

INvestigating TRiticeae EPIgenomes for Domestication (INTREPID)



Anthony Hall



Neil Hall



John Kenny



Mike Bevan



Klaus Mayer

HelmholtzZentrum münchen
German Research Center for Environmental Health



Rob Martienssen



Richard McCombie

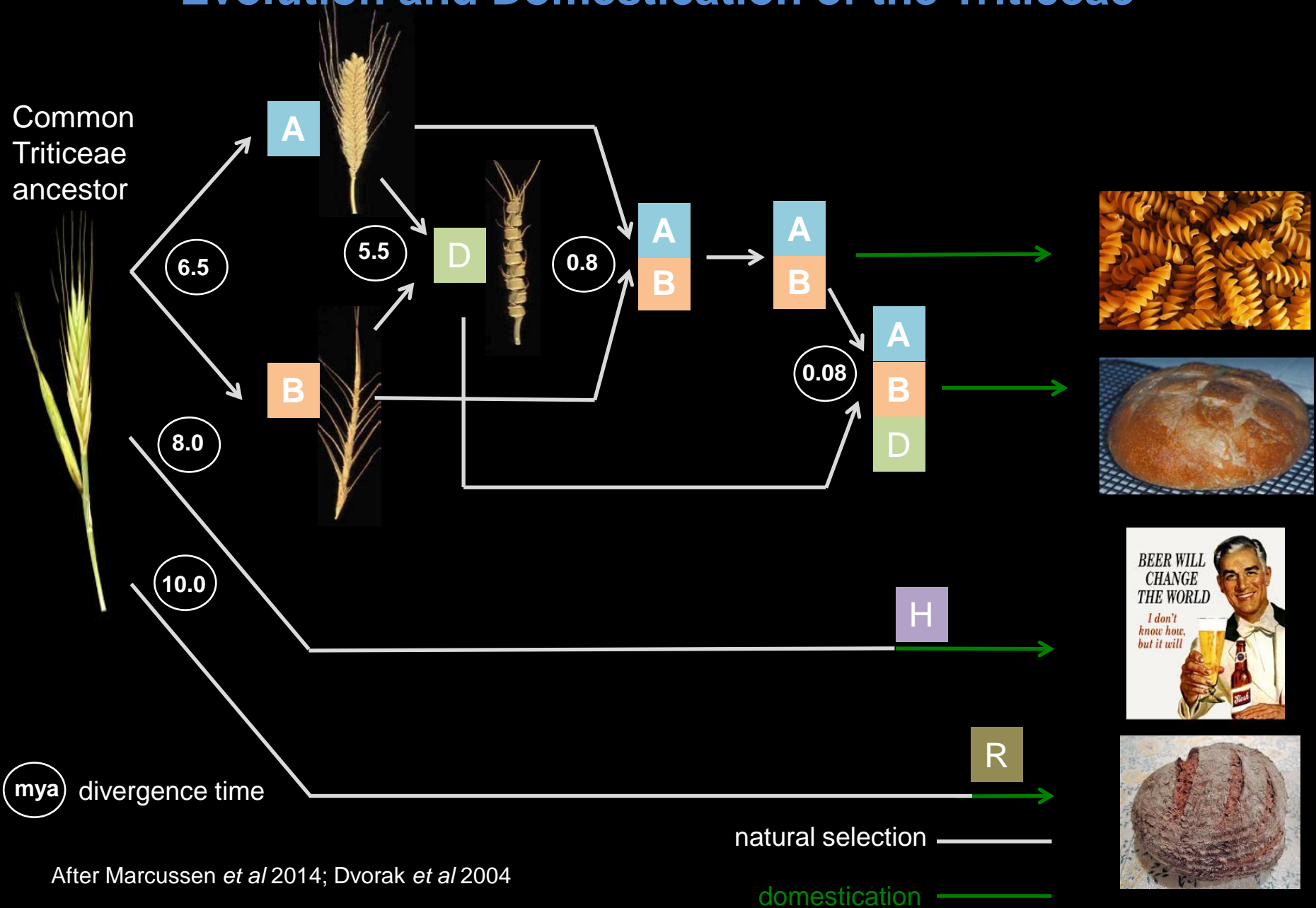


The crew of the INTREPID

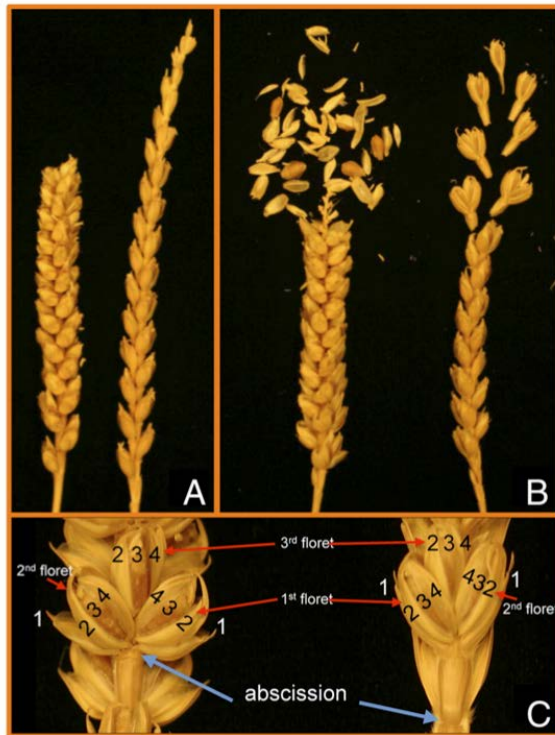
AIMS

- Measure epigenetic modifications to genes and relate this to gene expression
- patterns in hexaploid wheat and tetraploid and diploid progenitor species
- Survey the epigenomes of elite wheat varieties to map the influence of breeding
- on the epigenome
- Understand how hybridisation influences epigenetic modifications and how they
- are stabilised

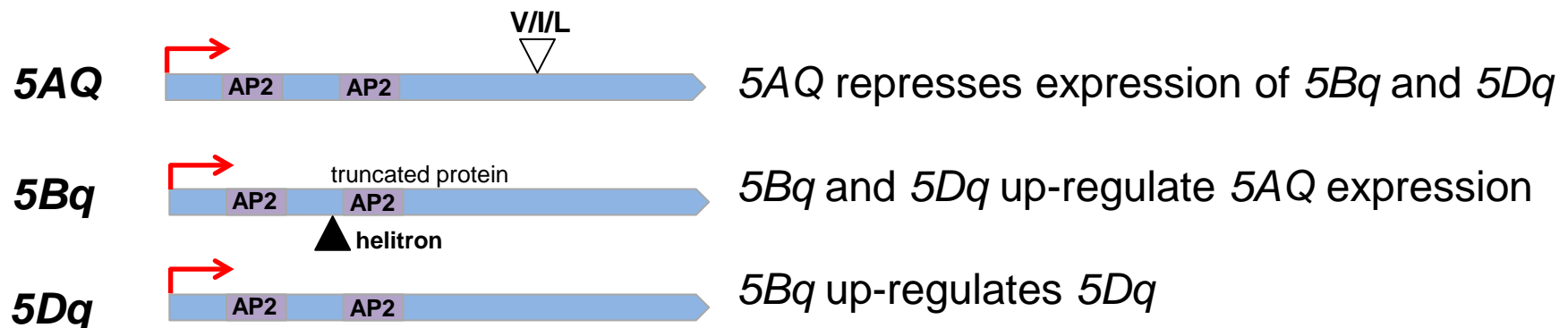
Evolution and Domestication of the Triticeae



New traits arise directly upon formation of new hexaploid wheat

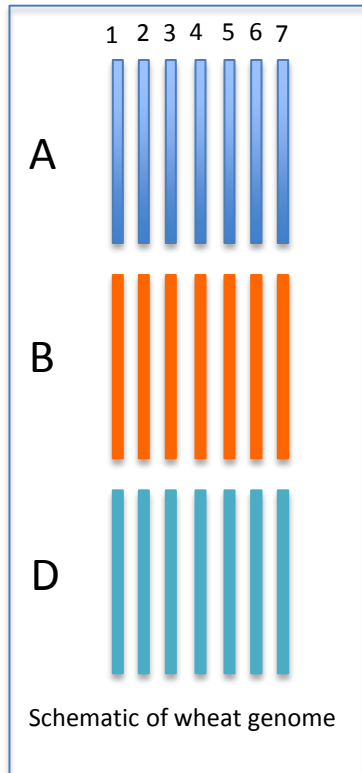


Zhang Z et al. PNAS 2011;108:18737-18742



The wheat genome

- The size of the wheat genome is 17GB
- Hexaploid bread wheat has 3 closely related and independently maintained genomes
- Mostly repeats (transposons, 89%)
 - Fragmented assemblies
 - Very limited long-range scaffolding



ARTICLE

doi:10.1038/nature11690

Analysis of the bread wheat genome using whole-genome shotgun sequencing

Rachel Brenchley¹, Manuel Spannagl², Matthias Pfeifer², Gary L. A. Barker³, Rosalinda D'Amore¹, Alexandra M. Allen⁴, Neil McKenzie⁵, Melissa Kramer⁶, Arnaud Kerhornou⁷, Dan Bolser⁸, Suzanne Kay⁹, Darren Waite¹, Martin Trick¹, Ian Bancroft⁴, Yong Gu¹, Nixin Huo¹, Ming Cheng Luo¹, Sunish Sehgal¹⁰, Bikram Gill¹⁰, Sharyar Kianian¹⁰, Olin Anderson¹¹, Paul Kersey⁶, Jan Dvorak⁶, W. Richard McCombie¹, Anthony Hall¹, Klaus F. X. Mayer², Keith J. Edwards¹, Michael W. Bevan¹ & Neil Hall¹

Brenchley *et al.* 2012

WHEAT GENOME

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)*†

IWGC 2014

of the A, B, and D subgenomes. Genetically, however, it behaves as a diploid because homeologous pairing is prevented through the action of *Ph* genes (8). Each of the subgenomes is large, about 5.5 Gb in size and carries, in addition to related sets of genes, a high proportion (>80%) of highly repetitive transposable elements (TEs) (9, 10).

The large and repetitive nature of the genome has hindered the generation of a reference genome sequence for bread wheat. Early work focused primarily on coding sequences that rep-

Chapman *et al.* Genome Biology (2015) 16:26
DOI 10.1186/s13059-015-0502-8

Genome Biology

METHOD

Open Access

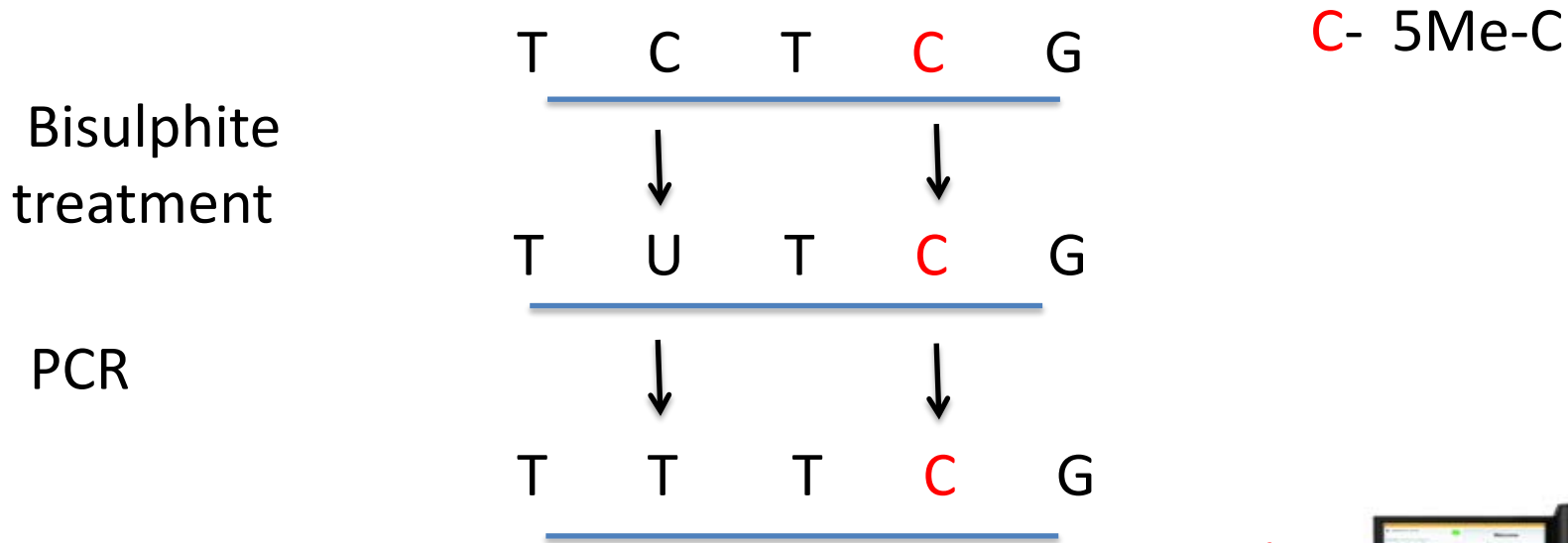
A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome

Jamod A Chapman^{1,2}, Moritz Mascher^{2,3}, Aydin Buluç², Kerrie Bory², Evangelos Georgantas^{2,4}, Adam Seaton⁵, Veronika Smadova², Jerry Jenkins^{1,2}, Sunish Sehgal^{1,2}, Leonid Olsh⁶, Jeremy Schmutz⁷, Katherine A Yelick^{1,4}, Uwe Scholz², Robbie Wough⁸, Jesse A Poland⁹, Gary J Muehlbauer¹⁰, Nils Stein¹¹ and Daniel S Rokhsar¹²

Next Release TGAC October 2015

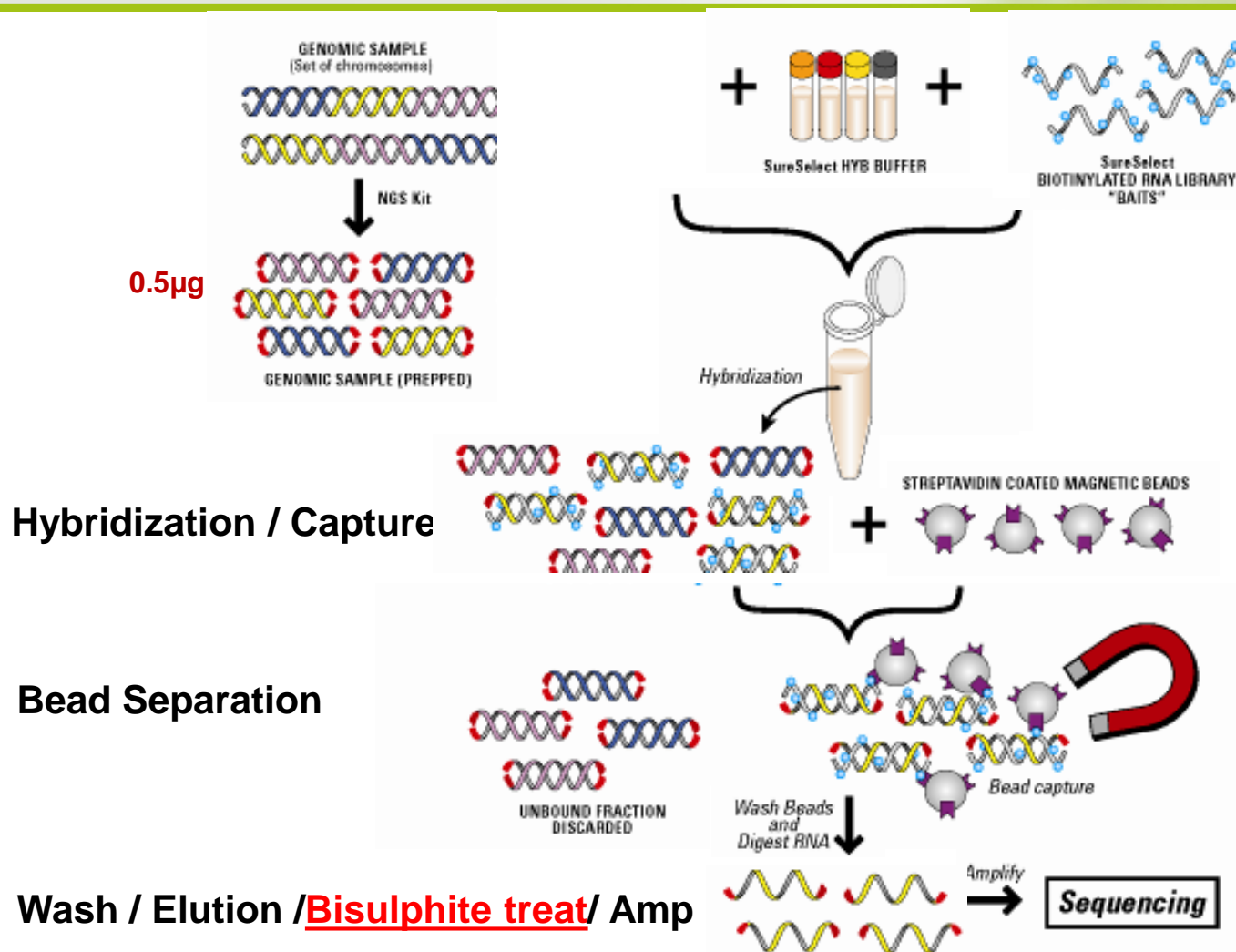
Chapman *et al.* 2015

Sequencing epigenomes



Map C/T with BISmark

TECHNOLOGIES FOR INVESTIGATING THE WHEAT EPIGENOME



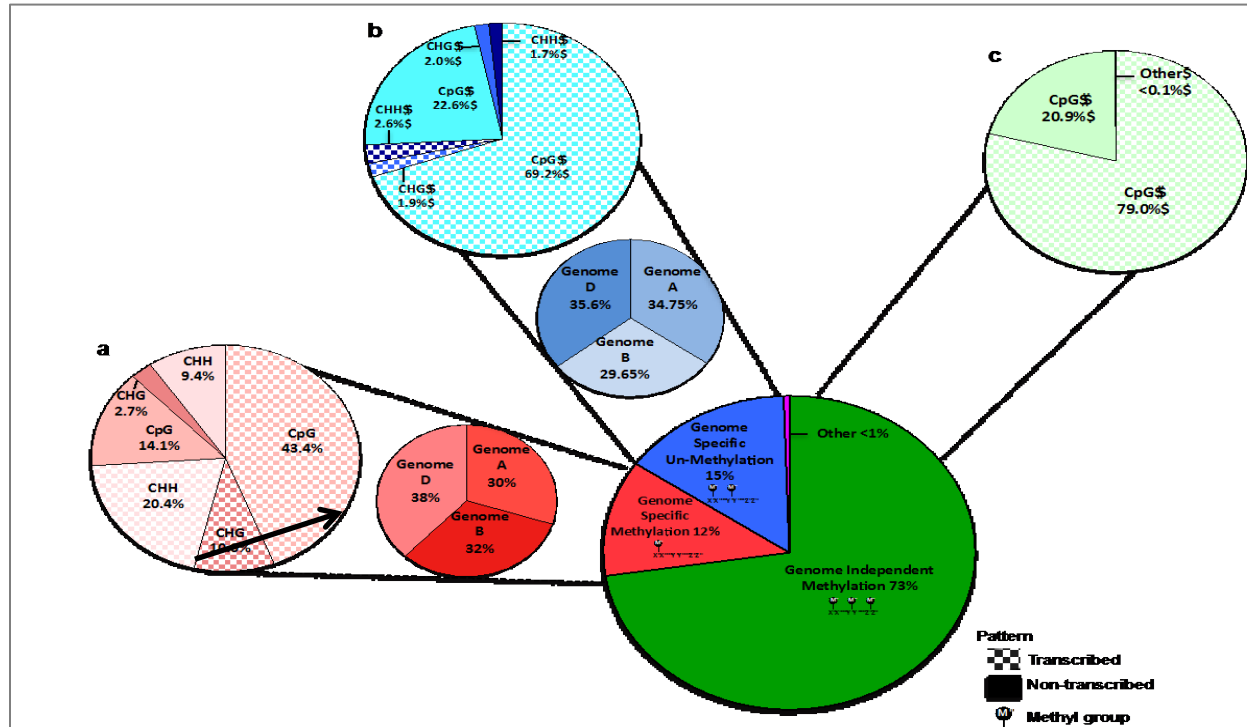
Baits:

- cRNA probes
- Long (120bp)
- Biotin labeled
- Designed to tile across the gene space of wheat

Sequences maps against 81Mb genic wheat assemblies
Off target sequence includes chloroplast and transposons

SureSelect Target Enrichment System: Workflow

PRELIMINARY ANALYSES



❖ 73% of methylation genome independent
 >99% CpG sites

❖ 27% methylation genome specific
 57% CpG, 30% CHH and 13% CHG sites

❖ D genome preferentially methylated

A & B genome GO term enrichment; signaling, metabolic & biosynthetic processes

D genome GO term enrichment; chromatin silencing, histone modification, methylation/regulation of gene expression, chromosome organization

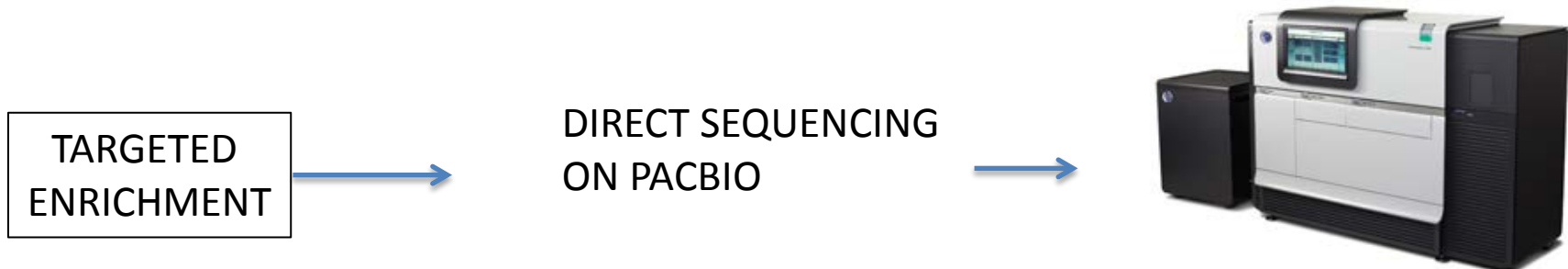
DESIGN A NEW PROBE SET FOR ANALYSING THE WHEAT EPIGENOME

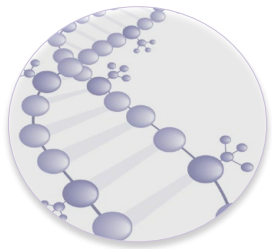
- Expand the design based on the latest IWGC assemblies
- Design a large probe set- 24Mb, capture ~120Mb, 360Mb
- Equates to the whole wheat genic space
- Transposons targeted and off target
- Off-target chloroplast DNA used as non-methylated standard

Main project pipeline Illumina sequencing



Modified bases slow down DNA polymerase





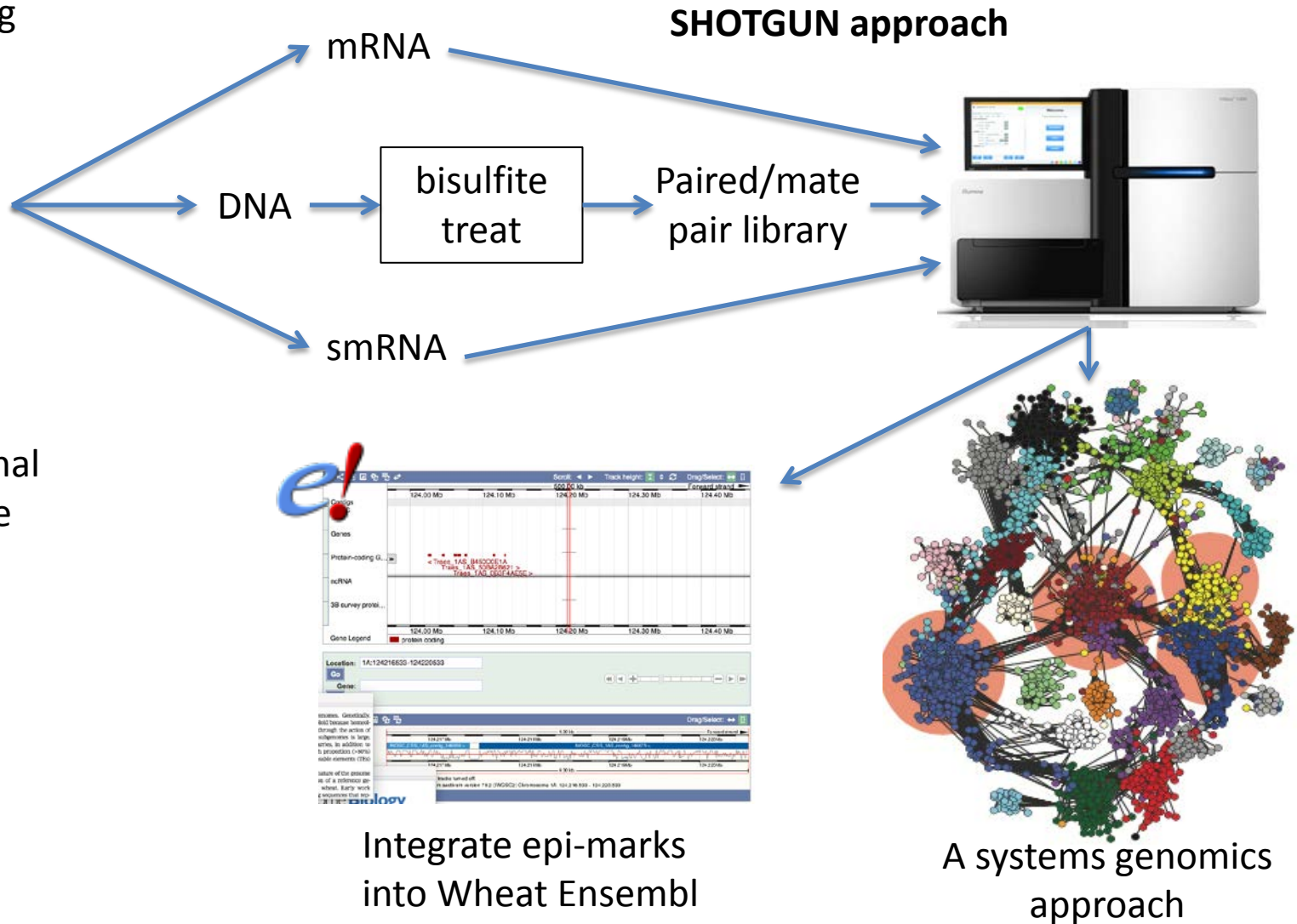
DEFINE THE WHEAT EPIGENOME FOR CHINESE SPRING BIN 42

The Project

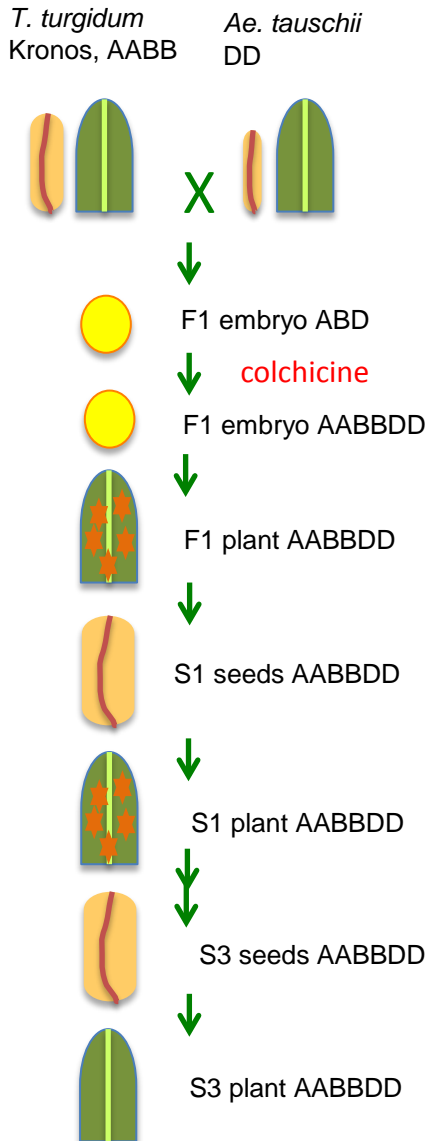
Single 7D seedling



CS42 International reference



Epigenome changes in newly formed hybrids



Hybrid formation and epigenome changes.

- New traits are formed directly as a result of new hybrid formation
- There is a period of genome instability for several generations shown by necrosis and death
- Test the hypothesis that methylation patterns are changed in new hybrids, and that methylation patterns are stabilised leading to genome stability

Experimental Plan

- Generate several independent hybrid wheat lines from different AABB and DD progenitors
- F1 hybrids (ABD), stabilised doubled hybrids (AABBDD) and S1, S2 and S3 selfed progeny will be sampled for epigenome, transcriptome, and sRNA levels
- Genes and repeats will be sequenced
- The type, extent and stability of epigenetic changes, and their consequences on gene family composition and expression.

Data Sharing Plan

- Established in 2008 by the NSF and a \$50M grant, renewed in 2014. UK iPLANT node established.
- **INTREPID** will use iPLANT to share data and pipelines across sites
- Upon publication all data, bioinformatics pipelines and documentation will be made available to the public *via* iPLANT



Data Store

Store, manage, access, and share all the data related to your research



Environment

Use hundreds of bioinformatics Apps and manage data in a simple web interface



Atmosphere

Create a custom cloud-based scientific analysis platform or use a ready-made one for your area of scientific interest