Unlocking the polyploid potential of wheat through genomics

Cristobal Uauy (cristobal.uauy@jic.ac.uk)
@CristobalUauy
On average, we eat 50 wheat plants every day.

https://www.jic.ac.uk/miracles/
Wheat is a young polyploid.

- **Triticum urartu** (A^U^A^U^) 5-7 MYA
- **T. turgidum** ssp. *dicoccoides* (AABB) 400,000 years
  - Wild emmer
  - *eg. Ae. speltoides* (“BB”) 10,000 years
- **T. turgidum** ssp. *durum* (AABB)
- **Ae. tauschii** (DD)
- **T. aestivum** (AABBDD) 400,000 years

Wheat is a young polyploid.
• Three copies of every gene
• ~97% sequence similarity
• Assumed similar expression patterns
• Buffering effects of the genome
In the past 36 months...

IWGSC 2014
Clavijo et al 2017
Zimin et al 2017

IWGSC 2017
Earlham (varieties)

HapMap: Jordan et al. 2015 Genome Biology

Gardiner et al. 2015 Genome Biology
Gardiner et al. 2017 bioRxiv

in silico TILLING: Krasileva et al 2017

Pearce et al 2015 BMC Plant Biol
Borrill et al 2016 Plant Physiol

Hassani-Pak 2017
Ramirez-Gonzalez et al under review

Wilkinson et al 2012/2016 BMC Bioinformatics

CerealsDB
polymarker

Ramirez-Gonzalez et al 2015 Bioinformatics
Home page

This website provides background information and practical resources to help both budding wheat scientists as well as researchers looking to expand their work into wheat.

As we move into the future there is a great need to improve crops to feed the world's growing population and to cope with climate change. One potential route to deal with this challenge is to translate fundamental plant biology research (e.g. from Arabidopsis thaliana) into crops such as wheat. However, learning even simple tasks such as growing and crossing wheat plants requires time and effort, while material and methods sections in published articles are often short and cannot substitute teaching aids. This is even more true for more complex topics such as the genomics aspect of wheat.

Here we provide information and training about:

- Growing wheat
- Wheat genomic resources
- TILLING mutant resources
- Useful wheat links
Talk outline

Homoeolog expression patterns

Grain size and yield

Polyploidy (!!!)

3 brakes  2 brakes  0 brakes
-5%    +5%    +??%
Over 850 samples; developmental time course, etc

Ramirez-Gonzalez, Borrill et al under review
Wheat eFP

http://bar.utoronto.ca/~asher/efp_wheat/cgi-bin/efpWeb.cgi

Wheat (Triticum aestivum L.) eFP Boswer at bar.utoronto.ca
RNA-seq data from Azhurnaya spring wheat

Nicholas Provart

Andv Sharpe

Yogendra Khedikar

Mark Davey
53,259 genes are present as 1:1:1 homoeologs.

<table>
<thead>
<tr>
<th>homoeologs</th>
<th>NAC-A1</th>
<th>NAC-B1</th>
<th>NAC-D1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPM (absolute)</td>
<td>10</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>TPM (relative)</td>
<td>0.33</td>
<td>0.27</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Balanced
Kinase-A1

Kinase-B1

Kinase-D1

homoeologs

TPM (absolute)

TPM (relative)

A genome

B genome

D genome

2.5

0

3

0.45

0.00

0.55

B suppressed
Defined seven expression categories (53,259 genes in triads)
How do we visualise?
Homoeolog expression patterns

- 21% A
- 7% B
- 73% D
- 30% A
- 30% B
- 40% D
30% of triads show non-balanced expression
Variation across tissues
Balanced triads are more highly expressed (absolute)
House-keeping (balanced) vs tissue-specific (dominant)
Polyploidisation and polyploidy affect these patterns

- Expected *in-silico*
  AS2255(BBAA) × AS60(DD)

- Observed SHW1 (BBAADD)
  AS2255 × AS60

- Observed Chinese Spring (BBAADD)
  Balanced
  A dominant
  A suppressed
  B dominant
  B suppressed
  D dominant
  D suppressed
Putative mechanism for homoeolog expression bias

![Diagram showing the expression bias of homoeologs under different conditions. The diagram includes graphs and boxes illustrating the relative contribution of genes in different states: Balanced, A suppressed, B suppressed, D suppressed, A dominant, B dominant, D dominant. The states are color-coded: Balanced = gray, Dominant = dark green, Non-Dominant = light blue, Suppressed = yellow, Non-Suppressed = brown.](image-url)
No effect from repeat elements in promoter (5 kb upstream)
Significant differences in histone methylation/acetylation

Etienne Paux
Lorenzo Concia et al.
Variation in homoeolog expression patterns across tissues
Variation in homoeolog expression patterns across tissues
Tissue-Stable and dynamic triads

![Graph showing the distribution of triads with different average distances between mean and tissue (%). The graph includes three categories: Stable 10%, Middle 80%, and Dynamic 10%.]
Stable triads are housekeeping genes

Cell cycle
Translation

Defence/external stimuli response
Secondary metabolite processes
Dynamic triads belong to all categories

94% Stable

31% Dynamic

94% Balanced
31% A suppressed
38% B suppressed
31% D suppressed
31% A dominant
31% B dominant
31% D dominant
Dynamic triads are under more relaxed selection pressure
Dynamic triads are under more relaxed selection pressure.
Dynamic triads: shorter, more frequent, and closer TEs

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of Genes</th>
<th>Distance to nearest TE (bp)</th>
<th>1.5 kb</th>
<th>Genes with $\geq 1$ TE (%)</th>
<th>Triads with $\geq 1$ TE (%)</th>
<th>Median Size of TE (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic 10%</td>
<td>4275</td>
<td>1113</td>
<td></td>
<td>58.5%</td>
<td>88.3%</td>
<td>220</td>
</tr>
<tr>
<td>Middle 80%</td>
<td>34221</td>
<td>1242</td>
<td></td>
<td>53.7%</td>
<td>81.1%</td>
<td>230</td>
</tr>
<tr>
<td>Stable 10%</td>
<td>4278</td>
<td>1234</td>
<td></td>
<td>54.9%</td>
<td>79.2%</td>
<td>259</td>
</tr>
</tbody>
</table>

*P value* 3.0E-06 1.1E-08 < 2.2E-16 1.5E-10
Expression asymmetry across chromosomes

Recombination
- High
- Infrequent

Expression Breadth
- Tissue Specific
- Ubiquitous

GO Enrichment
- Defence
- Housekeeping
Expression asymmetry across chromosomes

Recombination
- High
- Infrequent

Expression Breadth
- Tissue Specific
- Ubiquitous

GO Enrichment
- Defence
- Housekeeping

Homoeolog Expression Bias

Within Tissue
- Suppressed / Dominant
- Balanced

Across Tissue
- Dynamic
- Stable

Inter-Cultivar
- Variable
- Conserved
Talk outline

Homoeolog expression patterns

Grain size and yield

Polyploidy (!!!)
Yield and quality components

- Reproductive stage:
  - number of spikes (tillering)
  - grain number (spike architecture)

- Grains filling stage:
  - grain weight

Final grain yield
From QTL to NILs to field…

Doubled Haploid population
Charger x Badger

Number of SNPs
0
25
50+

Inner tracks
SNP variation (iSelect 90k chip)
1. Parental SNPs
2. BC₄ NILs SNPs

Outer track
Combined LOD score
(12 environments)

Simmonds et al 2014 BMC Plant Biology
Brinton et al. 2017 New Phytologist
5A+ NILs have increased TGW

Brinton et al. 2017 New Phytologist
Natural variation for grain size

Carpel / early grain development

Grain length (mm)

Average vs average

95.6 µm
107.3 µm

8.33%
P = 0.049
Grain length (mm)

Big vs average

93.1 µm

107.3 µm

9.53%
P = 0.015
Two mechanisms underlie QTL

GW-A

GL-A

Carpel / early grain development

- Cell expansion
- Cell division

Can we combine different grain size genes?

Grain length (5A QTL)
Brinton et al, *New Phytologist* 2017

Grain width (6A QTL)
Simmonds et al, *BMC Plant Biol* 2014

Cell expansion

Cell division

Cell expansion

Cell division
Combination: higher TGW, higher yield (stability?)

TGW

- Morley
- CF

GL-A
GW-A
Talk outline

Homoeolog expression patterns

Grain size and yield

Polyploidy (!!!)
**in silico TILLING**

EMS

\[ \text{Genomic DNA} \quad \text{Exome capture} \quad \text{Mutation identification} \]

\[ \text{M}_2 \text{ population} \]

Krasileva et al 2017 *PNAS*

Uauy et al 2009 *BMC Plant Biology*
Welcome to the *in silico* wheat Target Induced Local Lesions In Genome (TILLING) website

This resource consists of TILLING populations developed in tetraploid durum wheat cv 'Kronos' and hexaploid bread wheat cv 'Cadenza' as part of a joint project between the University of California Davis, Rothamsted Research, The Earlham Institute, and John Innes Centre.

We have re-sequenced the exome of 1,535 Kronos and 1,200 Cadenza mutants using Illumina next-generation sequencing, aligned this raw data to the IWGSC Chinese Spring chromosome arm survey sequence, identified mutations, and predicted their effects based on the protein annotation available at Ensembl Plants.

Search TILLING data

BLAST Scaffold

Detailed info on use at [www.wheat-training.com](http://www.wheat-training.com)

Krasileva et al 2017 *PNAS*
Mutants in TGAC gene models in *EnsemblPlants*
Combining mutations

• Simple scheme

Ramirez-Gonzalez et al. 2015 *Bioinformatics*
Speed Breeding collaboration with Lee Hickey/Amy Watson (Queensland) and Brande Wulff/Sreya Ghosh (JIC)

Cadenza 38 days after sowing

Accelerated growth

Normal growth

Watson, Ghosh et al. 2018 Nature Plants
Speed Breeding collaboration with Lee Hickey/Amy Watson (Queensland) and Brande Wulff/Sreya Ghosh (JIC)

Watson, Ghosh et al. 2018 Nature Plants
Fine mapping of the 5A QTL for grain length

Brinton et al. 2017 New Phytologist
Haplotype-based analysis of QTL interval

Related to Cadenza, Robigus or Claire

Bernardo Clavijo

RNA-Seq: Brinton et al. 2018 BMC Plant Biology
Double mutant increases TGW by ~20%

Candidate based on genome sequence of cultivars
Natural variation for grain size

+GW-A  -gw-A
Natural variation for grain size

+GW-A

-gw-A

+GW-A in green behind
Polyploidy, redundancy and the multi-brake hypothesis
Polyploidy, redundancy and the multi-brake hypothesis

- Few large effect QTL
  - 30-40% phenotypic difference
- Many subtle effect QTL
  - 5-10% phenotypic difference

Polyploidy

Genetic Architecture
Diploid rice: single gene effect

Song et al 2007 Nature Genetics
Polyploidy, redundancy and the multi-brake hypothesis

3 brakes

2 brakes +5%

0 brakes +??%

A genome

B genome

D genome

GW-A

GW-B

GW-D

Borrill, Adamski & Uauy 2015
New Phytologist

Simmonds et al 2016 TAG

50%

TGW (% increase)
Triple GW2 mutant increases TGW by >20%

TGW (%)

Wildtype - -
single + 6 %
double + 11 %
triple + 21 %
Dosage of quantitative trait

Dominance

Quantitative phenotypic variation (vs wild type)

# Mutant alleles

Gene Examples

PPD, VRN1, AP2

Full mutant effect

50%

Wild-type
Dosage of quantitative trait

Quantitative phenotypic variation (vs wild type)

# Mutant alleles

- 1
- 2
- 3

Gene Examples
- PPD, VRN1, AP2
- GPC, GW2, 5A, R
- Waxy

Dominance
- Full mutant effect
- 50%
- Wild-type

Dosage
- Additive
- Non-additive
Dosage of quantitative trait

**Quantitative phenotypic variation (vs wild type)**

- **Dominance**
  - Full mutant effect
  - 50%
  - Wild-type

- **Dosage**
  - Additive
    - # Mutant alleles: 1, 2, 3
    - Gene Examples: PPD, VRN1, AP2
  - Non-additive
    - # Mutant alleles: 1, 2, 3
    - Gene Examples: GPC, GW2, 5A, R

- **Functional Redundancy**
  - # Mutant alleles: 1, 2, 3
  - Gene Examples: SBE-II, MLO
Dosage of quantitative trait

Quantitative phenotypic variation (vs wild type)

Dominance

Dosage

Non-additive

Functional Redundancy

Gene Examples

QTL

Selected Variation

Hidden Variation

PPD, VRN1, AP2

GPC, GW2, 5A, R

Waxy

SBE-II, MLO
Dosage of quantitative trait

Quantitative phenotypic variation (vs wild type)

Dominance

Additive

Non-additive

Functional Redundancy

Full mutant effect

50%

Wild-type

# Mutant alleles

Gene Examples

PPD, VRN1, AP2

GPC, GW2, 5A, R

Waxy

SBE-II, MLO

Selected Variation

Hidden Variation

Wildtype

single

double

triple

TGW (%)
Concluding thoughts

Expression bias and its significance to phenotype

Mechanistic understanding of grain size (yield?)

Hidden variation in polyploids
Wheat (and human) diversity

Jemima Brinton
James Simmonds
Philippa Borrill
Ricardo Ramirez-Gonzalez
Sophie Harrington
Nikolai Adamski

@CristobalUauy

DESIGNING FUTURE WHEAT

International Wheat Yield Partnership
Research to Deliver Wheat for the Future