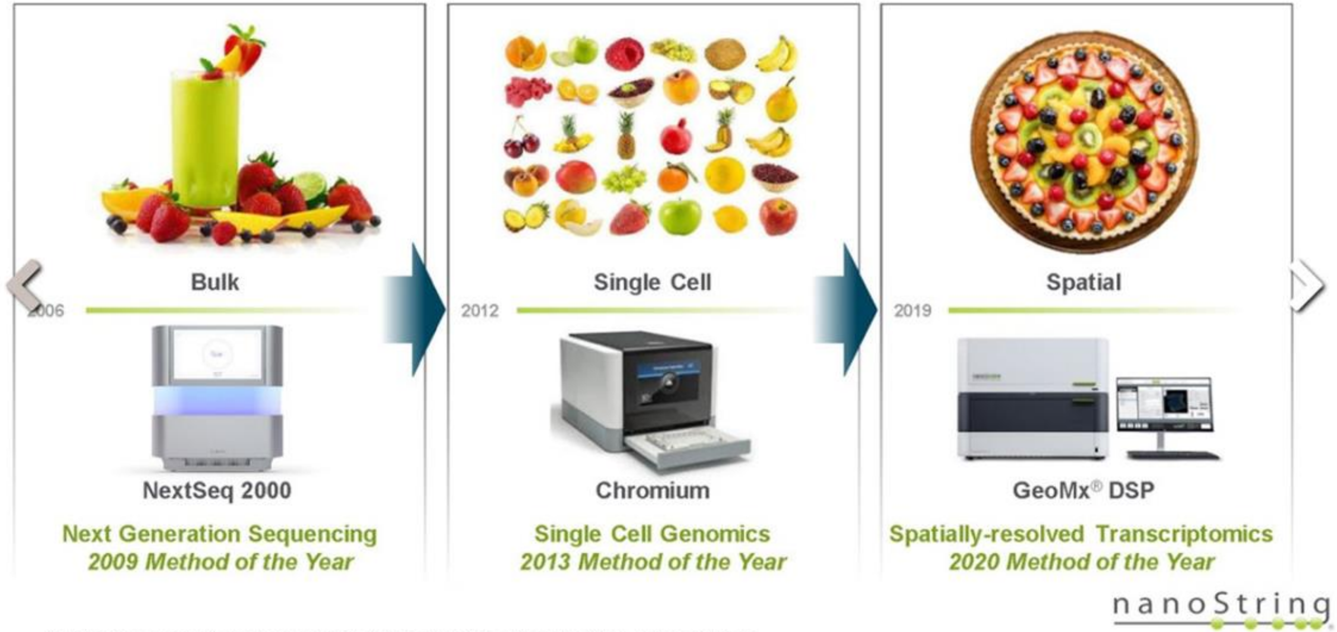
- Accelerate lead discovery and validation
  - Single Cell and Spatial technologies
  - *De novo* peptide sequencing



FUTURE: SINGLE CELL & SPATIAL TECHNOLOGIES



- Combination genomics & imaging innovations
- Measuring nucleic acids and proteins while maintaining spatial integrity even in 3D
- Main challenges
  - Cell isolation (tissues, species etc.)
  - Reporter lines for specific cell types
  - Data analysis
- Crop IB presentation Michiel Bontinck



Adapted from concept and images by Dr. Aviv Regev of The Broad Institute; *Nature Methods*

FUTURE: *DE NOVO* PEPTIDE SEQUENCING

- ‘Not to understand what could happen (genome) in crop traits, but moving towards what is happening right now (proteome)’ Jonathan M. Rothberg

- Single molecule Next-Gen protein sequencing technologies
  - Quantum-Si: time domain sequencing on semiconductor chip
  - Erisyon: fluorosequencing combined with Edman cycling
  - Encodia: ProteoCode technology

dick.roelofs@keygene.com

in @ KeyGene

f @ KeyGeneNV

@ KeyGeneInfo

@ _KeyGene