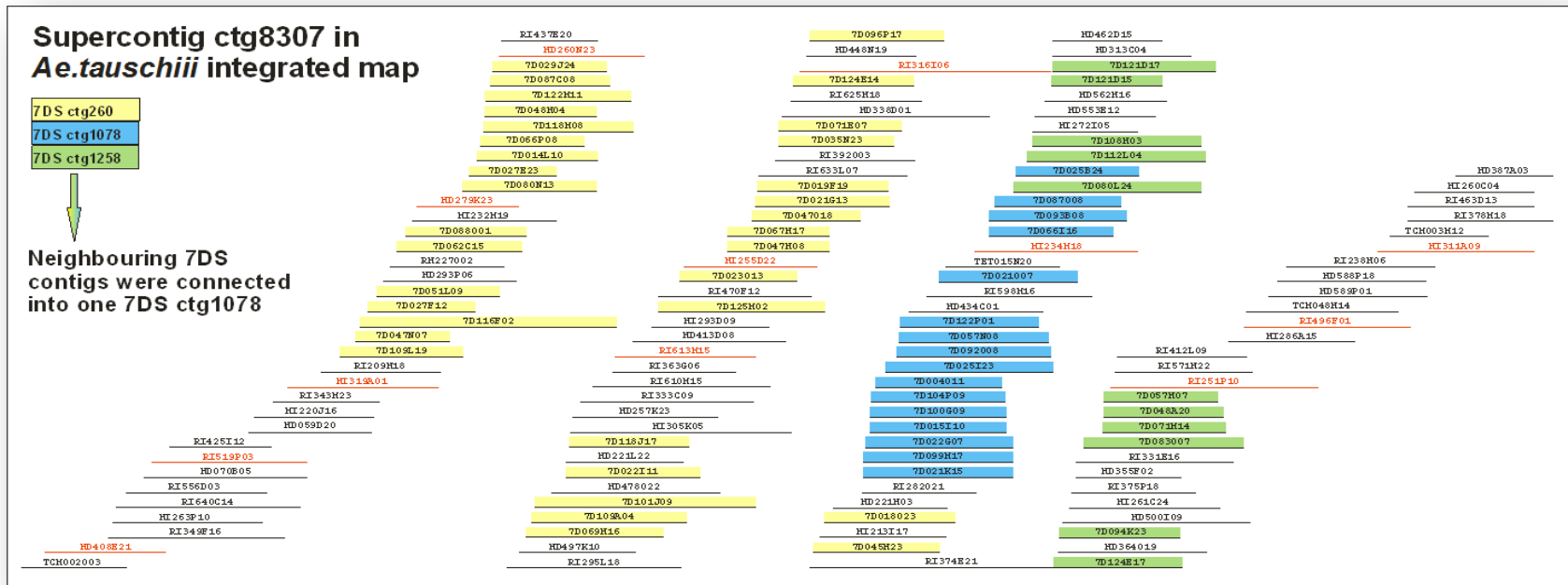


PRACTICAL ASPECTS OF BIONANO MAPPING OF THE WHEAT GENOME

Hana Šimková

**Institute of Experimental Botany
Olomouc**

7DS PHYSICAL MAP



Physical map

- Automatic assembly using FPC
- Manual end-merging of contigs based on integration of *Aegilops tauschii* whole-genome map with the map of the bread wheat 7DS chromosome arm
- Verification of the assembly using LTC

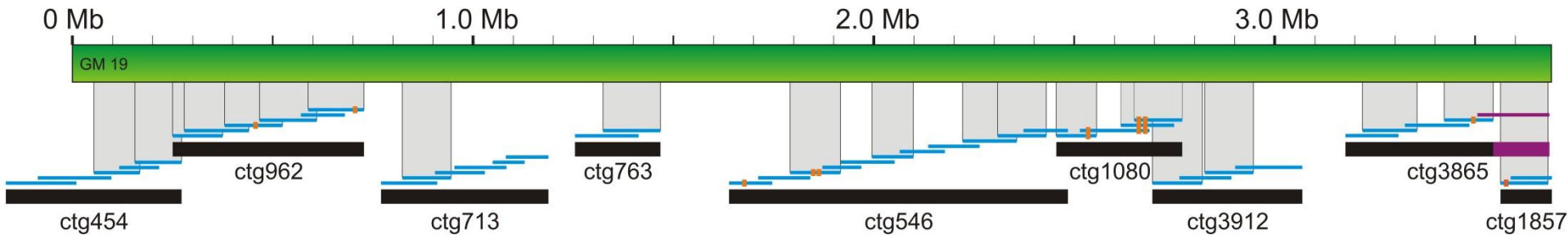
No. contigs > 2 BAC clones	904
No. contigs > 5 clones	652
Assembly length	360 Mb
7DS arm coverage	95%
Contig N50	548 kb
No. MTP BAC clones	4,608

PHYSICAL MAP ANCHORING

<u>Strategy</u>	<u>Marker type and map</u>	<u>No. markers</u>	<u>No. ctg anchored</u>
• Forward anchoring			
- Manual (3D pools screening)	SSR genetic (GrainGenes)	23	27
	STS genetic (<i>owm</i> , Staňková <i>et al.</i> , 2015)	11	8
- <i>In silico</i> (sequence-based)	SNP genetic <i>Ae. tauschii</i> (Luo <i>et al.</i> , 2013)	516	302
	DArTseq genetic (Kilian, unpublished)	556	295
	RJM – RH CS (Tiwari, unpublished)	205	150
	SNP – RH CS (Tiwari, unpublished)	202	144
	DArT – RH <i>Ae. tauschii</i> (Kumar <i>et al.</i> , 2015)	51	14
• Reverse anchoring			
- Based on BAC sequences	STS – RH <i>Ae. tauschii</i>	29	26

73% of the physical map have been **anchored** → 508 contigs through 1,601 markers

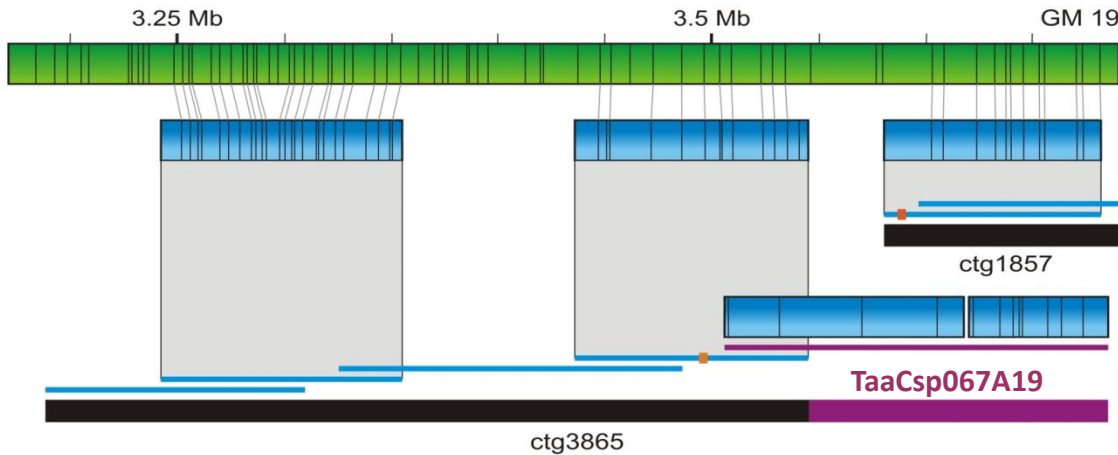
PHYSICAL MAP ANCHORING THROUGH BIONANO MAP



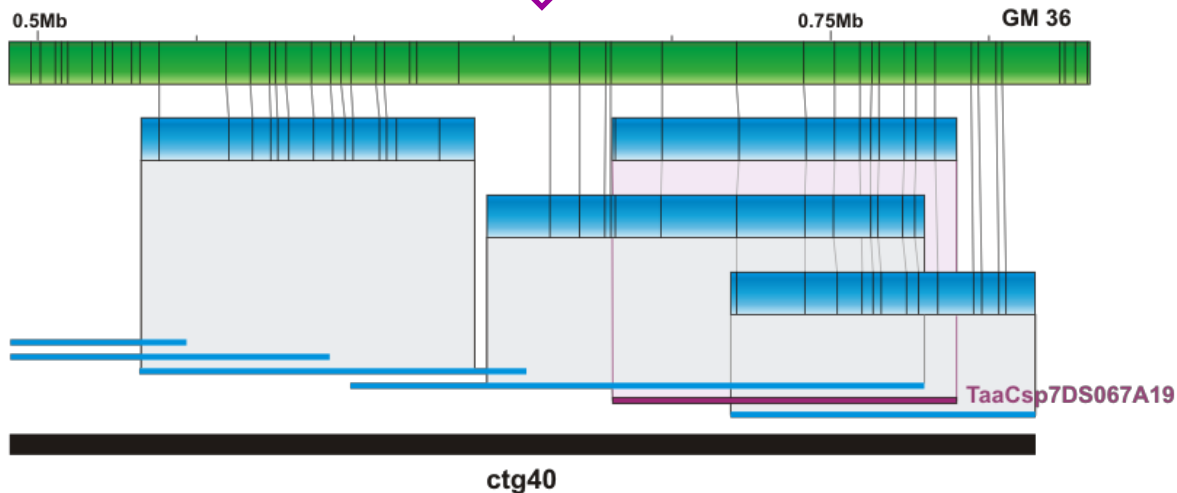
- Additional 36 BAC contigs anchored and oriented through BioNano map → 4% of assembly anchored through BioNano map only, the work is ongoing
- BioNano map resolved orientation of contigs with a single marker
- Useful in non-recombining regions BUT...

BIONANO MAP FOR PHYSICAL MAP IMPROVEMENT

□ Identifies and corrects misassemblies



- Clone TaaCsp067A19 incorrectly positioned in ctg3865



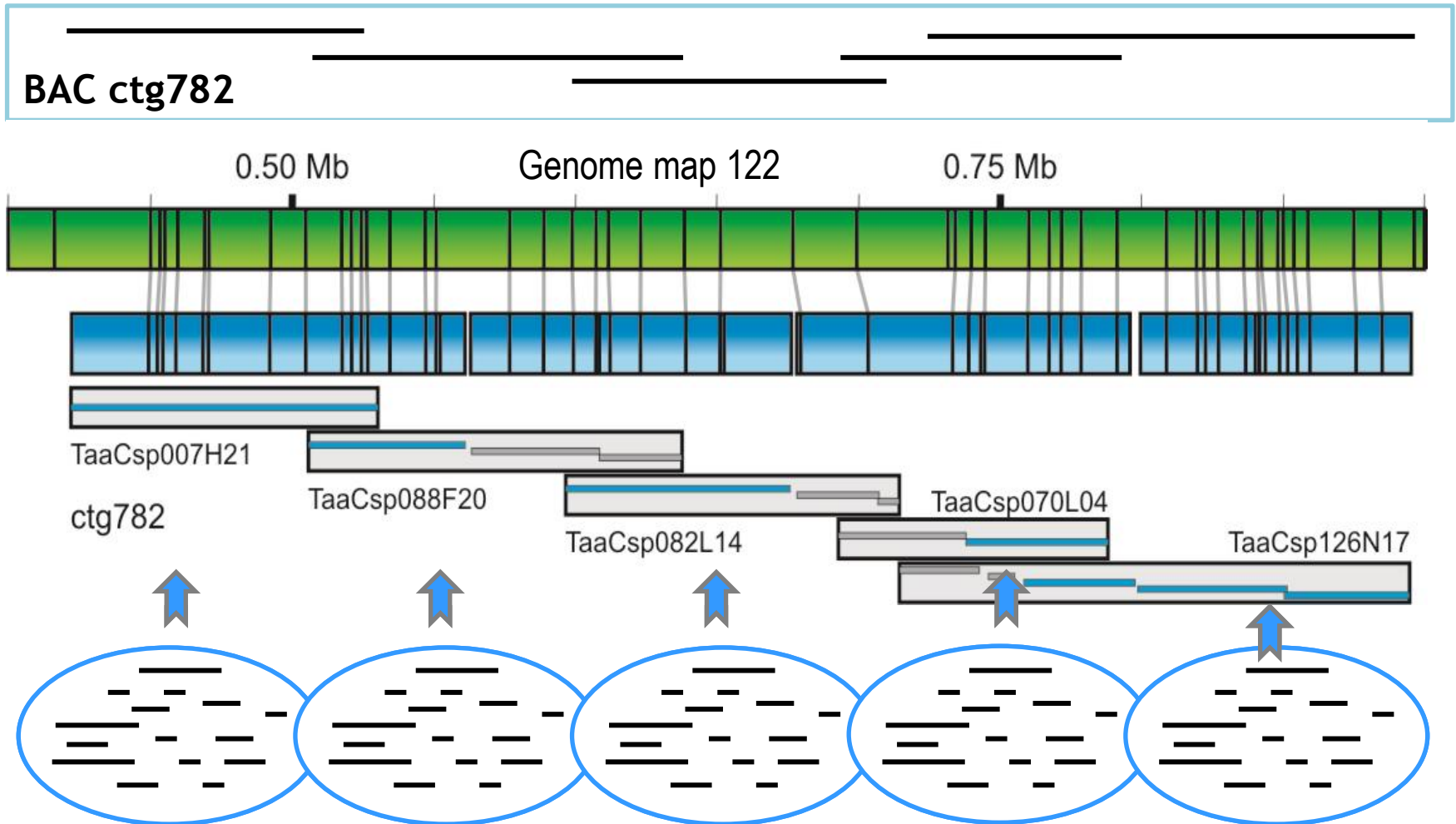
- Correct position found through BioNano map in ctg40

7DS SEQUENCING

- Sequencing pools of four non-overlapping MTP BAC clones
- Illumina pair-end sequencing
 - 550bp fragment size,
100 bp read length,
coverage >500x → Sassy
- Deconvolution through BAC end sequences
 - part of inner contigs unresolved (<10% assembly)
- Scaffolding through 6-10 kb mate-pair data obtained from MTP-plate pools (384 clones/pool) → SSPACE
- Pseudomolecule construction is underway

Contigs per BAC	1 - 17
Average No. contigs/BAC	1.9
Median No. contigs/BAC	1.5
Average scaffold size	56 kb
Scaffold N50	116 kb
% N in assembly	2.9 %
No. scaffolds	9,063

BIONANO MAP FOR POOL DECONVOLUTION



- 1) Sequences of five BAC pools aligned to the 7DS BioNano map
→ 1-3 contigs per clone anchored to genome map 122
- 2) The remaining contigs resolved through BLASTing BAC pools against each other

PRELIMINARY SEQUENCE ANNOTATION

The sequence assembly comprises 91% of gene models identified in 7DS CSS sequences



- 9,084 sequence scaffolds analyzed through TriAnnot pipeline version v5.1p02
- Supported by 7DS RNAseq data obtained from RNAseq mapping on the 7DS scaffolds

Total No. 7DS transcripts identified through Tuxedo suit	2,584
• No. transcripts mapped on 7DS scaffolds	2,364 (91%)
• No. transcripts mapped on 7DS CSS sequences	1,556 (60%)

PRELIMINARY SEQUENCE ANNOTATION

No. scaffolds in 7DS assembly	9,084
• No. analyzed scaffolds (>10 kb)	6,152
• No. scaffolds with at least one gene	1,873 (~30%)
No. predicted protein-coding genes	3,454
• No. high confidence genes	1,916 (~55.5%)
No. pseudogenes	63,262
No. TEs	59,668 (15 families)
• Class I – LTR: <i>Copia</i> & <i>Gypsy</i>	9,218 + 15,796 (~42%)
• Class II – TIR: <i>CACTA</i>	13,908 (~23%)



CONTRIBUTION OF SHORT BAC CONTIGS

2-BAC contigs - excluded from sequencing

No. 2-BAC contigs	550
Total length 2-BAC contigs	83,152 kb
% phys. map assembly	18.7%
Total assembly length	443 Mb (116% arm size)

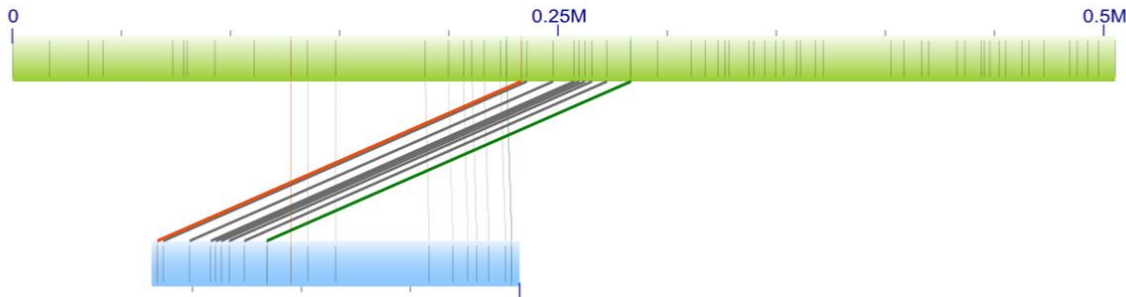
	3-BAC	4-BAC	5-BAC	Total
No. short contigs in 7DS physical map	169	58	25	252
No. anchored contigs	32	16	17	65
• through markers (+ BioNano map)	22	13	13	48
• through BioNano map only	10	3	4	17
Total length of short contigs [kbp]	25,800	10,605	4,214	40,873
• length of anchored (% phys. map assembly)	4,867 (1,4%)	3,254 (0,9%)	2,736 (0,8%)	11,111 (3%)
• length of non-anchored (% phys. map assembly)	21,187 (5,9%)	7,351 (2,0%)	1,478 (0,4%)	29,762 (8%)

➤ Significant part of the short contigs may not belong to the 7DS

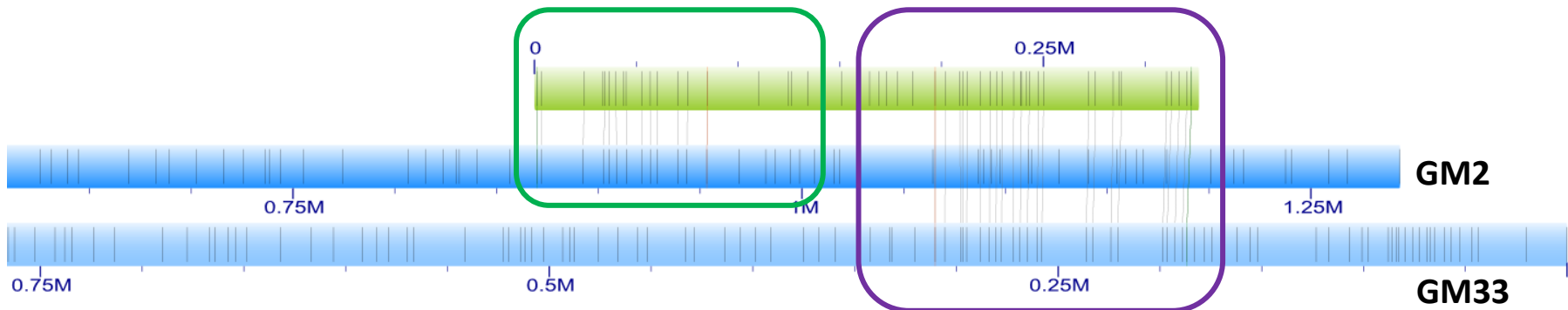
BIONANO MAP FOR SEQUENCE IMPROVEMENT

□ Identification of scaffold mis-assemblies

- Due to mis-orientation contigs in a scaffold



- Due to mis-joining of contigs within a sequencing pool or cross-contamination



AVAILABLE CHROMOSOMAL BIONANO MAPS

	7AS	7AL	7BS	7DS-CS	7DS-CI2401
Molecule Coverage	194x	230x	700x	180x	206x
Molecule N50 (kb)	206	232	236	344	219
No. Genome Maps	783	330	254	371	468
Avg. Genome Map Length (kb)	571	1,251	1,397	900	765
Genome Map N50 (kb)	1,553	2,069	1,999	1,300	1,355

- 7BL coming soon

Each arm - data from one Irys chip



GOOD NEWS FROM BIONANO

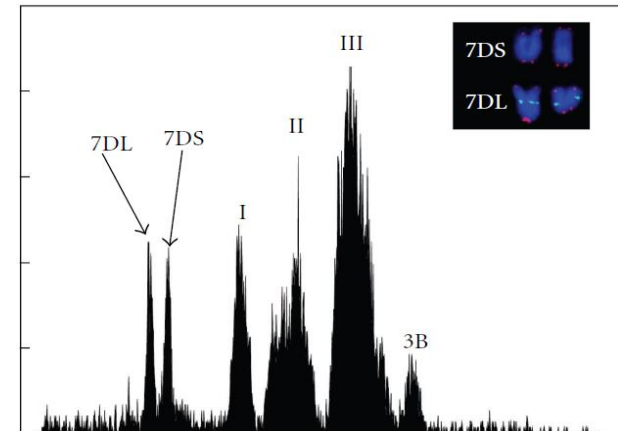
□ New tool for sample preparation will be available soon

- Reduces required DNA amount four times

- Sorting time reduced to 1-2 weeks per two arms (simultaneously sorted)

- Larger molecules

- Need for lower coverage
→ two arms could be done on one chip



□ New enzyme with higher labelling frequency
Nt.BssSI (CACGAG)

- Potentially additional information



ACKNOWLEDGEMENTS



Helena Staňková
Zuzana Tulpová
Jan Vrána
Marie Kubaláková
Jaroslav Doležel



Mingcheng Luo



Ajay Kumar



Alex Hastie
Saki Chan



Philippe Leroy
Hélène Rimbart



Andrzej Kilian



David Edwards
Paul Visendi
Jacqueline Batley
Satomi Hayashi



Kansas State University
Vijay Tiwari
Bikram Gill
Bernd Friebe



EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

