The wheat genome sequence: a key enabler to boost wheat research

Catherine Feuillet
Head of Trait research
September 13, 2016
Outline

- Wheat strategy @ Bayer
- Why do we need a reference sequence?
- How to get there? The IWGSC
- How do we leverage the genome sequence in our programs?
WHAT THE WORLD EATS: WHEAT

CIMMYT
INTERNATIONAL MAIZE AND WHEAT IMPROVEMENT CENTER

Food Security • Hunger & Malnutrition • Economic Development • Climate Change
The Wheat ID card

**Name**: Triticum aestivum L.

**Surname**: Bread wheat

**Birth place**: Fertile Crescent

**Family**: Grasses (rice, maize, sorghum, sugar cane, millet..)

**Parents**: T. urartu (A), unknown (B), Ae. tauschii (D)

**Address**: More than 160 countries (China, India, USA, Russia France, Canada, Australia, Germany, Pakistan, Turkey, Ukraine, Argentina, UK…)

**Weight**: >600 Million tons/year

**Size**: 16 billion bp

**Features**:
- Globally most important food crop
- Staple food for 30% of world population
- Provides up to 20% of calories and proteins
Future demand projections suggest growing productivity gap in wheat

- Yield growth rate lower than rate of population growth
- Area competition between crops
- At maturity position with “today’s” technology?
- Impacts from climate, resistance and water/nutrient scarcity

**Challenges**

**Way forward**

**Sustainable intensification** through investment in yield and yield stability:

- locally adapted, stress-tolerant, high-yielding varieties
- improved agronomic practices and storage

**Projected demand and wheat yield trends under several scenarios**

Source: CIMMYT 2014
Potential of Wheat Seed/Trait market

---

**Value**
- **Corn**: $17,537
- **Soybean**: $6,211
- **Cotton**: $2,020
- **Canola OSR**: $1,300
- **Wheat**: $2,271

**Total Crop Area**
- **Corn**: 2,020
- **Soybean**: 6,211
- **Cotton**: 2,020
- **Canola OSR**: 1,300
- **Wheat**: 2,271

Source: Sigma Seed Gfk (data 2013)
Bayer CropScience is committed to advancing wheat R&D

“We have mapped out a comprehensive 10-year plan to invest EUR 1.5 billion in the research and development of new solutions for wheat through 2020 – encompassing both seeds and crop protection products.”

- Liam Condon, CEO Bayer CropScience - Sept 2014
Future solutions will work across disciplines: Selected examples for wheat

<table>
<thead>
<tr>
<th></th>
<th>Chemicals</th>
<th>Biologicals</th>
<th>Traits</th>
<th>Seeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Weed Management Systems for grass &amp; broad-leaved weeds</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>2. Disease management which meets increasing regulatory demands, provides excellent disease control, &amp; yield benefits beyond disease control</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Deliver Yield gains</td>
<td></td>
<td></td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Wheat yield frontiers

Based on Bio-Physical limits

Potential Yield (Plot Trials)

Top Quartile Farmers

New technology gap

Expanding the yield frontier

Exploiting today’s yield potential

Technological possibility
Current technology potential
Best Farmer practice
Average yield

Systems gap

Bio-Physical limits

Potential Yield (Plot Trials)

Top Quartile Farmers

Best practice

Systems gap

Expanding the yield frontier

Exploiting today’s yield potential

Page 9 – Plant Genomics Congress USA- Sept 2016

Bayer CropScience
Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping

Crop Efficiency Research Strategy: focus on yield components

- Product concept
- Phenotypic trait
- Improved plants
- Biological process
- Tools to enable idea
- Idea to improve crop
- Crop & Breeding Expertise
- Biological expertise
- Native
- GM
- Non-GM
- Data mining
- Omics
- Phenotyping
Gene discovery in wheat...10 years ago

- Molecular markers
- ESTs and microarrays
- BAC libraries
- Genetic resources
- Gene validation
- Grass comparative genomics

~ 10 years for map based cloning single trait genes...
Gene discovery in wheat.....10 years ago
THE Challenge....

✓ Allohexaploid
✓ Large: 16 Gb

✓ 80-90% of TEs and repeats

✓ > 50% of non recombinogenic regions
THE Challenge....

✓ Allohexaploid
✓ Large: 16 Gb

✓ 80-90% of TEs and repeats

✓ > 50% of non recombinogenic regions
International Wheat Genome Sequencing Consortium

- 2005: **Launch** (Kansas Wheat)
- 2006: 1st **Strategic road map**
  - Involve industry
  - Hexaploid wheat
  - Chromosome-based approach (BAC by BAC)
- **Country based funded projects**
- **Key achievements**
  - 2008: 1st physical map of chr. 3B (1Gb)
  - 2012: first reference sequence chr. 3B
  - 2014: 21 chromosomes survey sequences
- 2015: NRGene de NovoMAGIC assembly from **whole genome sequence**
A chromosome-based draft sequence of the hexaploid bread wheat genome
The International Wheat Genome Sequencing Consortium

- 10.2 Gb assembled sequences
- 128Mb (1DS) – 639Mb (3B) assembled sequence per chromosome
- N50 contig* = 5.9 kb (1.7kb-8.9kb)
- 99,386 annotated genes assigned to chromosomes for the first time
- 50% gene ordered- In silico mapping
- Homoeologous gene expression studies
- Ancestral genome evolution studies

*After repeat masking

Download & BLAST search available at the IWGSC repository:
http://wheat-urgi.versailles.inra.fr/
Progress towards completion of a bread Wheat reference genome sequence: Status 2016

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

Completion expected for 2018-19
De novo assembly:

- **NRGene's DeNovoMagic-2** platform, total run time < 3 weeks, 1Tb RAM computer
- **Illumina short-read** sequence data only (200x coverage, paired ends + mate pairs)
- Sequence contigs / scaffolds assigned to chromosomes using **IWGSC CSS + POPSeq data**

<table>
<thead>
<tr>
<th>Assembly size:</th>
<th>14.5 Gbp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaps size:</td>
<td>262 Mbp (1.8%)</td>
</tr>
<tr>
<td>N50 (scaffolds):</td>
<td>7.1Mbp</td>
</tr>
<tr>
<td>Coverage in scaffolds: &gt; 100kb</td>
<td>14.2 Gb (4,442)</td>
</tr>
<tr>
<td>Scaffolds ordered by HiC map</td>
<td>3,975 (14.1 Gb)</td>
</tr>
</tbody>
</table>

- Assembly contains **95 to 99%** of the genes and TE based markers
- **Chromosome data** (physical maps, WGP tags, Bionano maps, HiC) used to correct chimeras, orient scaffolds and generate super-scaffolds
- Super-scaffolding extends linkage by ~ 3-fold - **N50 superscaffolds 23.8Mbp**
Roadmap to the Wheat Genome
Gold standard Reference Sequence

Illumina sequencing of individual chromosomes

- Chromosome Survey Sequence v2 (2014)
  - Whole genome mate pairs
    - IWGSC CSS v3 (2016)

Physical maps of individual chromosomes

- MTP sequencing (86%)
  - Pseudomolecule assembly (42%)
    - NRGene-Illumina WGS
      - IWGSC Whole Genome Assembly (2016)

BioNanoGenomics, optical, RH, HiC maps
- PopSeq, LD maps
- MTP sequence tags

Reference genome sequence (2017)
Wheat Research gets a boost! But ...

**Agronomy**
- Agricultural practices and policies

**Phenotyping**
- HT in controlled conditions
- HT in the fields
- Imaging technologies

**Genomics resources**
- Physical maps
- Genome sequences
- HT DNA markers platforms
- HT gene expression platforms
- All “omics” and bioinformatics

**Methodologies**
- Modeling/prediction
- Genome engineering
- New breeding techniques

**Genetic resources**
- Mapping populations (RILs, NAM, MAGIC, RH….)
- Association panels
- Genebanks
- Mutant populations

---

Page 20 – Plant Genomics Congress USA- Sept 2016
Success will depend on our ability to translate data into knowledge and use technologies.
Computational biology is key to ensure integration and translate data into knowledge.

Field Trials data

Farming & sensor data

- Omics data

Gene expression data

metabolic data

Computational Life Science

PRODUCE DATA

STORE DATA

ANALYSE DATA

INTEGRATE UNDERSTAND MODEL

EXPERIMENTAL DESIGN
Translating knowledge into products for a next generation of (complex) traits

Herbicide, Insect, Disease resistance...
Oil composition, Pod shatter reduction...
Crop Efficiency, water use efficiency...
Bayer CropScience allocates grants for the exploration of attractive, novel solutions to increase crop productivity.

Grants4Traits™

Novel solutions to increase crop productivity

At Bayer, we have a successful history of collaborating with external parties to deliver safe and innovative seed products to markets. We are interested in learning from you and partnering with you to sustain this innovation.

Apply now for a grant and develop further some of your ideas towards products that will increase sustainable crop output to meet the demands of an ever growing population and changing environment.

Next submission deadline is October 31st 2016

Submit Proposal

https://innovate.bayer.com/
Thank you
Forward-Looking Statements

This presentation may contain forward-looking statements based on current assumptions and forecasts made by Bayer Group or subgroup management.

Various known and unknown risks, uncertainties and other factors could lead to material differences between the actual future results, financial situation, development or performance of the company and the estimates given here. These factors include those discussed in Bayer’s public reports which are available on the Bayer website at www.bayer.com.

The company assumes no liability whatsoever to update these forward-looking statements or to conform them to future events or developments.