



# **A high quality chromosome-based reference sequence for hexaploid bread wheat**

## **Wheat Initiative Associated Programme update**

Catherine Feuillet

IWGSC

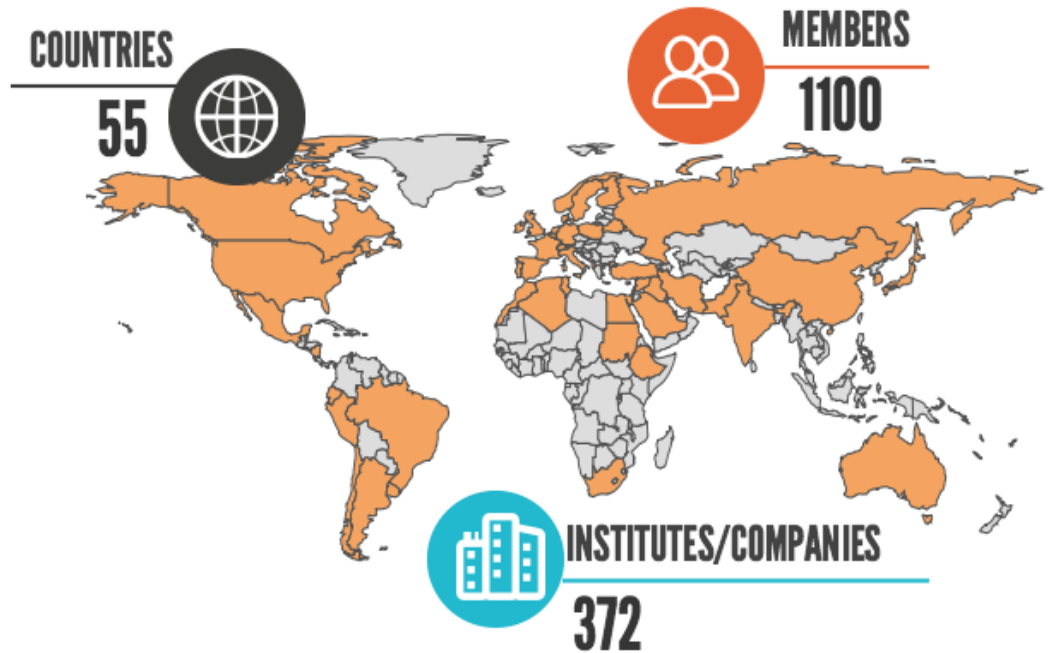
8 January 2016 – Wheat Initiative Research Committee



# The International Wheat Genome Sequencing Consortium



# 2015



**SPONSORS**  
21



# Vision and road map



## Goal

- Lay a foundation to accelerate wheat improvement
- Increase profitability throughout the industry

## Vision

- High quality annotated genome sequence, comparable to rice
- Physical map-based, integrated and ordered sequence

Chromosomal-based  
Physical Maps

MTP Sequencing

Chromosomal  
Survey

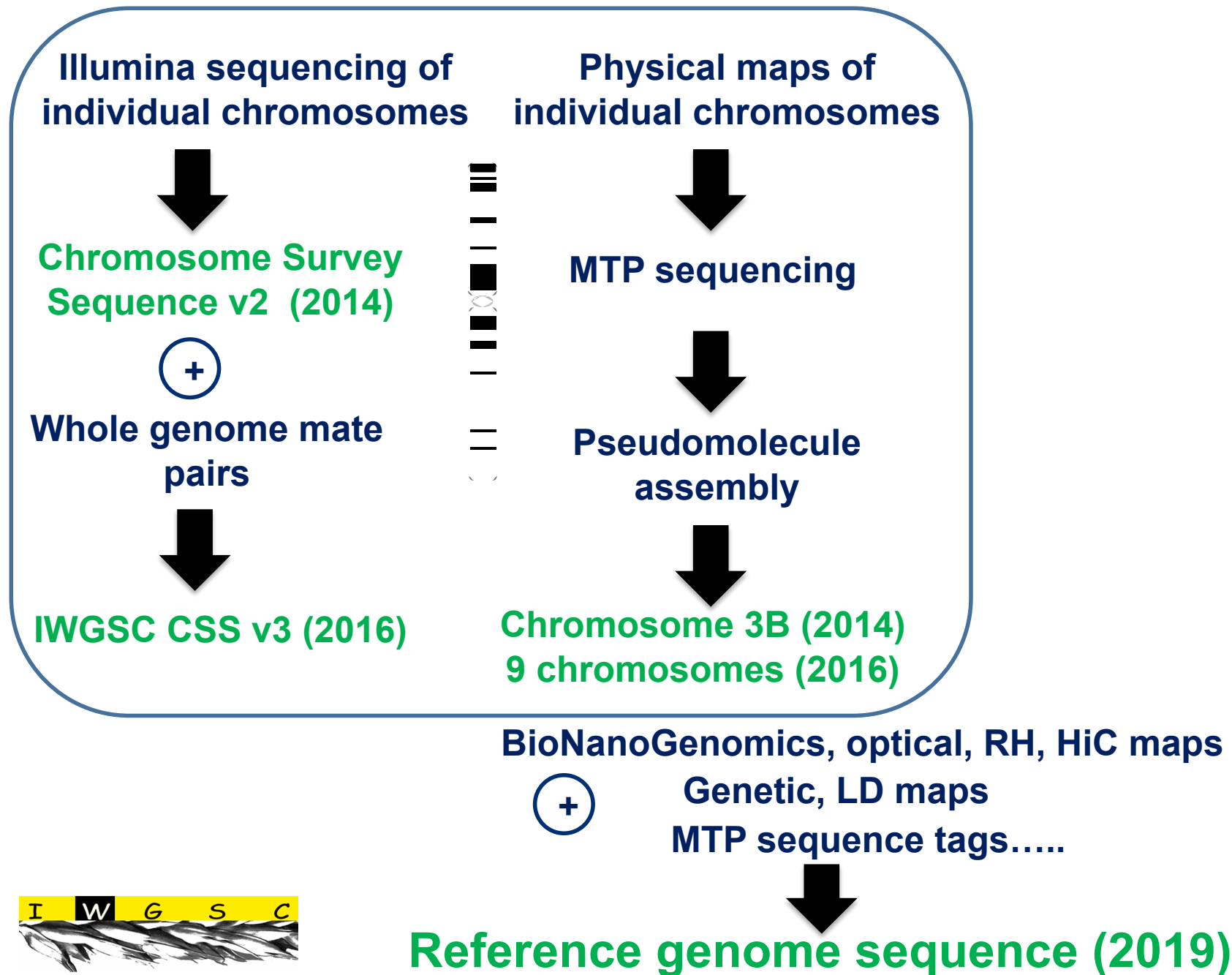
Sequence

Whole Genome  
Shotgun Assembly

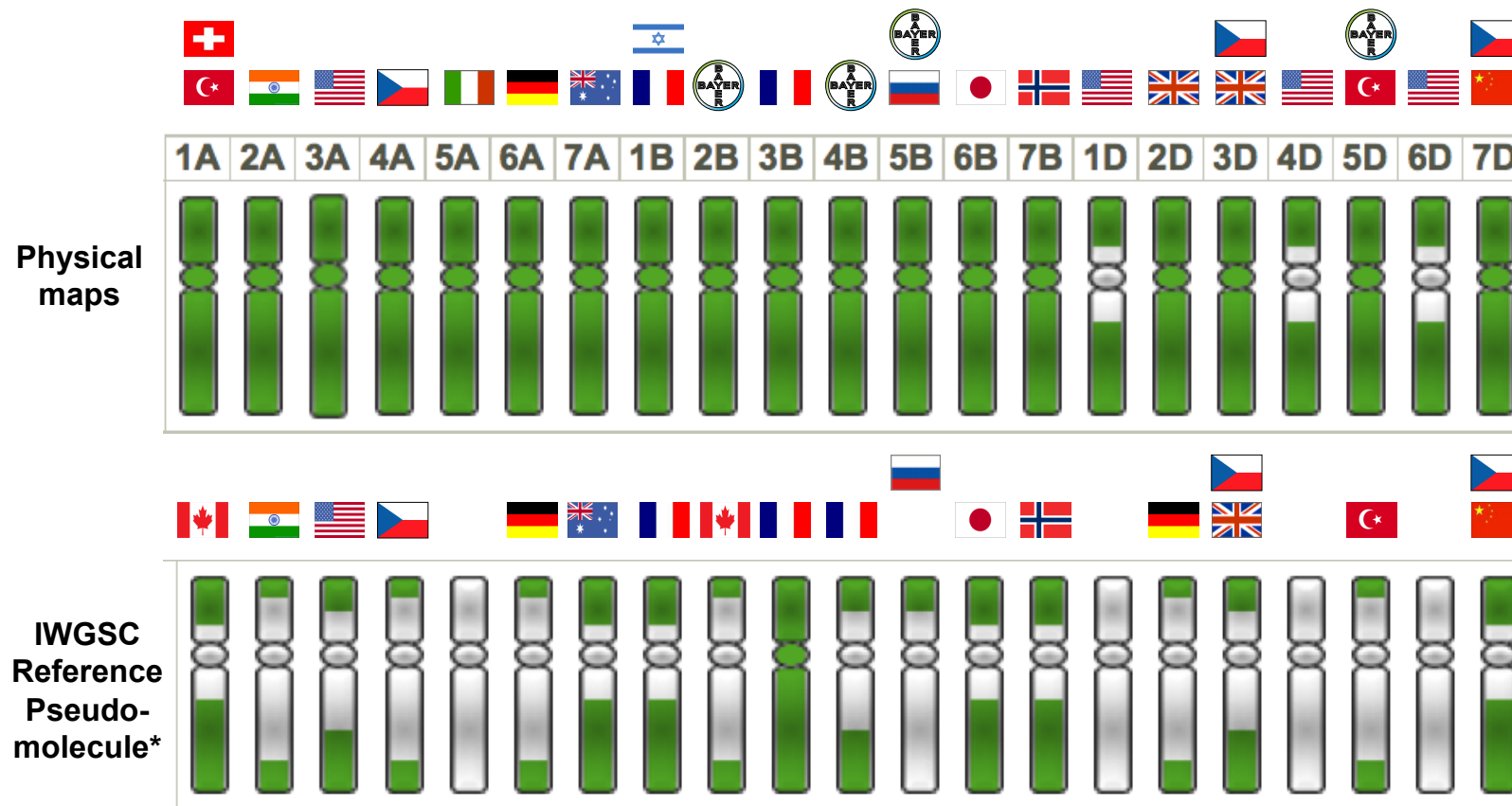
A REFERENCE SEQUENCE LINKED TO GENETIC AND  
PHENOTYPIC MAPS



# Roadmap to the Wheat Genome Sequence



# Chromosome based sequences- Progress



Physical maps available at IWGSC repository for all except 3AL

1 chromosome completed, 15.5 underway.

\*Flags represent countries where work is underway with funding, as of January 2016

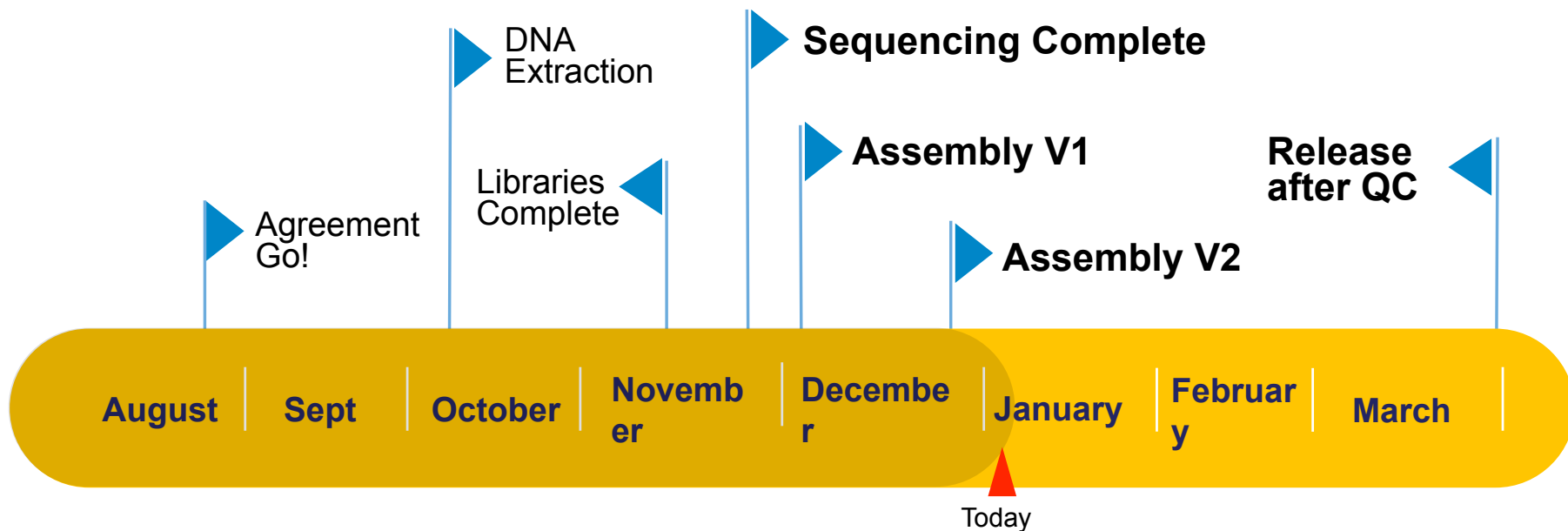


# Whole Genome Shotgun- IWGSC WGA project

IWGSC coordinated - Led by Nils Stein, Curtis Pozniak, Jesse Poland



~2 months from data accumulation to completion of first assembly



4 talks planned in main workshops at PAG



# Whole Genome Shotgun- IWGSC WGA project



## De novo assembly:

- NRGene's DeNovoMagic-2 platform, total run time < 3 weeks, 1Tb RAM computer
- illumina short-reads sequencing data only (x200 coverage)

---

<b>Assembly size:</b>	<b>14.5 Gbp</b>
<b>Gaps size:</b>	<b>262 Mbp</b>
<b>Gaps %:</b>	<b>1.80</b>
<b>Total # scaffolds:</b>	<b>138,484</b>
<b>L50:</b>	<b>7.1Mbp</b>
<b>N50 (#sequences):</b>	<b>566</b>
<b>L90:</b>	<b>1.3 Mbp</b>
<b>N90 (#sequences):</b>	<b>2,363</b>
<b>MAX Scaffold:</b>	<b>45.8 Mbp</b>

---

- 95 to 99% of the genes, TE based markers present in assembly
- A number of chimeric assemblies to be resolved





# IWGSC Whole Genome Assembly Data Release

IWGSC Data Repository hosted by INRA-URGI

<http://wheat-urgi.versailles.inra.fr/Seq->

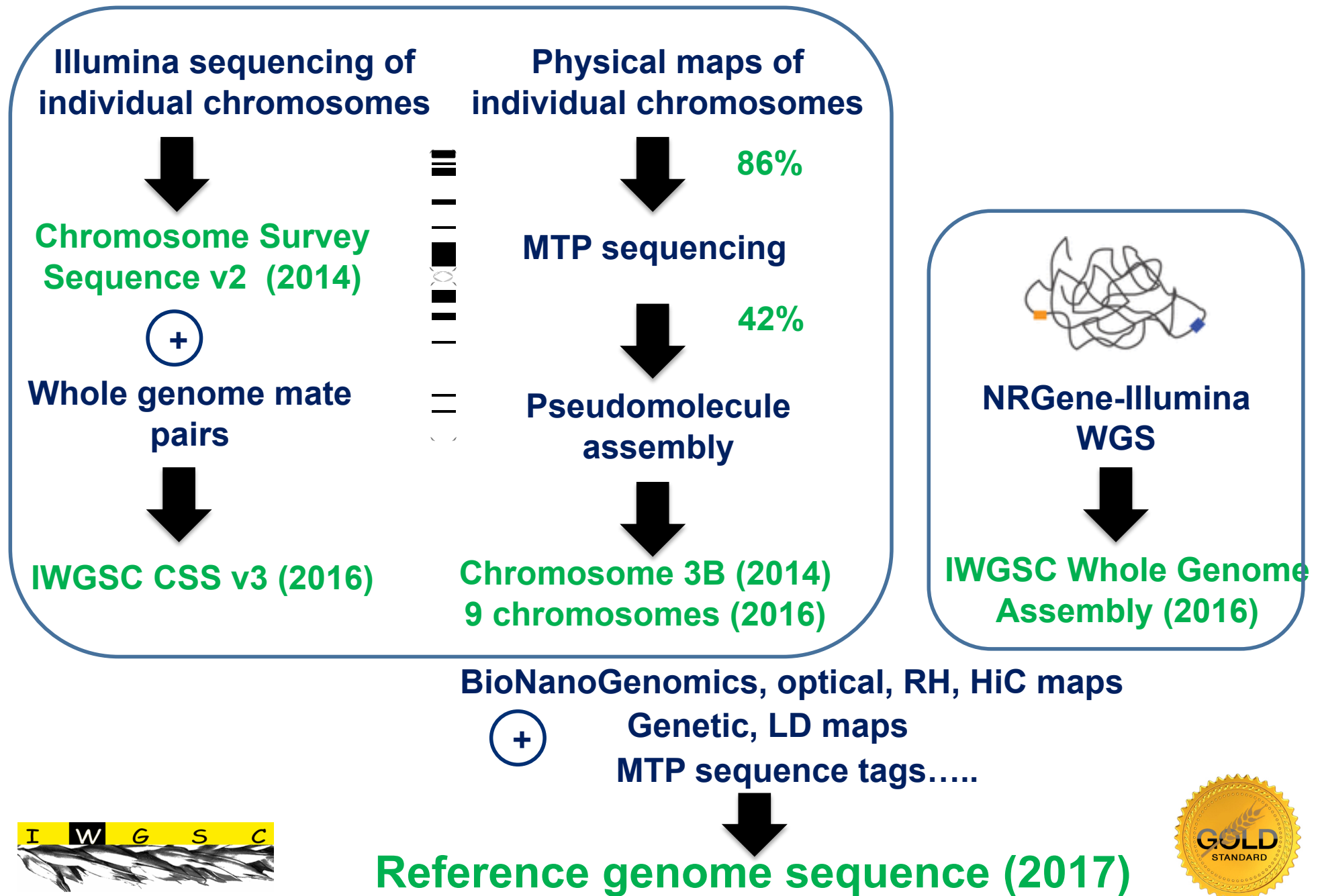
The screenshot shows the 'Seq Repository' page of the International Wheat Genome Sequencing Consortium (IWGSC). The page features a navigation menu with options like 'Projects', 'Data', 'Tools', 'Seq Repository', and 'About us'. A sidebar on the left lists various data types such as 'Reference sequence', 'Genes & annotations', and 'Physical maps'. The main content area displays a grid of 21 wheat chromosomes, labeled 1A through 7D, arranged in three rows of seven. Each chromosome is represented by a vertical bar with colored segments indicating different genomic regions. A legend at the top of the grid explains that clicking on a chromosome allows users to download, BLAST, or display sequences. The IWGSC logo is visible in the top left corner of the page.

After Quality Control (QC),  
IWGSC WGA will be  
released under Toronto  
agreement through the  
IWGSC sequence repository  
– March/April 2016





# Roadmap to the Wheat Genome Sequence



# Budget for Achieving Gold Standard Reference Genome Sequence by 2017

Category	2-Year Budget
Pseudomolecule Assembly	€550,000
Automated annotation support	€88,800
Data Integration & Display	€100,000
Storage & computing power	€1,000,000
Pseudomolecule Assembly Jamborees & exchange visits for rapid completion	€276,000
Management & Operations	€620,000
<b>Total Costs:</b>	<b>€2,634,800</b>

Addendum (20 January 2016): estimated cost is now less than €1 million

Note: does not include €600,000 for contingencies



# IWGSC Sponsors



MONSANTO



I W G S C

