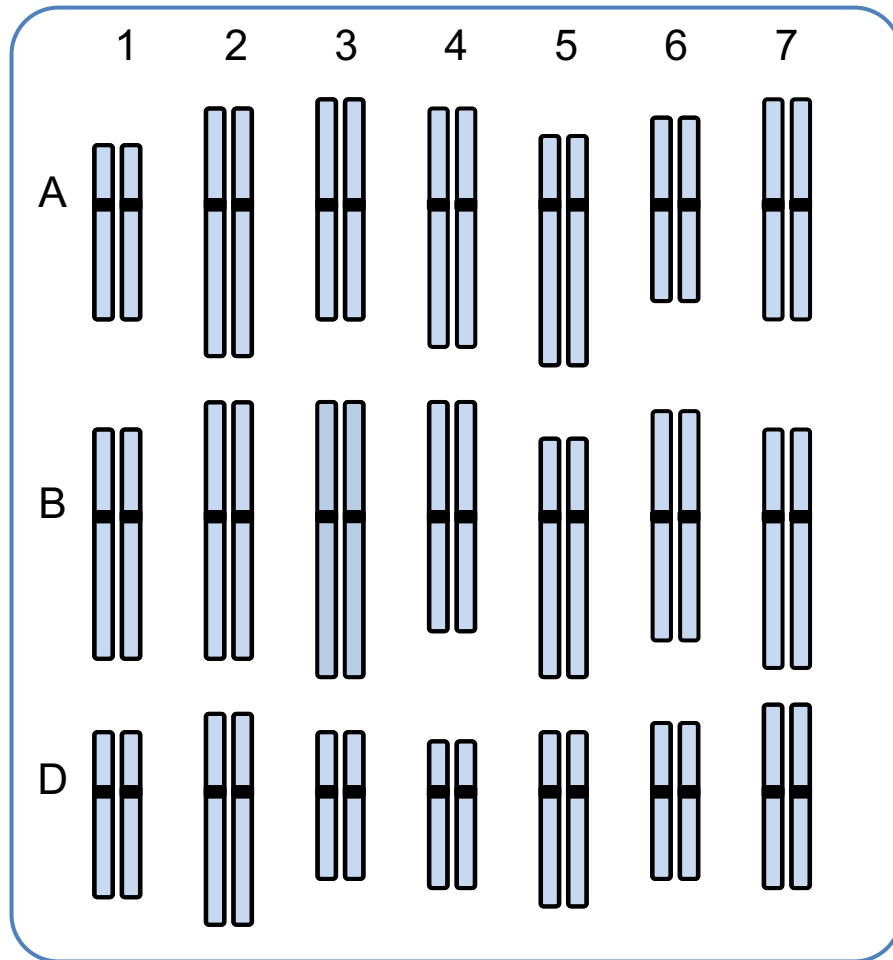
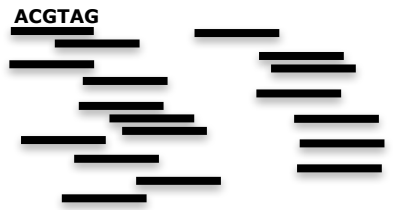


Bread wheat whole genome sequence assembly

Frédéric Choulet







reads



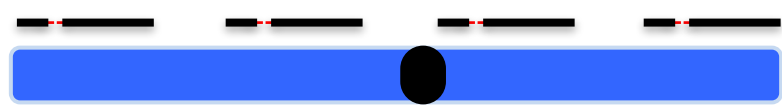
contigs



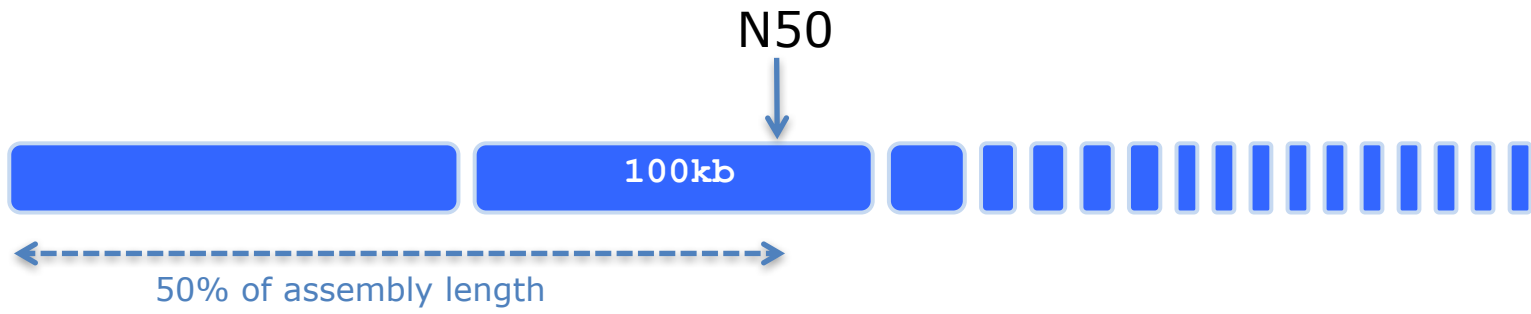
scaffolds



superscaffolds



pseudomolecule

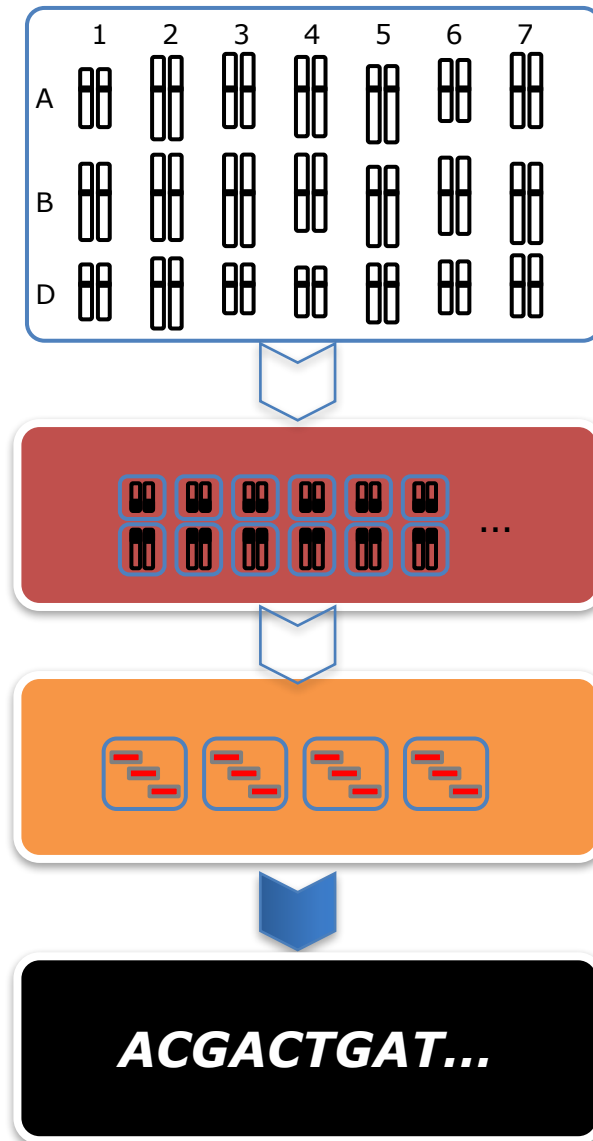




Launched in 2005

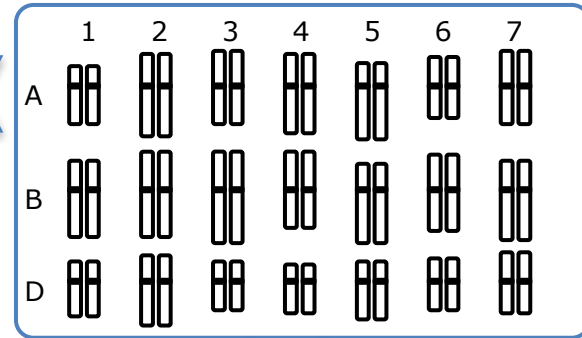
- Produce a **high quality** genome sequence of **hexaploid** wheat

Chr-by-chr approach



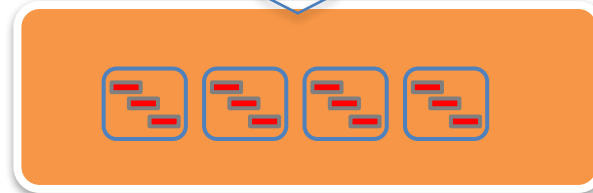
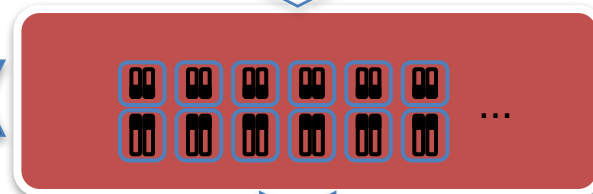
WGS

AABBDD *Brenchley et al. 2012*
AA *Ling et al. 2013*
DD *Jia et al. 2013*
AABBDD *Chapman et al. 2015*



CSS = Chr. Survey Seq

IWGSC Science 2014



BACseq

Choulet et al. Science 2014

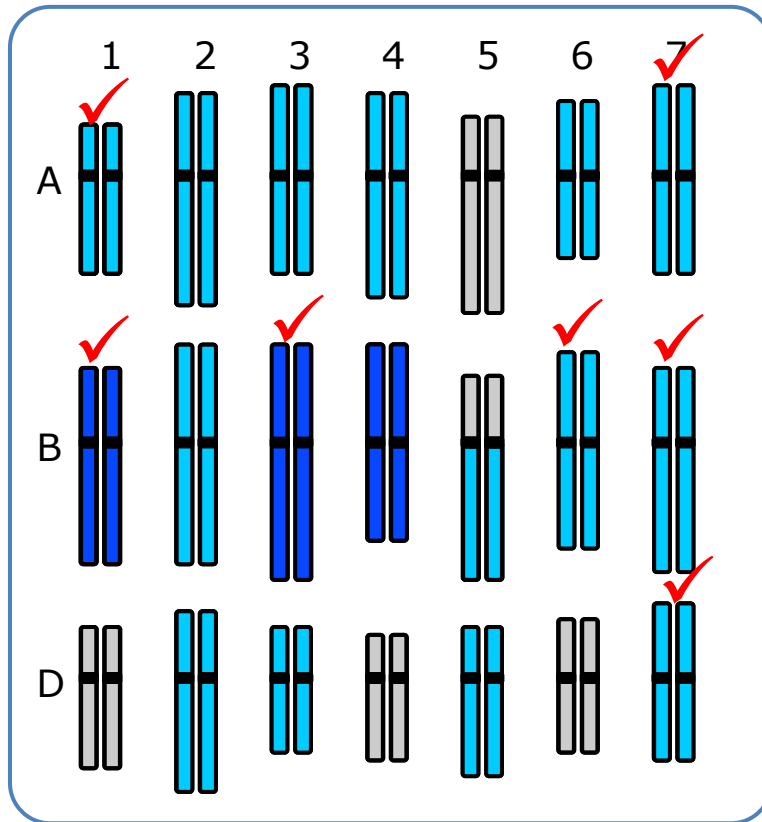
3B

ACGACTGAT...

N50

2012	AABBDD	BBSRC	1 kb
2013	AA	BGI	64 kb
2013	DD	BGI	58 kb
2014	AABBDD	IWGSC-CSS-draft	2 kb
2014	3B	GDEC/Genoscope	892 kb
2015	AABBDD	JGI	21 kb

○ Chr-by-chr approach in 2015



GDEC
Génoscope
URGI

3B



1B & 4B



○ Wild Emmer Wheat (AABB)

U Tel Aviv, NRGene, U Sask...

NRGene:

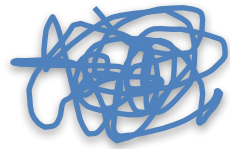
WGS Illumina 180x, 3 MP lib

DeNovoMAGIC™ 2.0

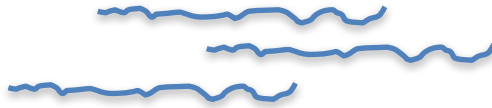
N50

2012	AABBDD	BBSRC	1 kb
2013	AA	BGI	64 kb
2013	DD	BGI	58 kb
2014	AABBDD	IWGSC-CSS-draft	2 kb
2014	3B	GDEC/Genoscope	892 kb
2015	AABBDD	JGI	21 kb
2015-Jul	AABB	Wild-Emmer-NRGene	7000 kb

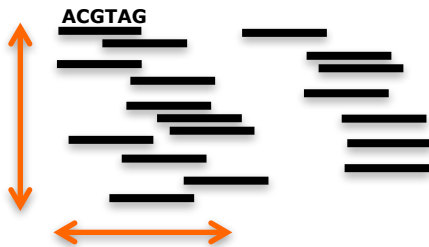
○_○



Library Prep.



Sequencing



Assembly

DeNovoMAGIC™ 2.0



Anchoring



○ IWGSC-Whole Genome Assembly (AABBDD)



N. Stein



C. Pozniak



J. Poland



A. Diestelfeld



A. Scharpe



G. Ronen



M. Thompson

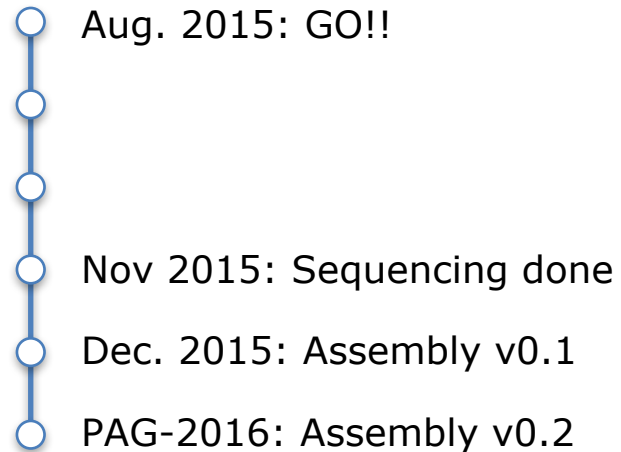


K. Eversole, J. Rogers



F. Choulet

+ *IWGSC chr. leaders*



N50

2012	AABBDD	BBSRC	1 kb
2013	AA	BGI	64 kb
2013	DD	BGI	58 kb
2014	AABBDD	IWGSC-CSS	2 kb
2014	3B	GDEC/Genoscope	892 kb
2015	AABBDD	JGI	21 kb
2015-Jul	AABB	Wild-Emmer-NRGene	7000 kb
2016-Jan	AABBDD	IWGSC-WGA	7394 kb
-	-	TGAC	80 kb

o_o

○ IWGSC-Whole Genome Assembly (AABBDD)

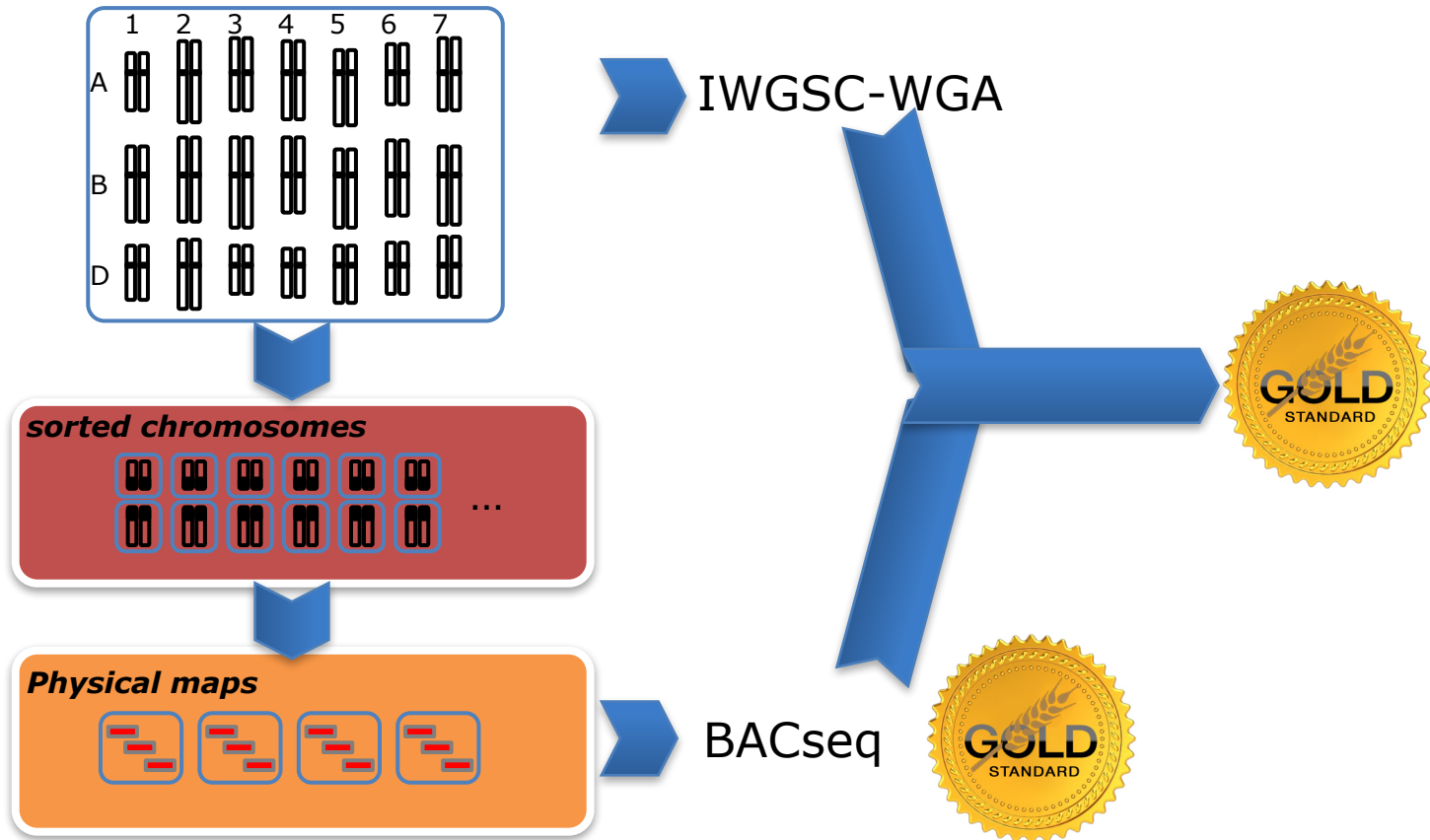
1. QC

- Contiguity: +++
- Completeness: +++ (97..99%)
- Chimeras?

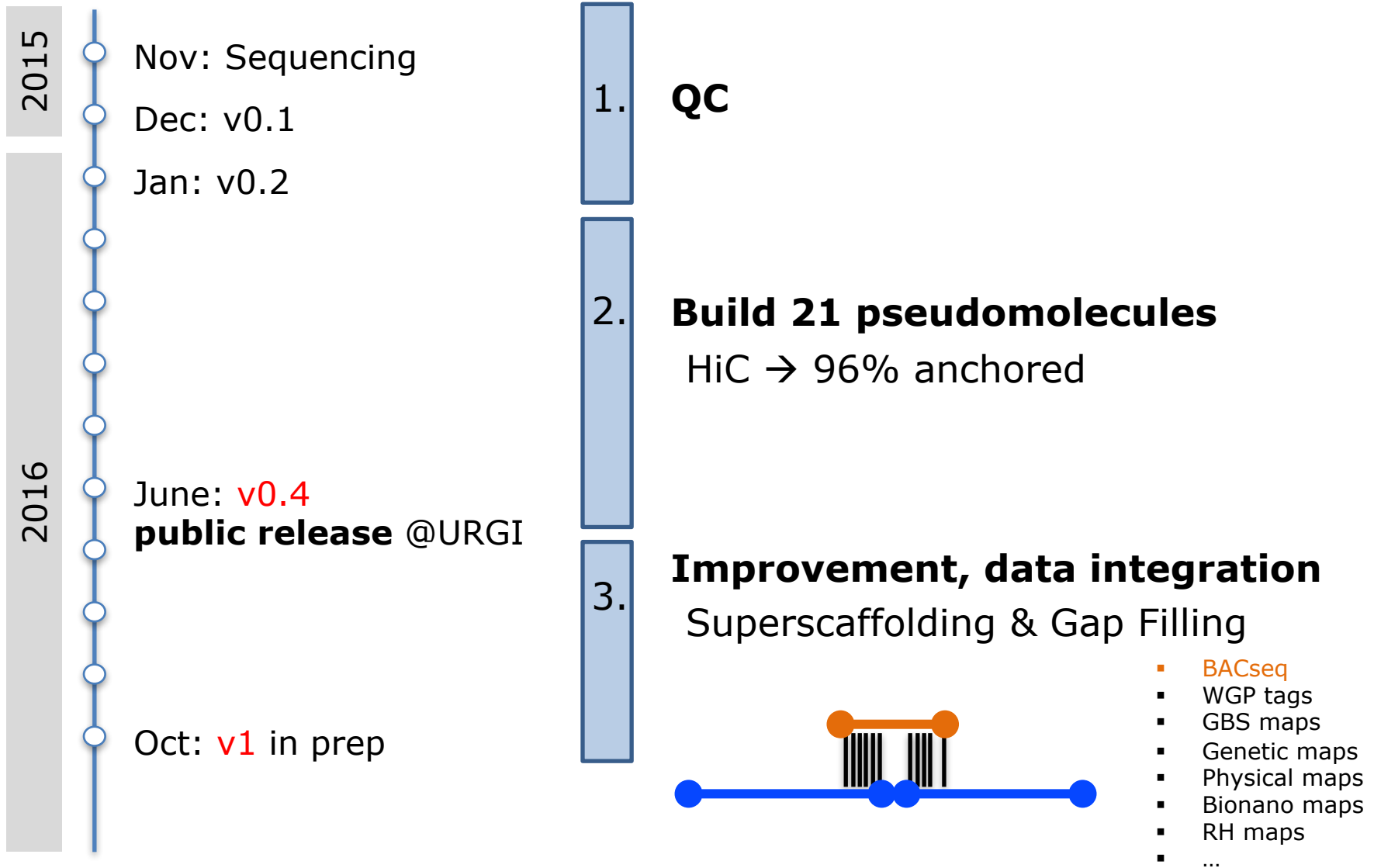


~**200** chimeric scaff (1.5 Gb) corrected

- Update the IWGSC road map



○ IWGSC-Whole Genome Assembly



○ Assembly Metrics

v0.5.4 (sep2016)

total size	14.5 Gb
Ns (gaps)	1.8%
N50	7 Mb (571 scaff)
N99	3 kb (24,037 scaff)
max scaff size	46 Mb

21 pseudomolecules:

total size	14.1 Gb (97%)
nb superscaffolds	1608 (77/chr)

○ **Current work and Perspectives**

- **Freeze: v1 assembly**
- **Annotation**
 - GDEC (TriAnnot)
 - MIPS
- **Biological analyses**
- **Publication in 2017**

Thanks



IWGSC-WGA working team

GDEC

- **Ambre-A. Josselin**
- **Hélène Rimbart**
- Philippe Leroy
- Nicolas Guilhot
- Benoit Darrier
- Pierre Sourdille
- François Balfourier
- Romain De Oliveira
- Jonathan Kitt
- Etienne Paux

IPK, Germany

- **Martin Mascher**
- Nils Stein

U Murdock, Australia

- **Gabriel Keeble**
- Rudi Appels

U Sask., Canada

- **Kevin Koh**
- Curtis Pozniak

Kansas State U, US

- Jesse Poland

MIPS, Germany

- Klaus Mayer
- Manuel Spannagel
- Sven Twardziok

NRGene

- Gil Ronen, et al.

INRA URGI

- Michael Alaux

UZH, Switzerland

- Beat Keller

U Tel Aviv, Israel

- Assaf Distelfeld

GIFS, Canada

- Andy Scharpe

IWGSC

- Catherine Feuillet
- Jane Rogers
- Kellye Eversole

