Wheat genome analysis using Illumina paired read sequencing

David Edwards
University of Queensland
Dave.Edwards@uq.edu.au

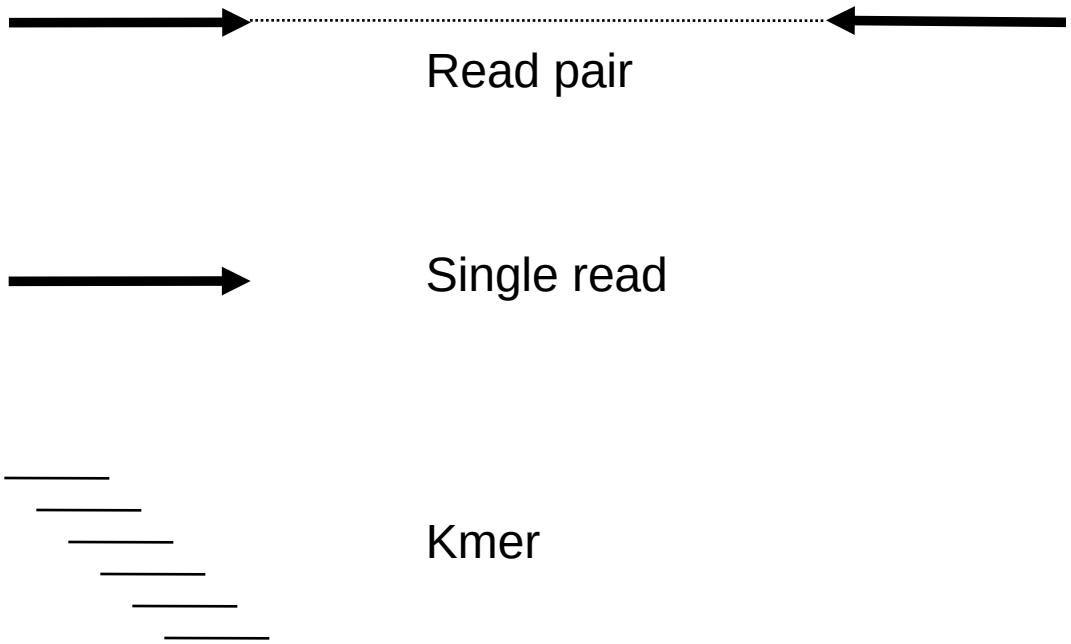
Outline

- Second-generation sequence data
- TAGdb and gene assembly
- Comparison of 7DS data with rice

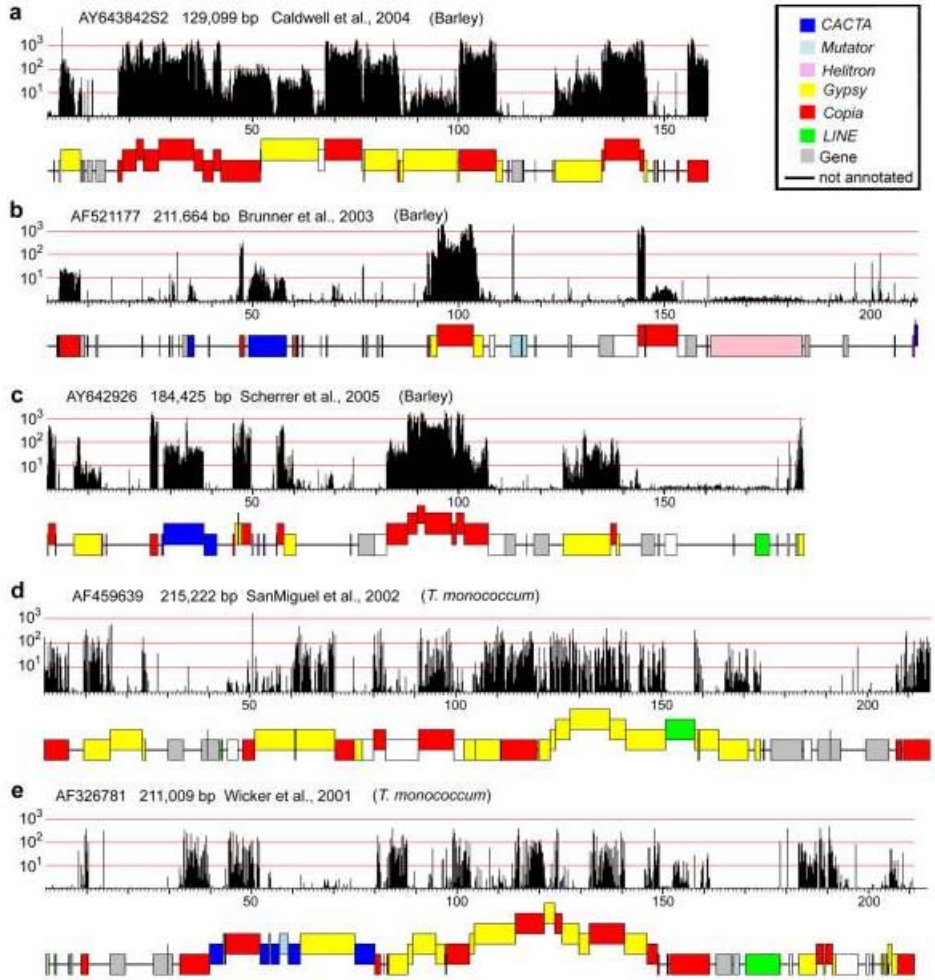
Paired-end short reads

- Illumina GAIIx
- Read length (35bp – 75bp)
- Insert size up to 10kbp
  - ~ Normal distribution
  - Standard deviation ~ 10% mean

Paired-end short reads



Kmer analysis



Wicker et al. (2008)

TAGdb – Data

Wheat WGS	2.2 Gbp	0.1x
Wheat 7DS	2.5 Gbp (6.3 Gbp)	6.2x (16x)
Barley WGS	2.5 Gbp	0.5x
Brassica rapa	5.5 Gbp	10.4x
Brassica oleracea	1.1 Gbp	1.7x
Brassica nigra	1.2 Gbp	1.8x
Pongamia pinnata	0.5 Gbp	0.3x
Nicotiana alata	5.1 Gbp	1.1x

We welcome more data!

http://flora.acpfg.com.au/tagdb/

Welcome to ACPFG Bioinformatics TAGdb.

This service performs BLAST alignment between a single query and short pair reads of selected species.

Please enter a valid email address

Note: Your result will be sent to the specified address.

Sequence data

Either: Select the sequence file to upload:

Browse...

Otherwise: Enter a sequence in FASTA format:

Note: Query sequence must be less than 5000 nucleotides.

Species selection

Please choose a query species

- Barley
- Brassica
- Nicotiana
- Pongamia
- Wheat

Short paired-read library selection

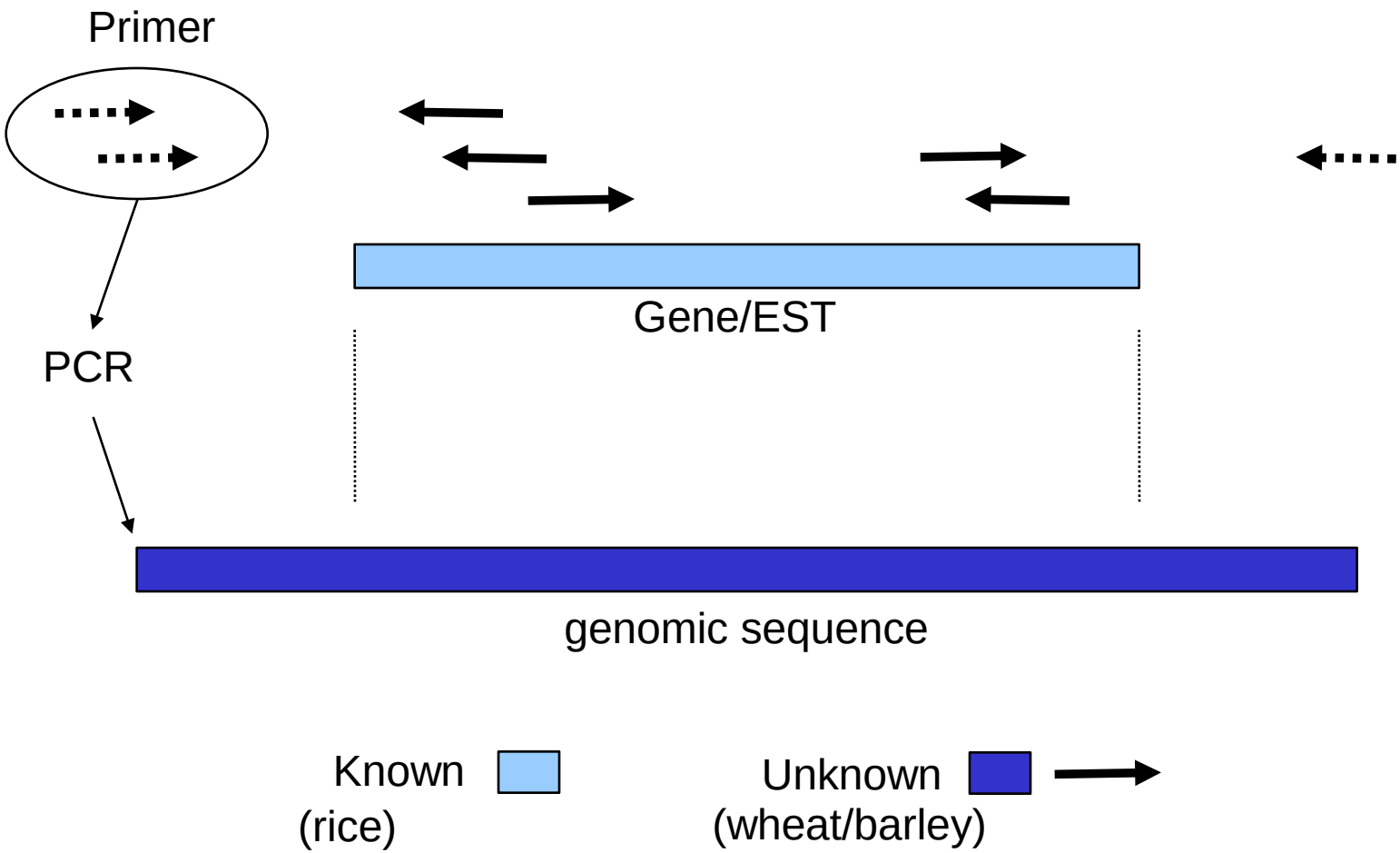
Please select one or more paired-read libraries to search

- Chinese Spring - 36 - 300 - WCs_03_001
- Chinese Spring - 76 - 320 - 7DS_03_002
- Chinese Spring - 76 - 3700 - WCs_37_001
- Chinese Spring - 76 - 300 - WCs_03_002
- Chinese Spring - 76 - 320 - 7DS_03_001
- Chinese Spring - 75 - 320 - 7DS_03_003

Format: SourceName - ReadLength - InsertSize - LibraryName

Start

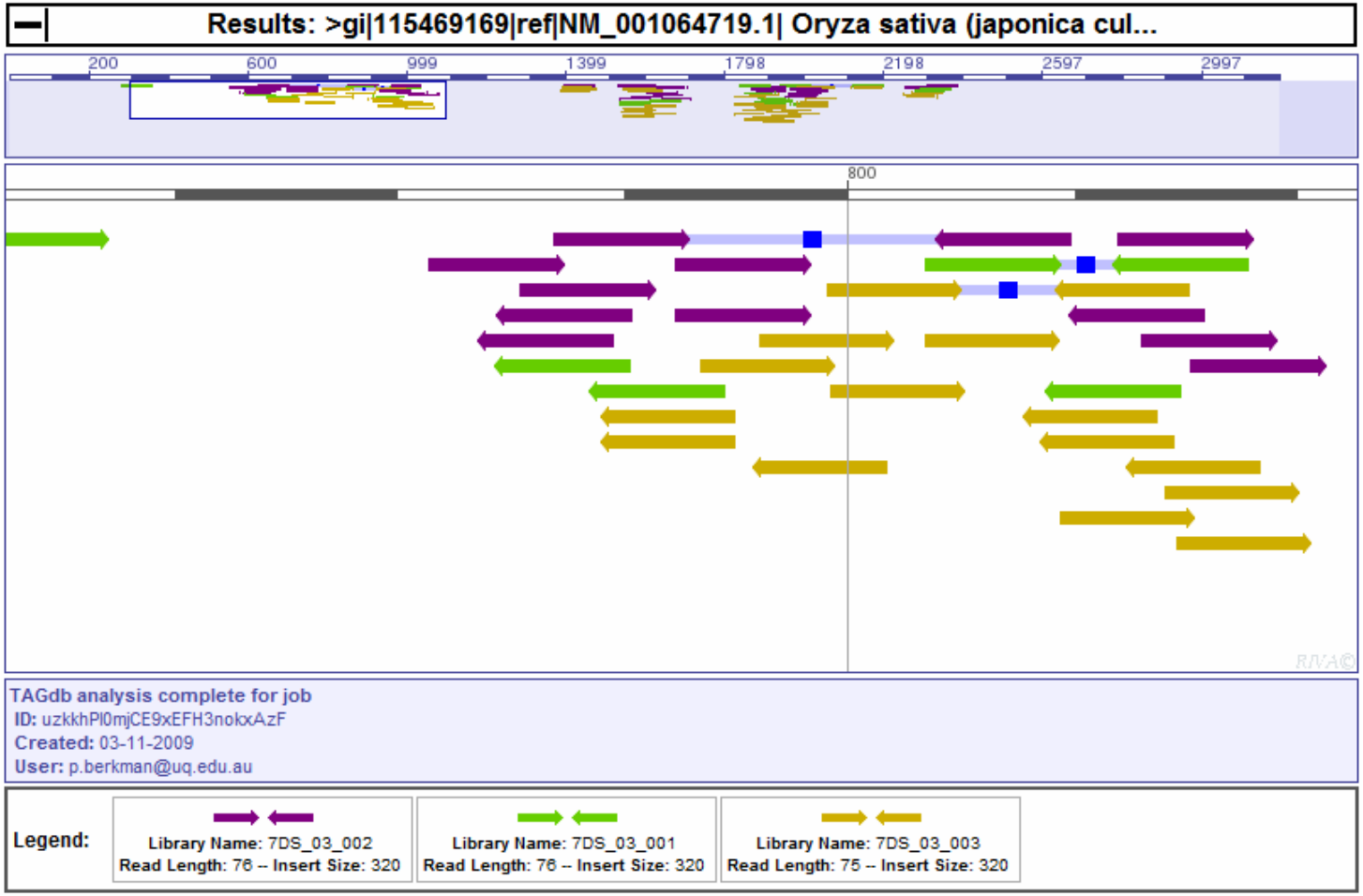
TAGdb – Gene finding and extension



TAGdb example – OVP1 ortholog in wheat

- OVP1 gene – rice gene orthologous to AVP1
  - Vacuolar H+-pyrophosphatase
- Compare rice gene sequence with wheat chromosome arm 7DS data
- Assemble sequences
- Repeat

TAGdb – OVP1 vs. wheat 7DS



http://flora.acpfg.com.au/tagdb/cgi-bin/results?jobID=uzkkhPl0mjCE9xEFH3nokxAzF

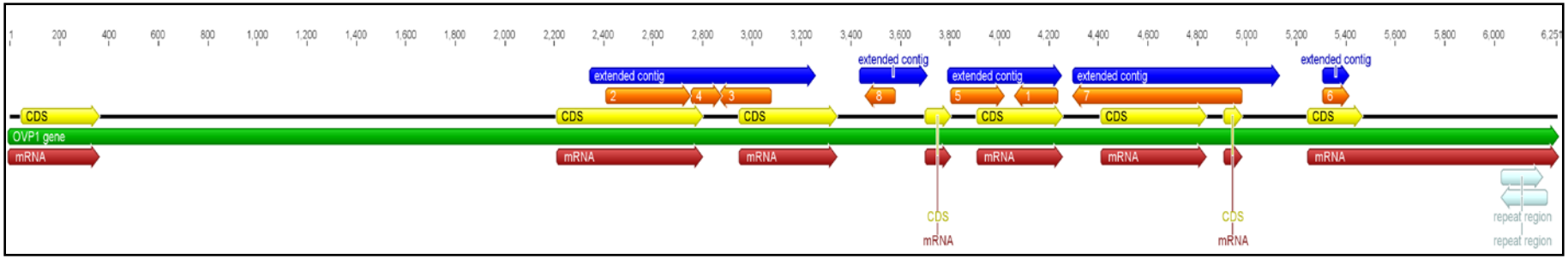
TAGdb – 7DS OVP1 ortholog assembly

- 93 7DS read pairs exported from TAGdb
- Assembled in velvet (kmer size 31)
- 8 contigs assembled
  - maximum length of 670
  - total length of 1670

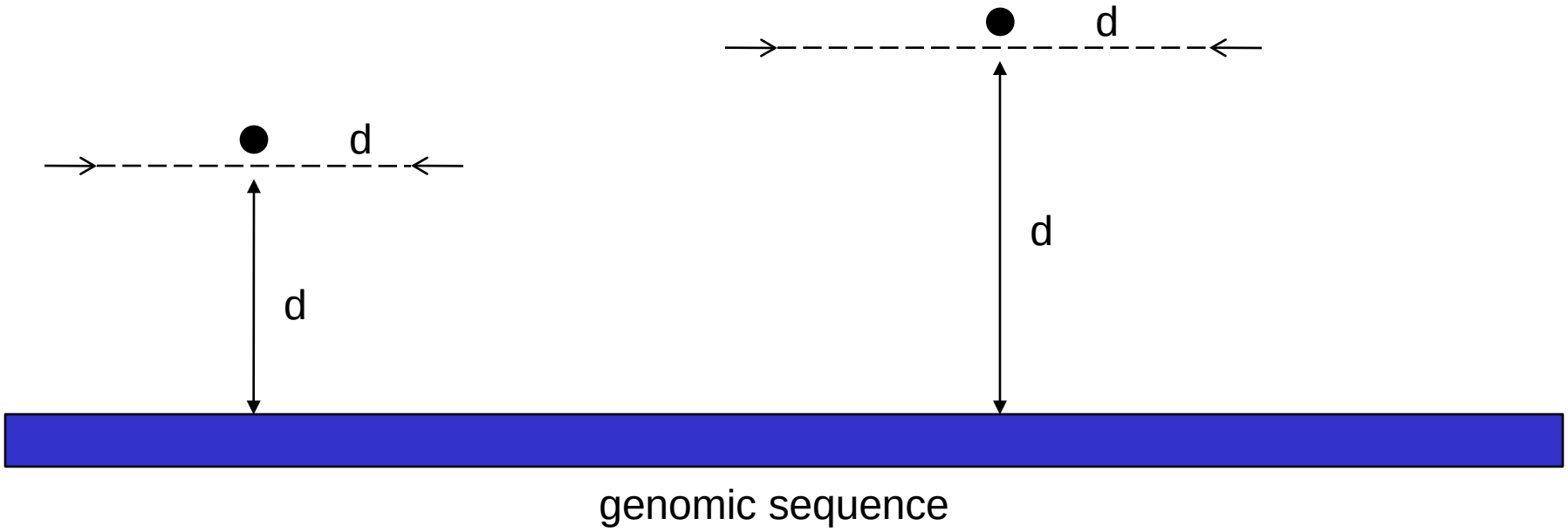
TAGdb – 7DS OVP1 ortholog assembly

- Contigs extended through TAGdb & velvet
  - 5 contigs, total 3240 bp

- Contigs indicate 7DS ortholog genomic context

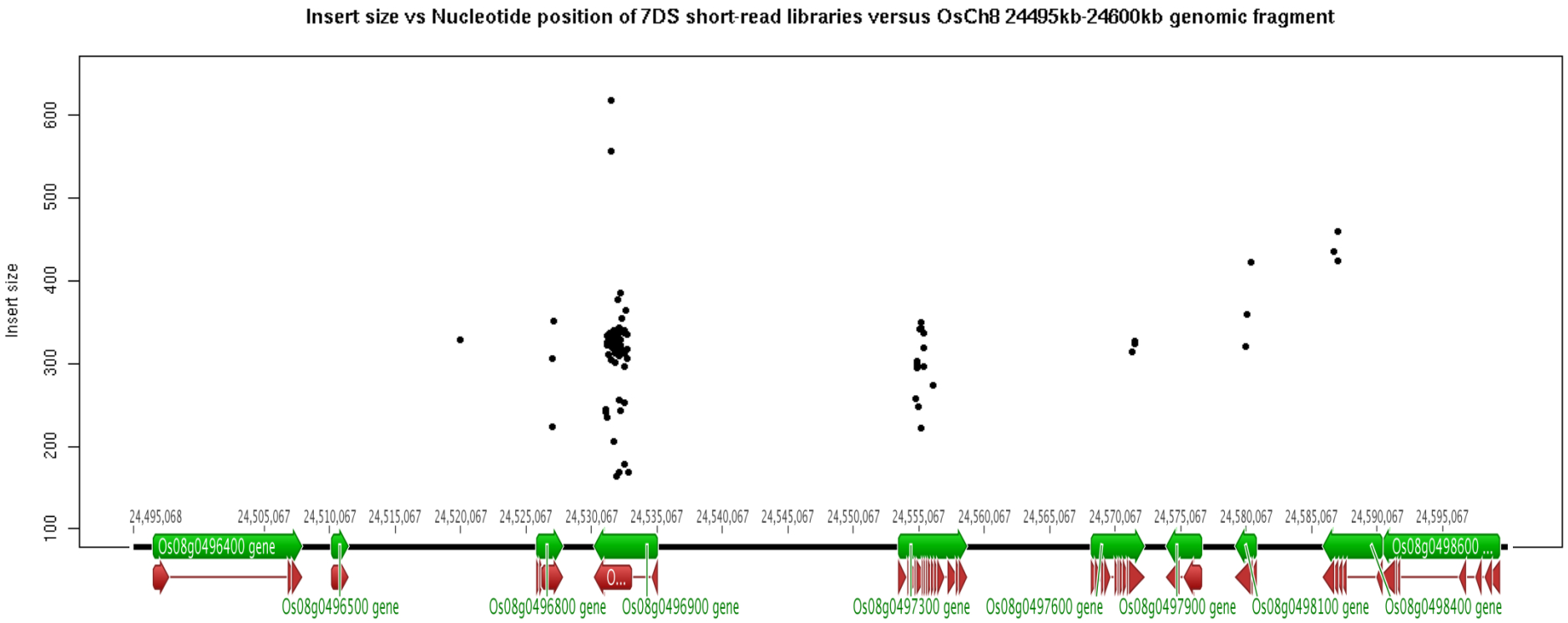


Comparison with rice



THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

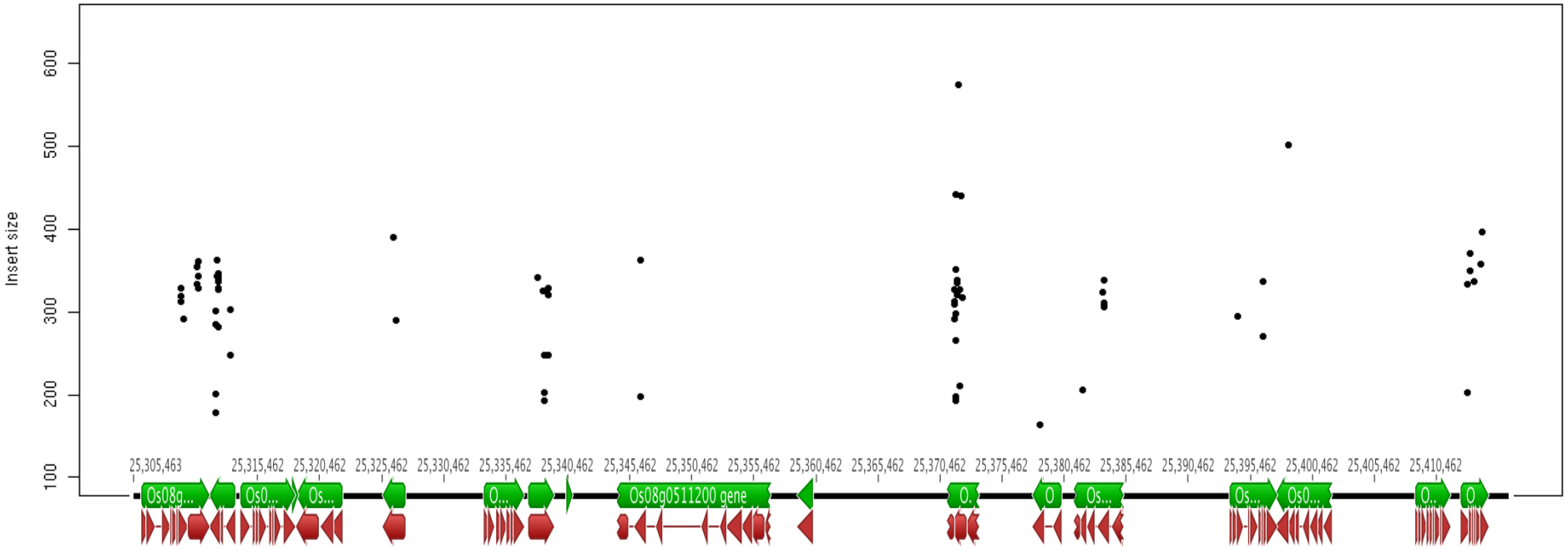
ACPFG
AUSTRALIAN CENTRE FOR PLANT FUNCTIONAL GENOMICS PTY LTD



THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

ACPFG
AUSTRALIAN CENTRE FOR PLANT FUNCTIONAL GENOMICS PTY LTD

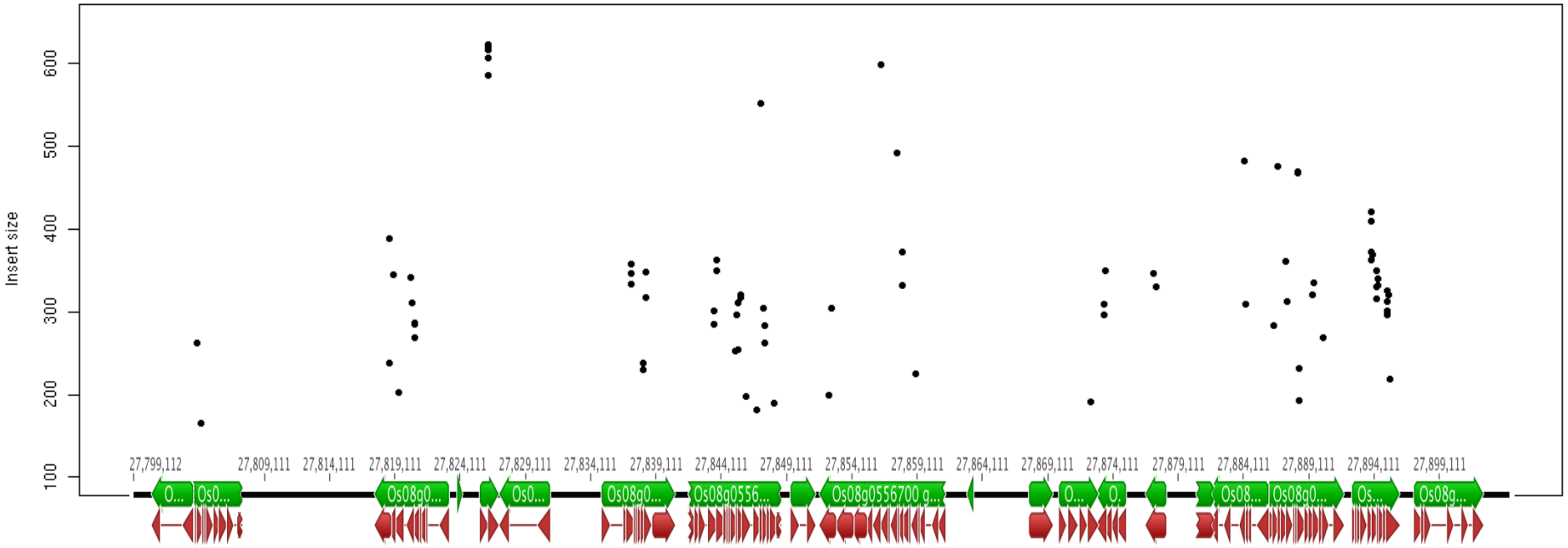
Insert size vs Nucleotide position of 7DS short-read libraries versus OsCh8 25305kb-25415kb genomic fragment



THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

ACPFG
AUSTRALIAN CENTRE FOR PLANT FUNCTIONAL GENOMICS PTY LTD

Insert size vs Nucleotide position of 7DS short-read libraries versus OsCh8 27800kb-27905kb genomic fragment



Summary

- Second-generation sequence data is very useful
- TAGdb can be used to identify and sequence chromosome arm specific genes
- Comparison with rice identifies genes on 7DS

Acknowledgements

Paul Berkman
Chris Duran
Dominic Eales
Michael Imelfort
Jiri Stiller
Terry Clark
Michal Lorenc
Edmund Ling
Megan McKenzie

Daniel Marshall
Zoran Boskovic
Nikki Appleby
Ping Zhang
Chang-Pyo Hong
Kaye Basford
Jaroslav Dolezel
Jacqueline Batley

Contact:
Dave.Edwards@uq.edu.au