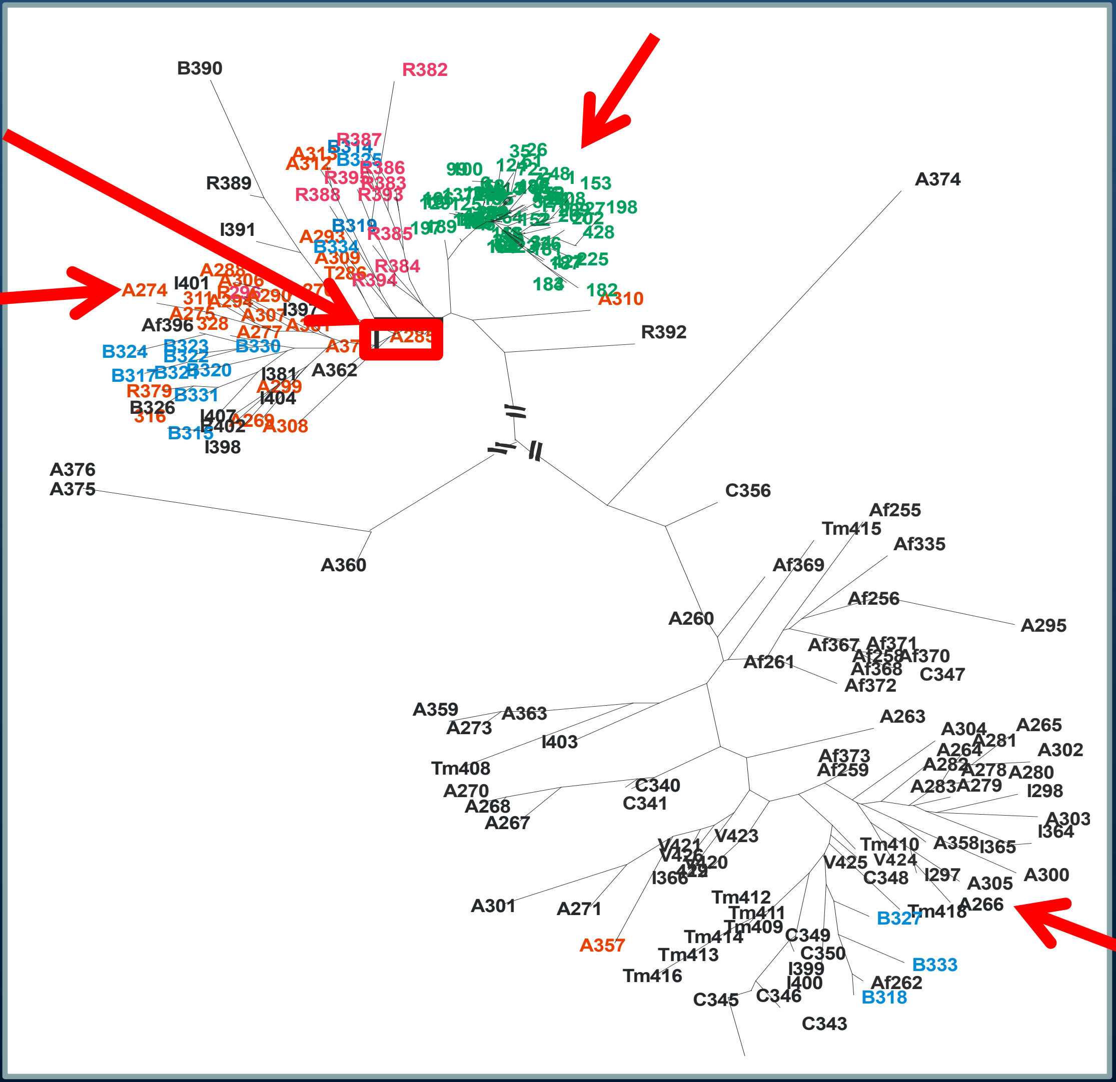
Sequencing of the *Aegilops tauschii* genome

IOS-1238231

subspecies *strangulata* accession AL8/78

Genetic relationships of accession AL8/78



Hexaploid wheat

AL8/78

Ae. tauschii
ssp. *strangulata*

Ae. tauschii
ssp. *tauschii*

Collection site of AL8/78 and the putative geographic area of hexaploid wheat origin

AL8/78

Hexaploid wheat origin

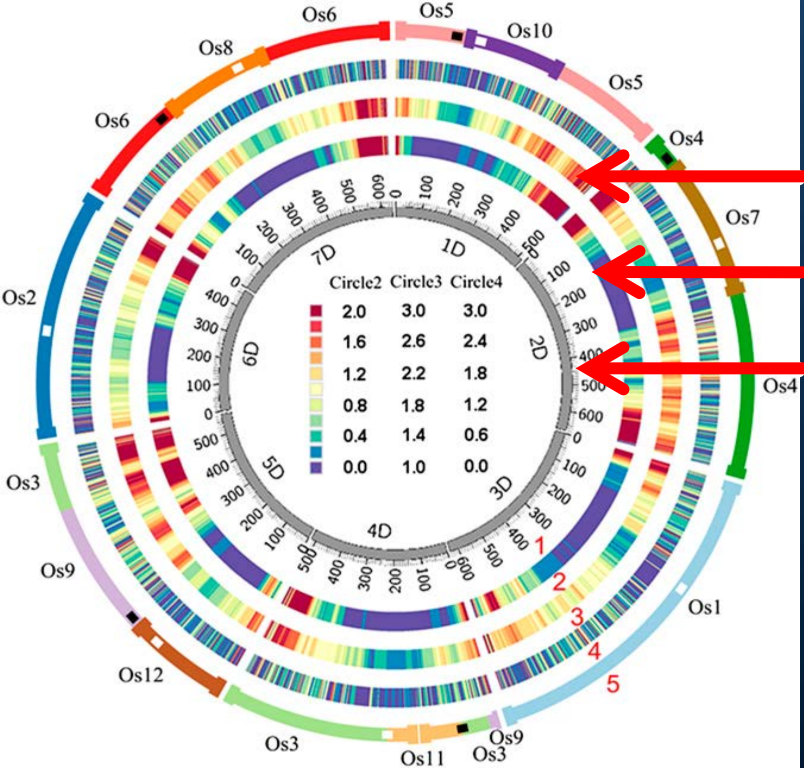
Aegilops tauschii genome

- Haploid genome ~ 4 to 5 Gbp
- ~ 80 to 90% repeated sequences

Ordered-clone sequencing approach

Aegilops tauschii physical map

- 3,578 BAC contigs
- 2,263 anchored contigs (84.2% total contig length)
- MTP = 42,882 BAC clones



3 = gene density

2 = recombination rates

1 = physical map

Luo *et al*., PNAS 110, 2013

Division of labor among participants

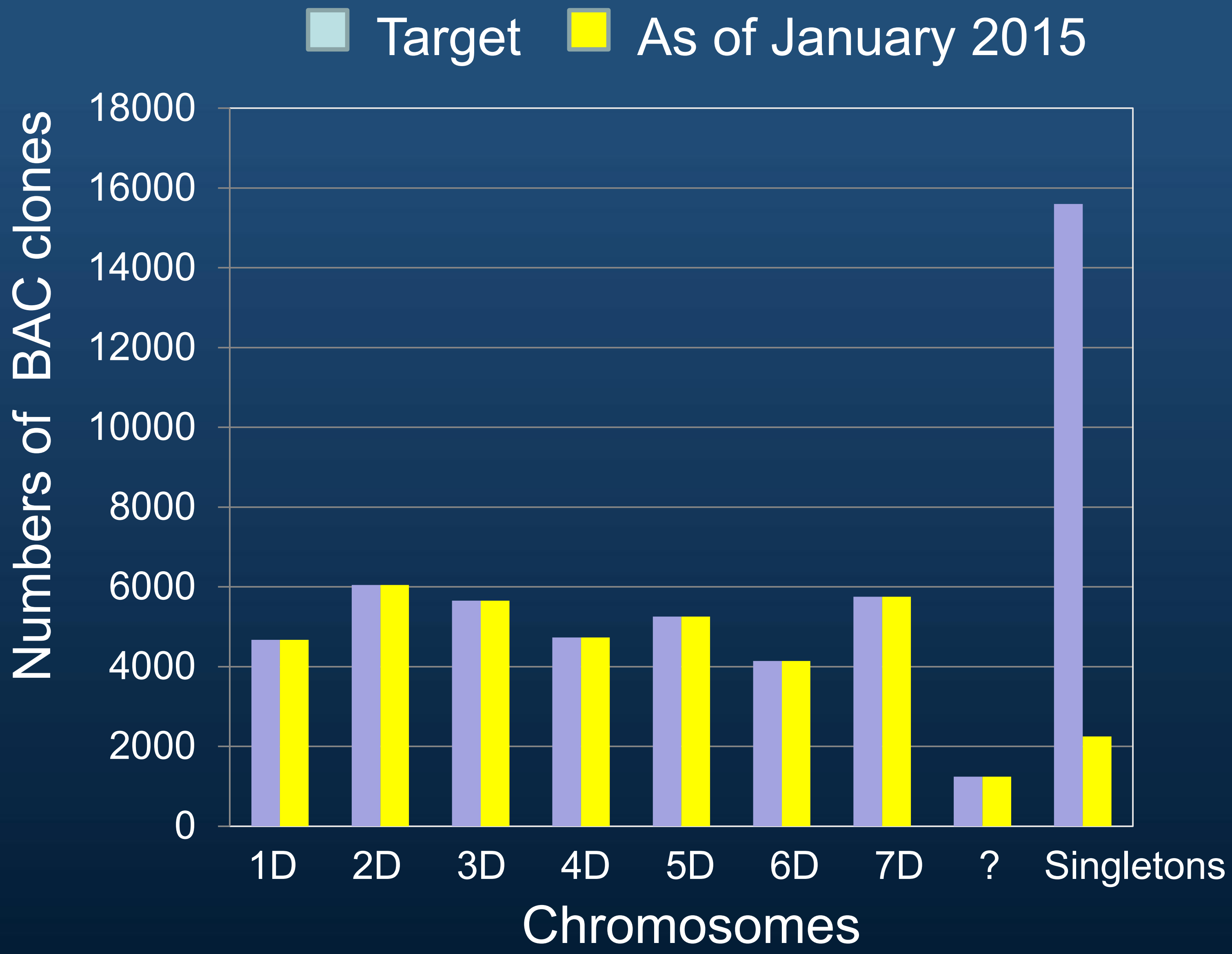
By chromosome

By task

Tasks

- Validate and pool 8 BAC clones
- Index each pool, combine 48 pools, sequence pools
- Assemble our short pair-end reads together with BGI long pair-end reads
- Preliminarily annotate TEs and genes
- Construct nanomap and validate scaffolds
- Produce pseudomolecules
- Finally annotate TEs and genes
- Conduct community gene annotation

MiSeq MTP sequencing and assembly status



Scaffold assembly v.1

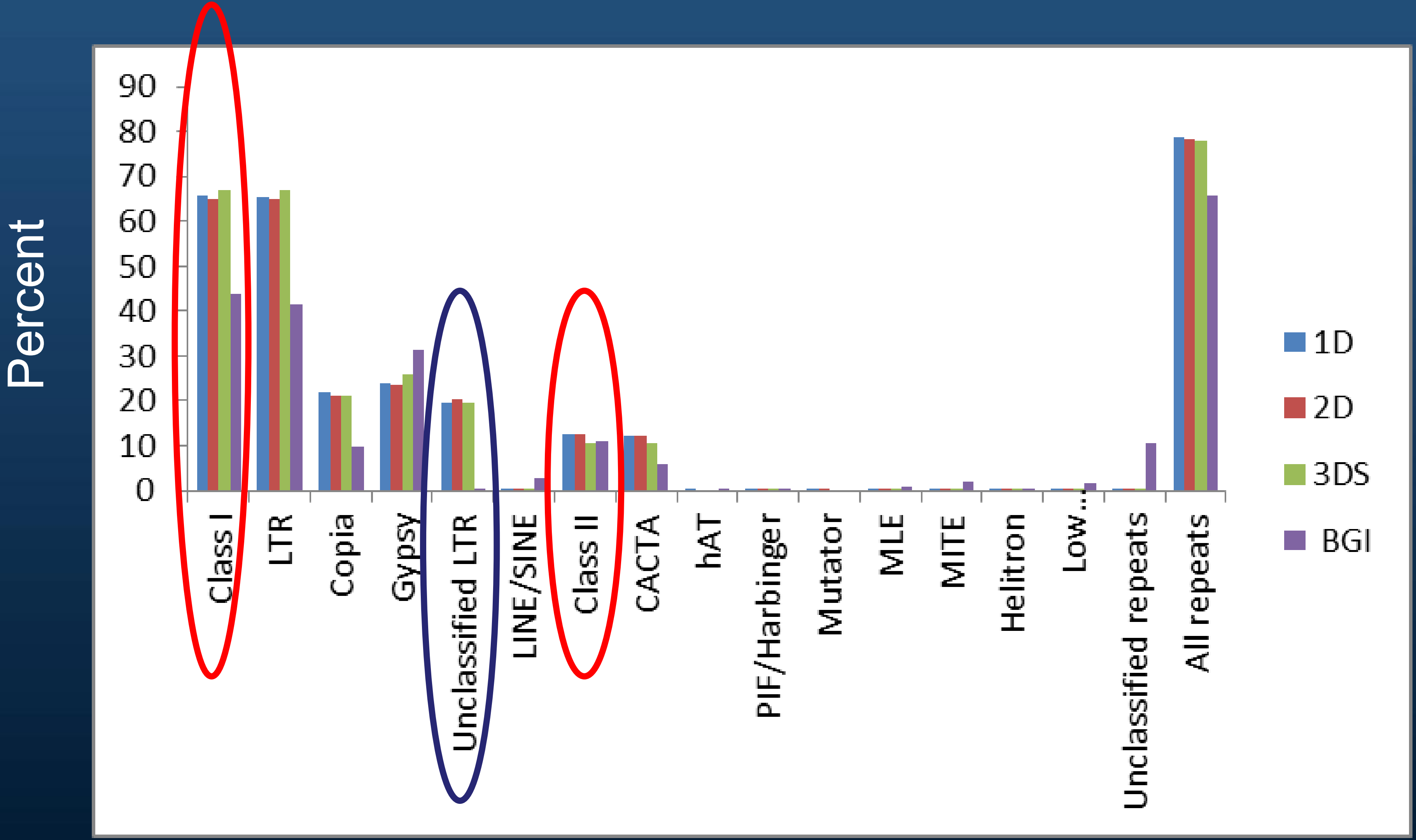
Average number of scaffolds > 2Kb per BAC pool: 18

Total scaffold length: 5.7 Gb

Average Scaffold N50 length: 203 Kb

Assembly pipeline (Poster #577)

Transposable element annotation



- TE classes and families

Gene annotation

Numbers of high-confidence genes

	Ae. tauschii			Wheat survey sequence	
Chrom.	**Anchored BAC contigs**	**Corrected for unanchored BAC contigs**	**A**	**B**	**D**
1	3,237	3,808	4,155	4,178	3,772
2	5,499	6,469	5,626	6,196	6,375
3 short	1,763	1,763	1,859		1,906

Gene annotation validation

Manual validation of annotation

- MIPS
- TriAnnot
- Maker
- Other

Annotation pipeline comparisons (Poster #576)

Pseudomolecule assembly

Chalenges

Scaffold validation, ordering, and orientating

Ordering BAC contigs in low recombination regions

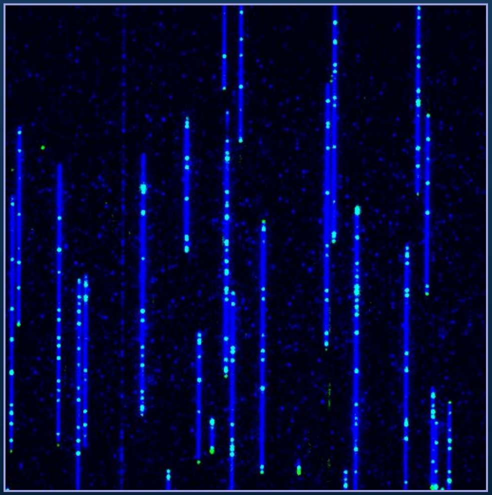
Closing gaps

New technologies
Nanomap

Nanomaps BAC contigs

(Hastie *et al.*, PloS ONE e55864, 2013)

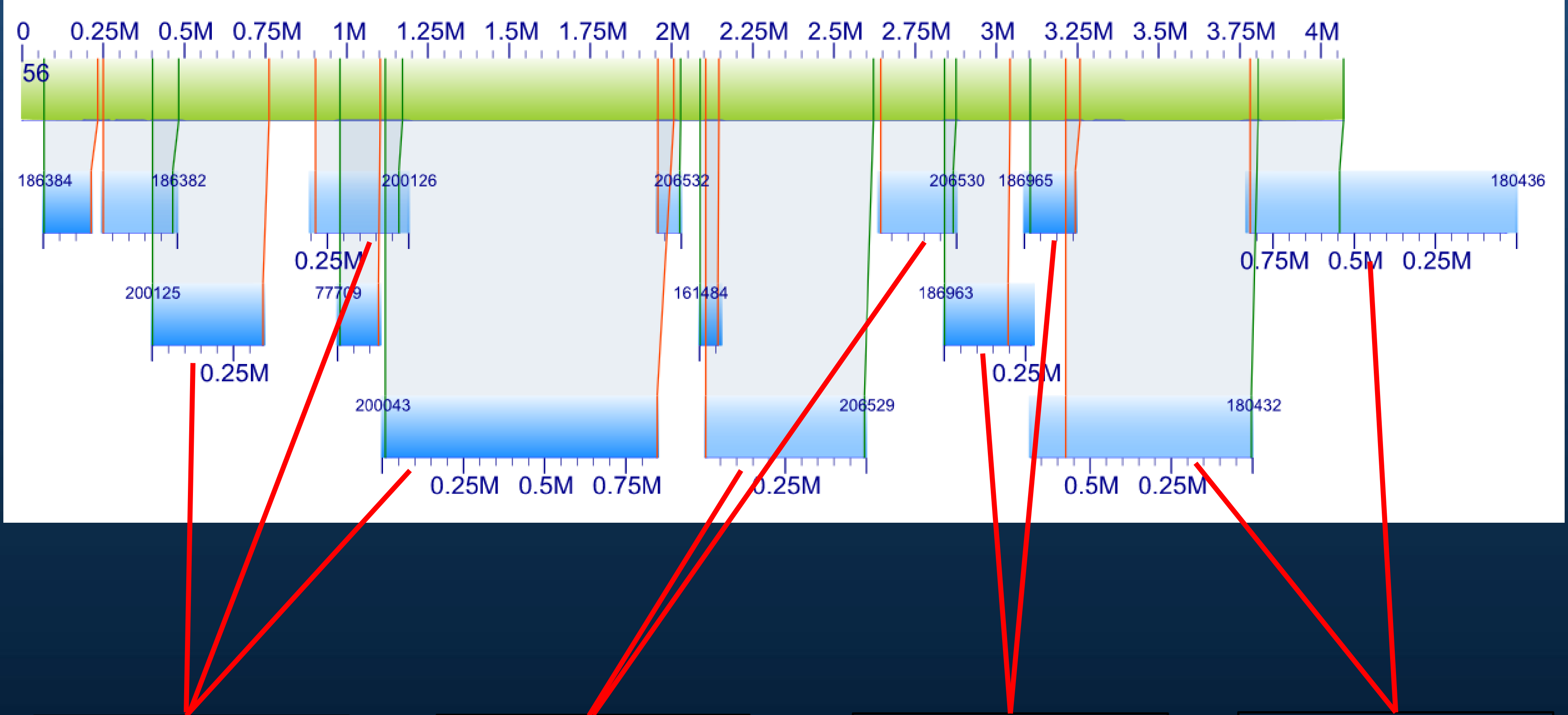
Global nanomap of the *Ae. tauschii* genome



Nanomap construction (Poster #574)
M.C. Luo’s presenations in Tuesday’s IWGSC technical workshop

Ordering and orienting sequence scaffolds on the global nanomap

🟩 Nanomap contig 🟦 Sequence scaffolds



ctg1715
6D (76.097 cM)

ctg12344
6D (76.097 cM)

ctg6115
6D (76.097 cM)

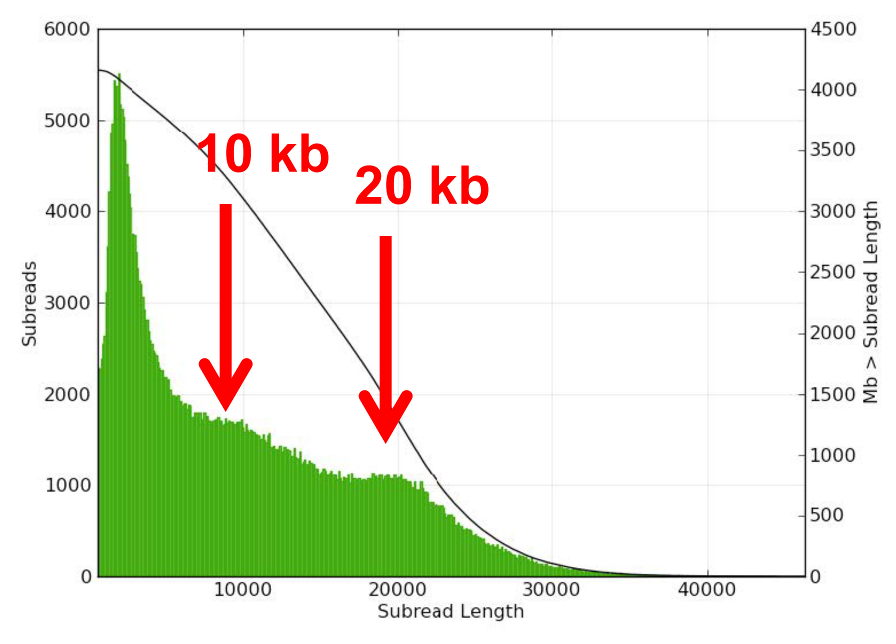
ctg195
6D (75.686 cM)

Whole genome shotgun sequence

PacBio P6 chemistry

5 SMRT cells	
Total reads	4.1 Gbp
Reads > 10 Kb	3.1 Gbp

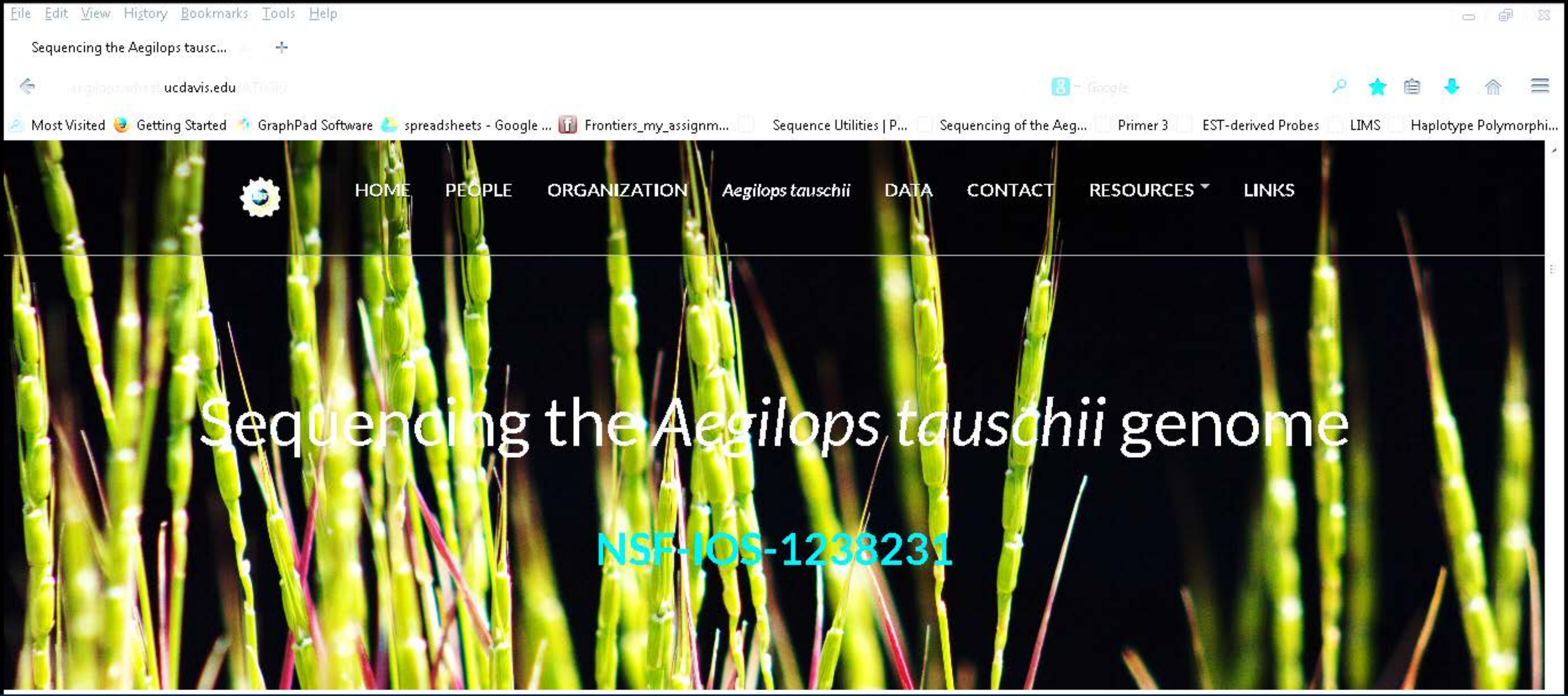
Number of reads



Read length

Where can I access data?

BLAST: http://aegilops.wheat.ucdavis.edu/ATGSP/data.php



Sequencing the *Aegilops tauschii* genome

NSF-IOS-1238231

Batch download of scaffolds also available

Acknowledgements

National Science Foundation Plant Genome

NSF

UCDAVIS

Karin Deal
Pat McGuire
Ming-Cheng Luo
Naxin Huo, Yi Wang
Chad Jorgensen
Tingting Zhu, Sonny Van
Lichan Xiao, Luxia Yuan
Luis Curiel, Scott Liu
JC Rodriguez, Thanh Ngo
Armond Murray

JOHNS HOPKINS UNIVERSITY

Steven Salzberg
Daniela Puiu
Geo Pertea

BIONANO GENOMICS

Alex Hastie
Andrew Anfora
Palak Sheth

ars

Olin Anderson
Yong Gu

Agriculture and Agri-Food Canada

Frank You

iPlant Collaborative
Empowering a new plant biology

Eric Lyons

中國農業大學
China Agricultural University

Shuhong Ouyang
Yong Liang
Zhenzhong Wang
Zhiyong Liu
Qixin Sun

University of Zurich UZH

Thomas Wicker

CSH Cold Spring Harbor Laboratory

Richard McCombie

1785
The University of Georgia

Katrien Devos
Hao Wang
Jeffrey Bennetzen

INRA

Philippe Leroy

mips
munich information center for protein sequences

Klaus Mayer
Matthias Pfeifer

Nanjing Agricultural University

Zhengqiang Ma

中國农业科学院
CHINESE ACADEMY OF AGRICULTURAL SCIENCES

Long Mao

IEB

Jaroslav Doležel

ARIZONA

Cari Soderlund