

2.2. Post-assembly analyses

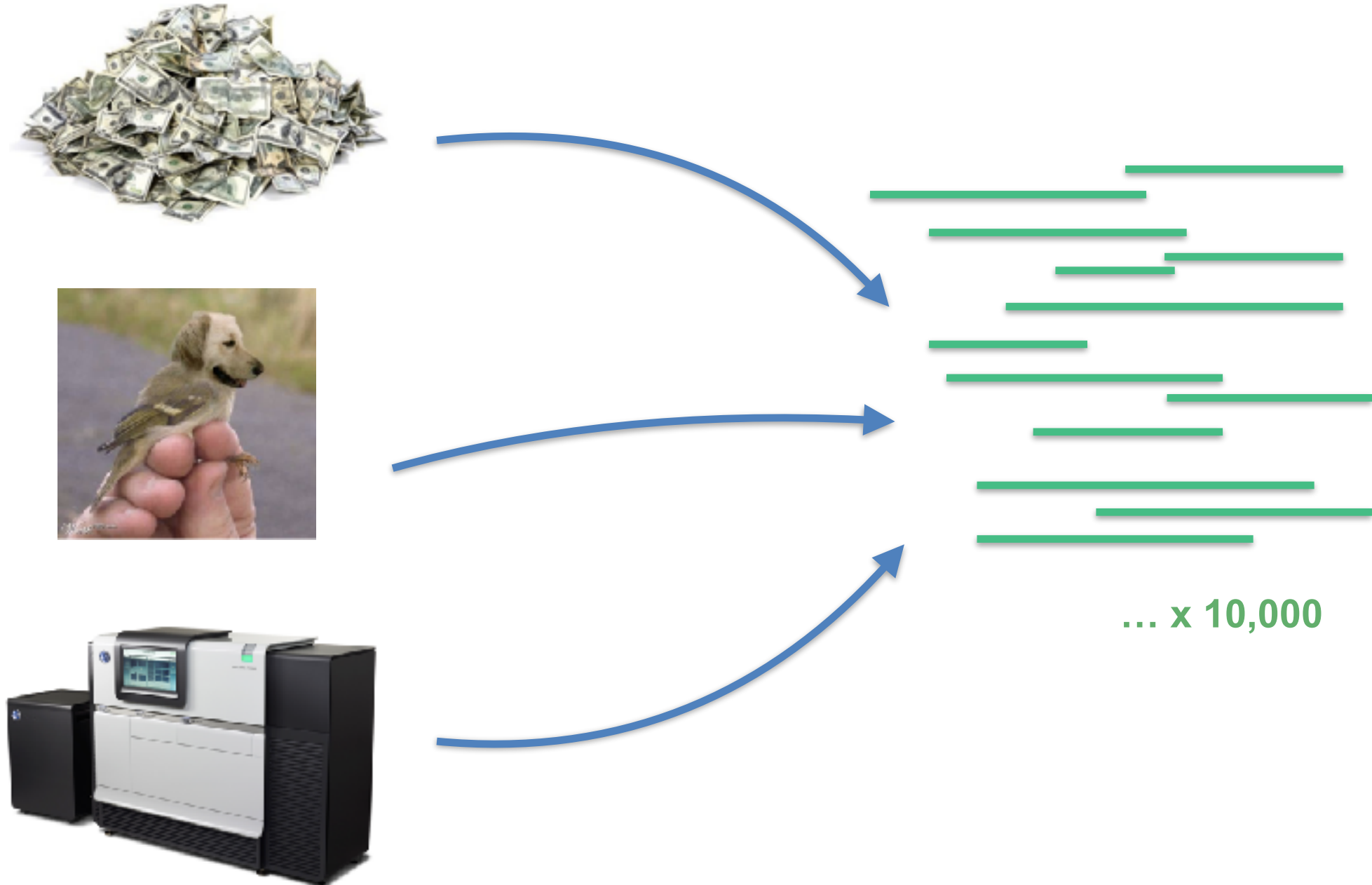
1. Long range scaffolding of sequence contigs

- Hi-C
- Optical mapping

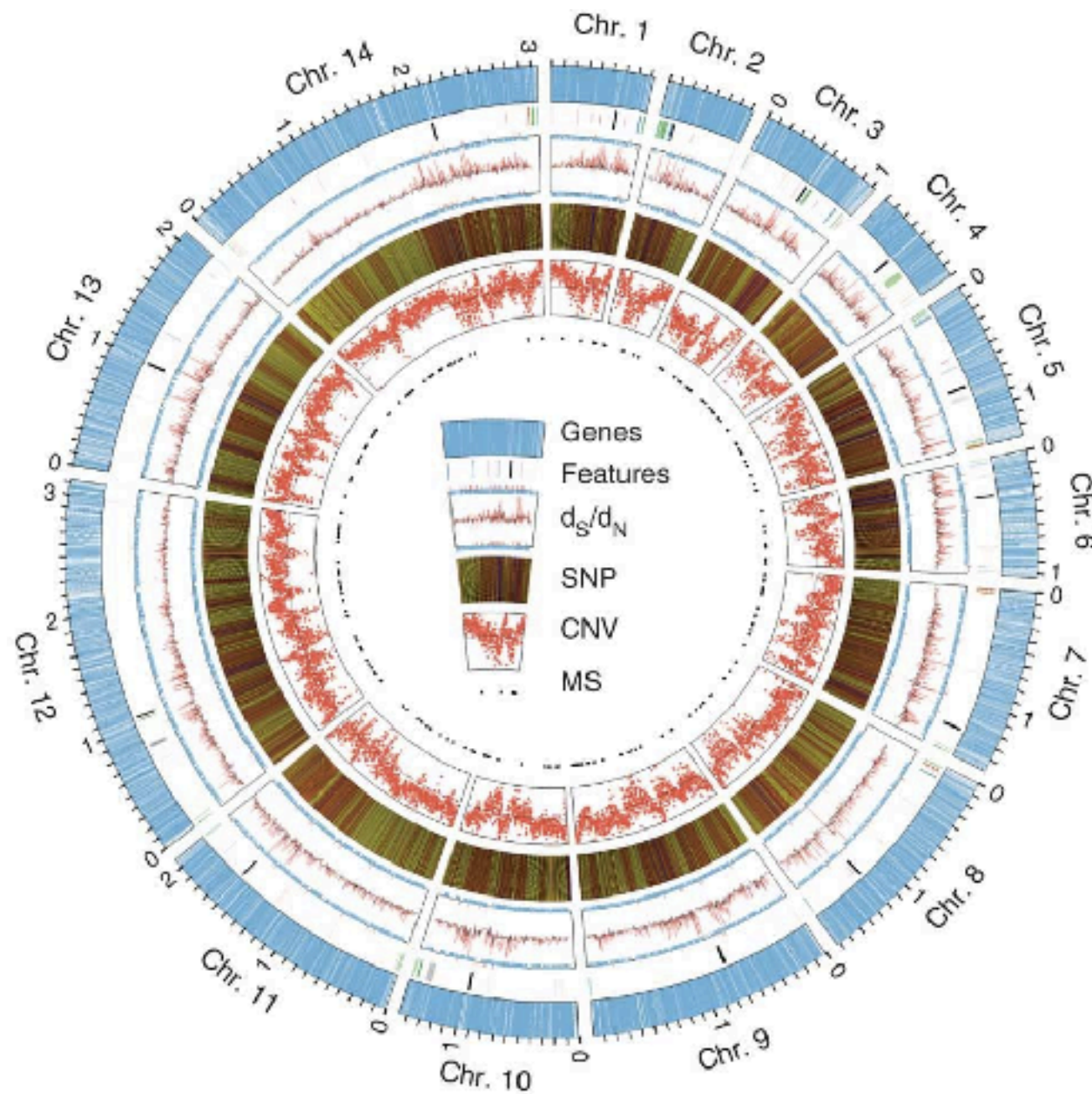
2. Anchoring and ordering your scaffolds along chromosomes

- Creating and using linkage maps
- Hands on. . .

2.2. Post-assembly analyses



2.2. Post-assembly analyses



2.2.1 Long range scaffolding

Genome



Reads



Contigs

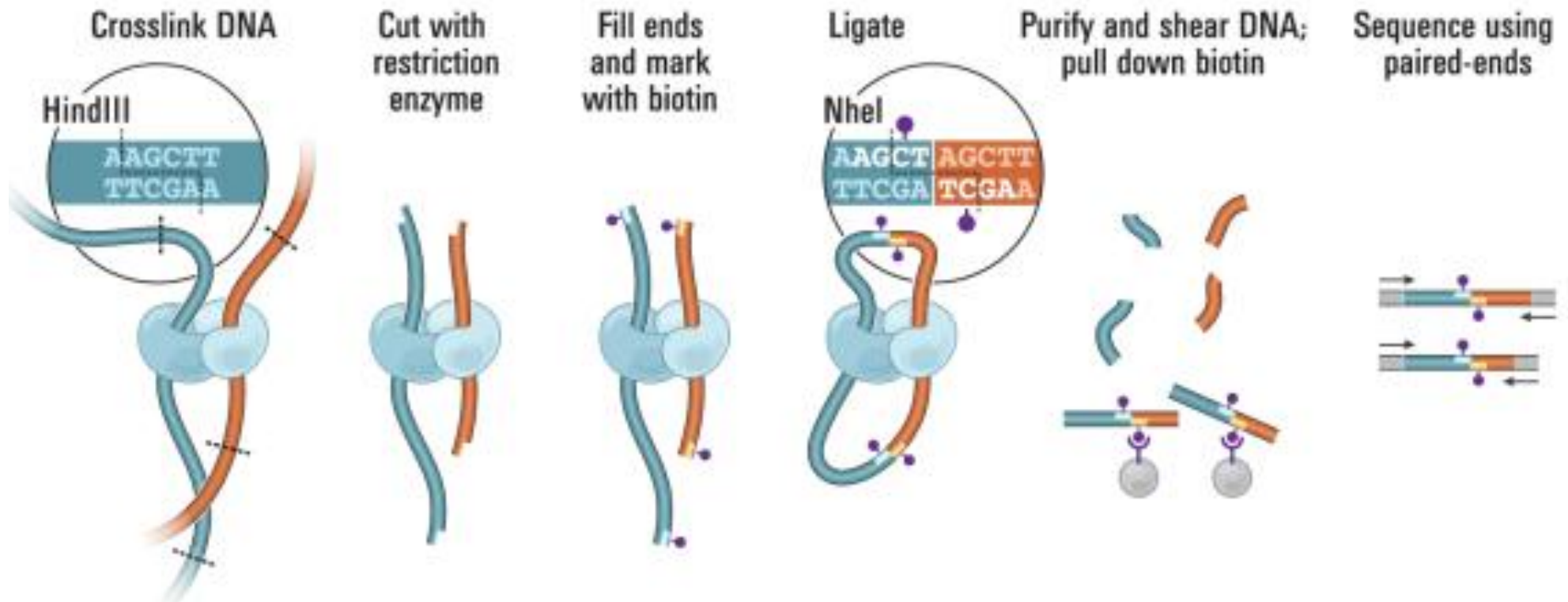


Scaffolds



2.2.1 Long range scaffolding

- Hi-C



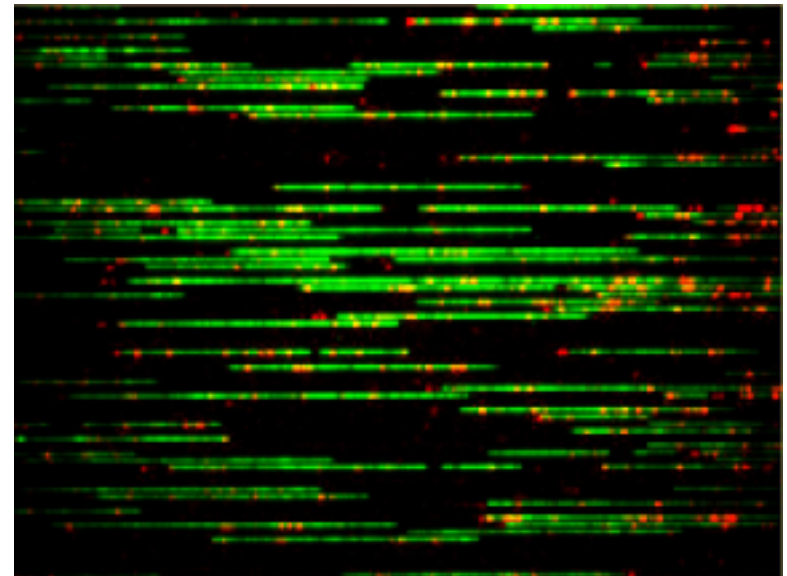
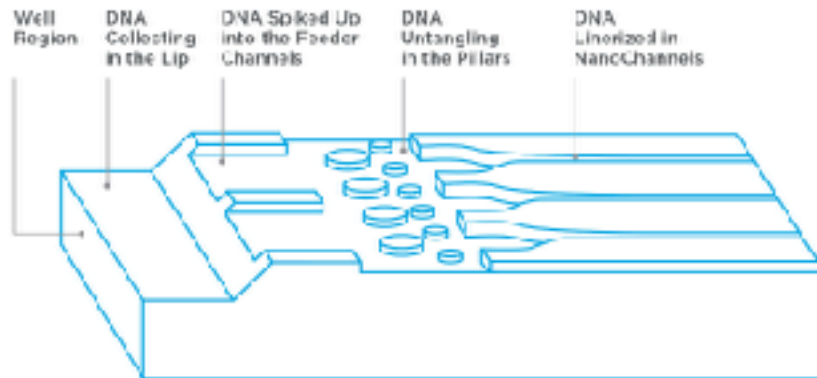
2.2.1 Long range scaffolding

Optical mapping



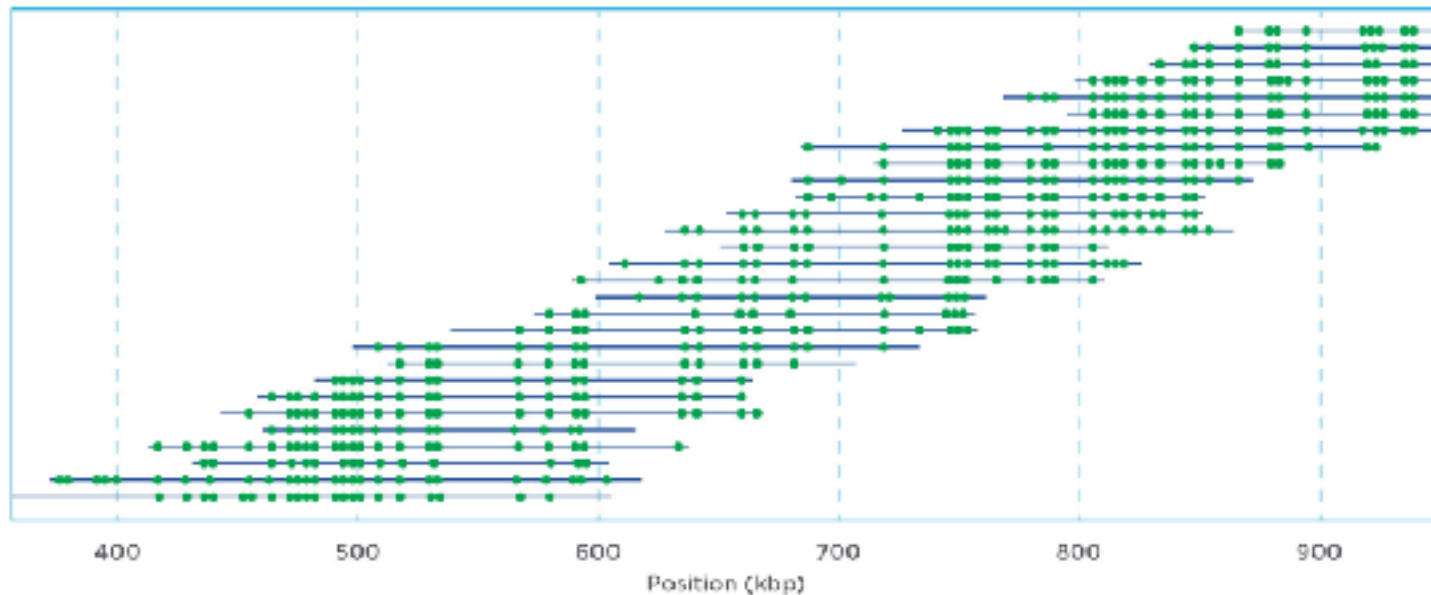
<https://bionanogenomics.com/products/saphyr/>

SAPHYRCHIP™ SCHEMATIC



2.2.1 Long range scaffolding

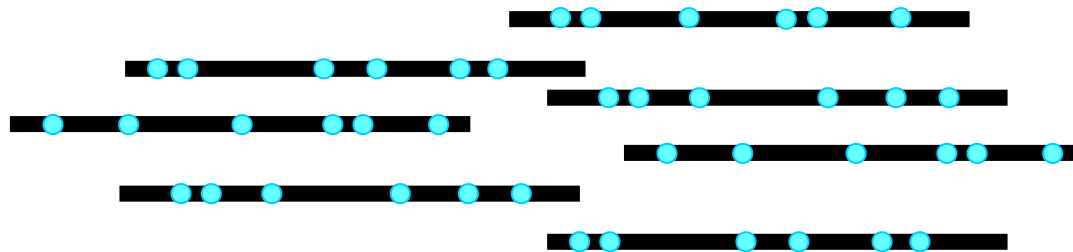
- Assemble Maps



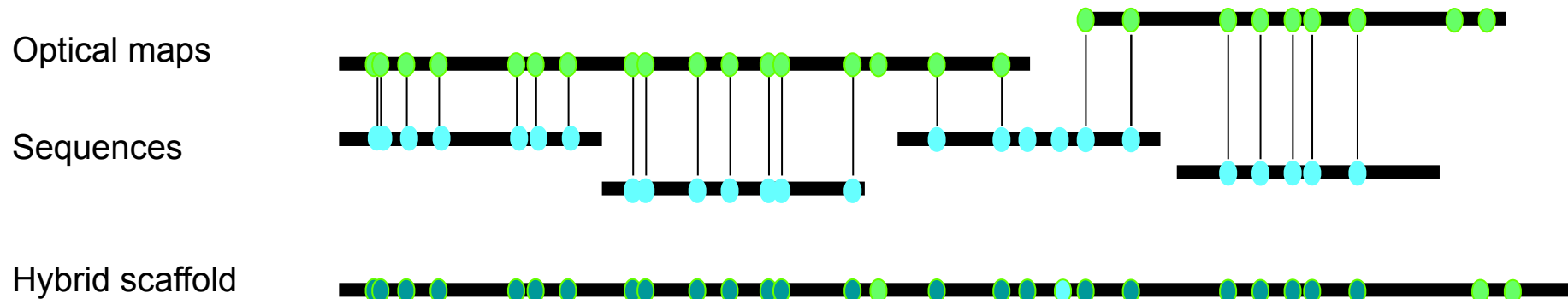
2.2.1 Long range scaffolding

- In-silico digest of existing sequence assembly

ATCGCTGAC**CCAGT**GACTAGCTCCATCTTAGG**CCAGT**ATCTCCAGATCGCTG**CCAGT**AGACTAGCTCCATA



- Combine with optical maps to make hybrid assembly



2.2.1 Long range scaffolding

Some examples



Single-molecule sequencing and chromatin conformation capture enable *de novo* reference assembly of the domestic goat genome

© All rights reserved.

Derek M Bickhart^{1,18}, Benjamin D Rosen^{2,18}, Sergey Koren^{3,18}, Brian L Sayre⁴, Alex R Hastie⁵, Saki Chan⁶, Joyce Lee⁵, Ernest T Lam⁵, Ivan Liachko⁸, Shawn T Sullivan⁷, Joshua N Burton⁶, Heather J Huson⁸, John C Nystrom⁸, Christy M Kelley⁷, Jana L Hutchison², Yang Zhou^{2,10}, Jijie Sun¹¹, Alessandra Crisà¹², F Abel Ponce de León¹³, John C Schwartz¹⁴, John A Hammond¹⁴, Geoffrey C Waldbieser¹⁵, Steven G Schroeder², George E Liu², Maitreya J Dunham⁶, Jay Shendure^{6,16}, Tad S Sonstegard¹⁷, Adam M Phillippy³, Curtis P Van Tassel² & Timothy P L Smith⁹

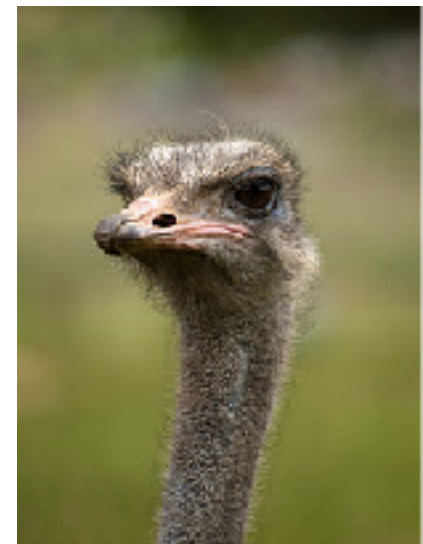
PacBio = 3.7 Mb —→ **OM** —→ N50 = 20.6 Mb

Improving the ostrich genome assembly using optical mapping data

Zhang et al. 2015

(<https://gigascience.biomedcentral.com/articles/10.1186/s13742-015-0062-9>)

Illumina N50 = 3.59 Mb —→ **OM** —→ N50 = 17.71 Mb



Hummingbird



PacBio contig N50 = 6.09 Mb

One enzyme scaffold N50 = 14.78

Two enzyme scaffold N50 = 25.08

Rana temporaria

PB N50 = 0.64Mb

OM staff N50 = 2 Mb

Fragile sites



A few more examples

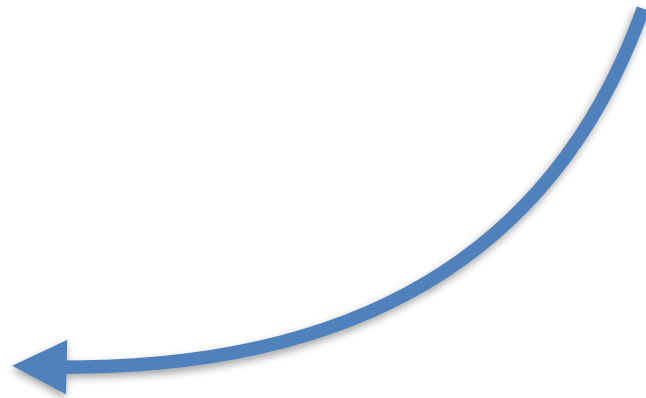
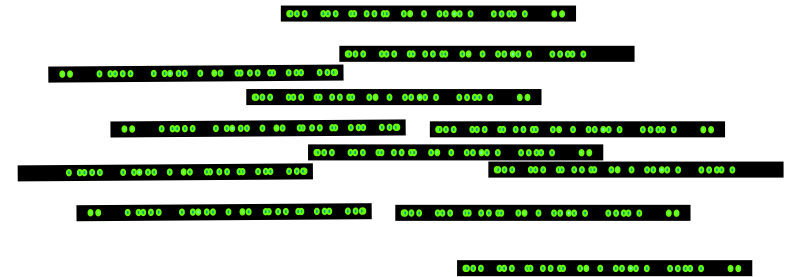
Sea Bass: <http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1005954>

Wheat: <http://onlinelibrary.wiley.com/doi/10.1111/pbi.12513/full> (With chromosome sorting)

Human (Korean): <http://www.nature.com/nature/journal/v538/n7624/full/nature20098.html#supplementary-information>

2.2.2. Anchoring contigs/scaffolds

Using linkage maps



ALLMAPS
Tang et al 2015

Integrate one OR MORE genetic maps
to order and orient genome scaffolds

2.2.2. Anchoring contigs/scaffolds

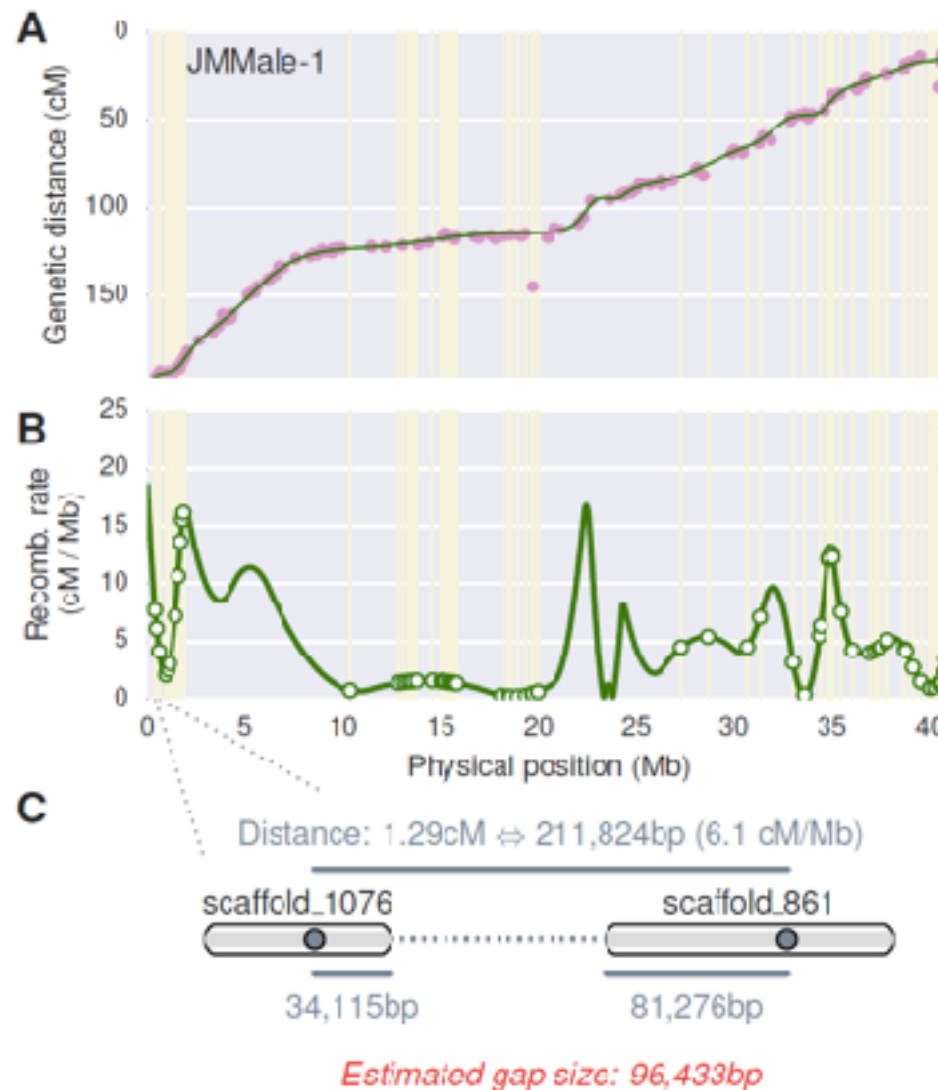
HANDS ON

Go to “TUTORIAL: Anchoring contigs and scaffolds using linkage maps”



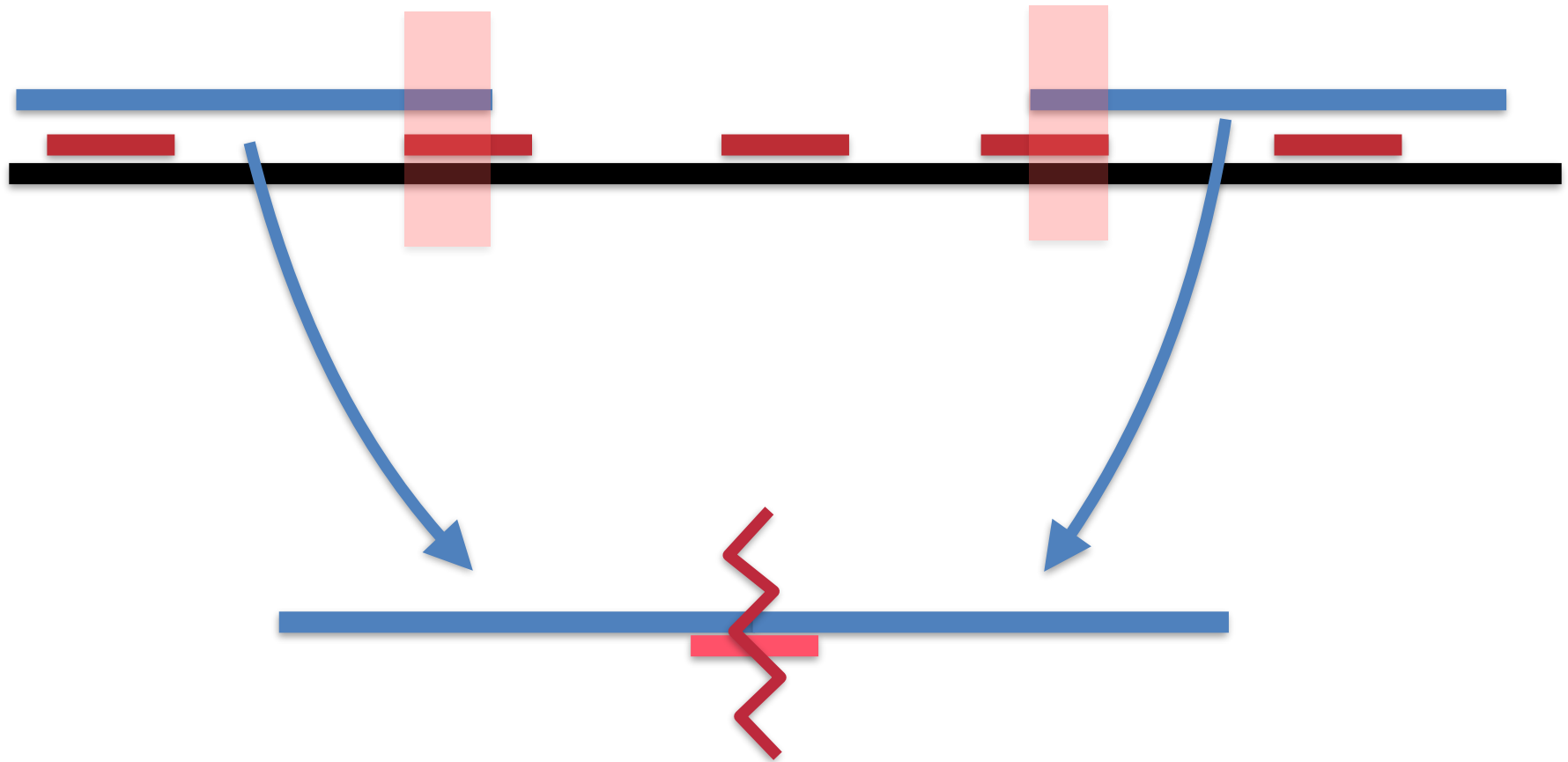
2.2.2. Anchoring contigs/scaffolds

Estimate gap sizes



2.2.2. Anchoring contigs/scaffolds

Split chimeric contigs



2.2.2. Anchoring contigs/scaffolds

Split chimeric contains



<http://catchenlab.life.illinois.edu/chromonomer/>

