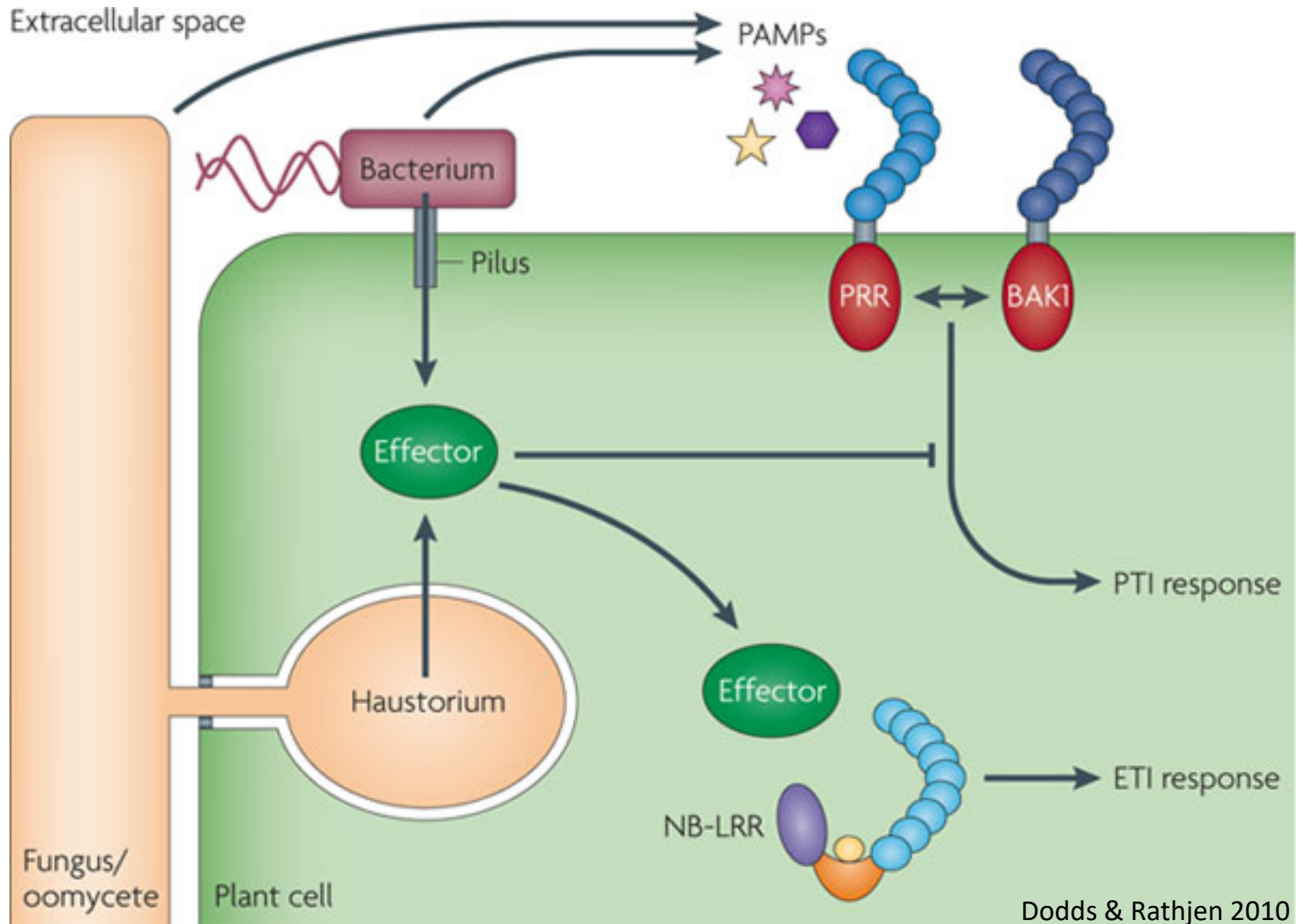# The intracellular immune receptor repertoire of wheat

Burkhard Steuernagel

Plant Pathogen Interaction

Dodds & Rathjen 2010

Consensus structure of a CC-NLR

Identification and localisation of the NB-LRR gene family within the potato genome

Florian Jupe[1,2], Leighton Pritchard[3], Graham J Etherington[2], Katrin MacKenzie[4], Peter JA Cock[3], Frank Wright[4], Sanjeev Kumar Sharma[1], Dan Bolser[5], Glenn J Bryan[1], Jonathan DG Jones[2] and Ingo Hein[1*]

| 17 | 16 | | 1 | 6 | 4 | 5 | 10 | 3 | 12 | 2 | 8 | 7 | 9 | 11 | 19 | Jupe *et al.* |

| Coiled Coil | NB-ARC | Linker | LRRs | |

| | EDVID | | P-loop | RNBS-A | Kin-2 | RNBS-B | RNBS-C | GLPL | RNBS-D | MHDV | Motif1 LDL | Meyers *et al.* |

Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily

Blake C. Meyers   , 1 2, , Allan W. Dickerman,   3, Richard W. Michelmore,
1 Subramoniam Sivaramakrishnan,   1 4 Bruno W. Sobral,   and 3 Nevin D. Young 5,
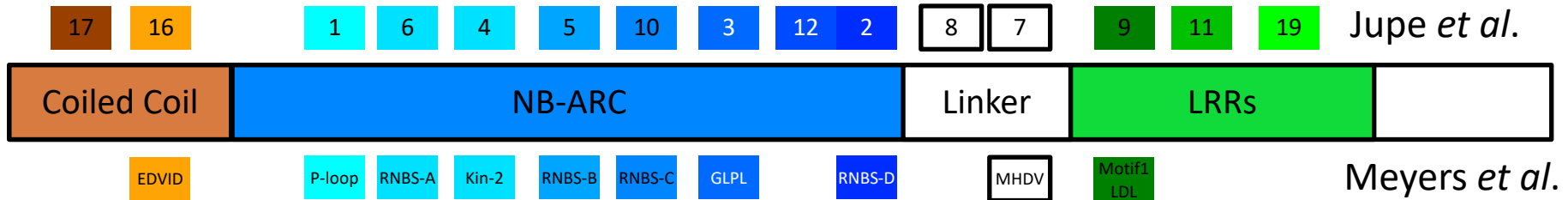
NLR-Parser

BMC Genomics

RESEARCH ARTICLE **Open Access**

Identification and localisation of the NB-LRR gene family within the potato genome

Florian Jupe[1,2], Leighton Pritchard[3], Graham J Etherington[2], Katrin MacKenzie[4], Peter JA Cock[3], Frank Wright[4], Sanjeev Kumar Sharma[1], Dan Bolser[5], Glenn J Bryan[1], Jonathan DG Jones[2] and Ingo Hein[1*]

Bioinformatics Advance Access published February 5, 2015

## NLR-parser: rapid annotation of plant NLR complements

Burkhard Steuernagel[1,*,†], Florian Jupe[2,*,†,‡], Kamil Witek[2], Jonathan D.G. Jones[2] and Brande B.H. Wulff[1]

OXFORD

	nplete	strand	start	end	1st-motif-frame	NB-ARC	2-NBLRR-Signal	MotifList
	tial	forward	38	740	2	N/A	FALSE	12,2,8,7,11,11,9
	tial	forward	18	1297	0	LCFSVIGMHGVGGYGKSTL	FALSE	1,6,4,5,10,3,2,8,7,9
	tial	forward	132	837	0	N/A	FALSE	12,2,8,7,9,11
	nplete	reverse	33	3933	1	KYCADGLTMLHIVGPGGIG	TRUE	17,16,9,1,6,4,5,10,3,12,2,8,7,15,19,11,11,11,11,11,11,11,11
	tial	forward	320	908	2	MDARKKLKVVSIVGIGGLG	FALSE	17,16,14,1
	tial	reverse	246	2387	1	NSRDGLCKVICVVGMGGL	FALSE	1,6,4,5,3,12,2,8,7,9,11
	tial	reverse	37	2520	1	PDNNRGCKVVSVWGMGG	FALSE	17,6,4,5,1,6,4,5,10
	tial	reverse	270	1032	0	N/A	FALSE	16,11,6,4,5
	tial	reverse	52	934	1	N/A	FALSE	3,12,2,8,7,9,11,9
	tial	reverse	72	774	0	NSLTKHRKVLSIVGFGGLG	FALSE	17,16,1,6
	tial	reverse	286	976	2	DDKEMQLKVVAIVGSGGL	FALSE	17,16,14,1,6
contig_3289_1	CNL partial	reverse	6	711	2	GSDDVECSLITLTGMGGLG	FALSE	17,16,1,6
contig_17782_1	CNL partial	reverse	945	1992	2	N/A	FALSE	5,10,3,12,2,8,7,9,9
contig_678_1	CNL partial	reverse	264	2882	2	VDNDYLLSVVSIVAMGGA	FALSE	17,1,6,4,5,10,3,12,2,8,7,9,9,9,11,11
contig_1224_2	CNL partial	reverse	185	1025	0	N/A	FALSE	3,12,2,8,7,11,11,9
contig_8213_1	CNL partial	reverse	83	1775	2	N/A	FALSE	19,17,16,14
contig_10715_1	CNL partial	reverse	338	1115	1	EERMTELPVLPIVGMGGV	FALSE	16,1,6,4,5
contig_11902_1	CNL partial	reverse	263	3965	0	ETKDEGISVFAIVGNGGVG	FALSE	1,6,4,5,10,3,12,2,8,7,9,11,11,11,11,11,11,11,11,11
contig_15355_1	N/A partial	reverse	163	424	0	N/A	FALSE	5,10,3
contig_9193_1	N/A partial	forward	167	1964	2	N/A	FALSE	9,4,9,11,11,20,9,11,11,11,11,9,11,11
contig_10690_1	TNL partial	forward	475	849	1	QAHHQVFLFWW*FELVV	FALSE	1,4,5
contig_9296_1	CNL partial	forward	5	740	2	N/A	FALSE	3,12,2,8,7,9
contig_2126_1	CNL partial	reverse	1273	2610	0	N/A	FALSE	10,3,12,2,8,7,11,9,9,20
contig_5402_1	N/A partial	reverse	96	438	1	N/A	FALSE	4,5,10,3
contig_3210_1	CNL partial	forward	224	983	2	DLEQQQLKVVAVLGPSCV	FALSE	16,1,6,4
contig_3416_1	CNL partial	reverse	126	1474	1	FIKDYLVGMVGIWGPGGV	FALSE	1,6,4,5,10,3,2,7,9,11,9
contig_4215_1	CNL partial	forward	982	2804	1	AIAREGISIIPIVGLGGLGKS	FALSE	17,16,1,6,4,5,10,3,2
contig_1121_1	CNL partial	forward	271	1288	1	KEGNTNIMVVAVVGVGGI	FALSE	17,16,1,6,4,5,10
contig_1327_1	CNL partial	reverse	189	1209	2	KEVNDGIMVVAIVGVGGI	FALSE	17,16,1,6,4,5,10
contig_2229_1	CNL partial	forward	549	1365	0	QQHQQKVHVLPIFGMGG	FALSE	16,1,6,4,5,10
contig_4861_2	CNL partial	forward	438	861	0	NKDLQHFHVISLWGMGG	FALSE	1,6,4,5,10

NLR-Annotator

Chop sequence to overlapping fragments

NLR-Parser
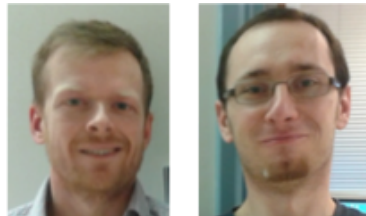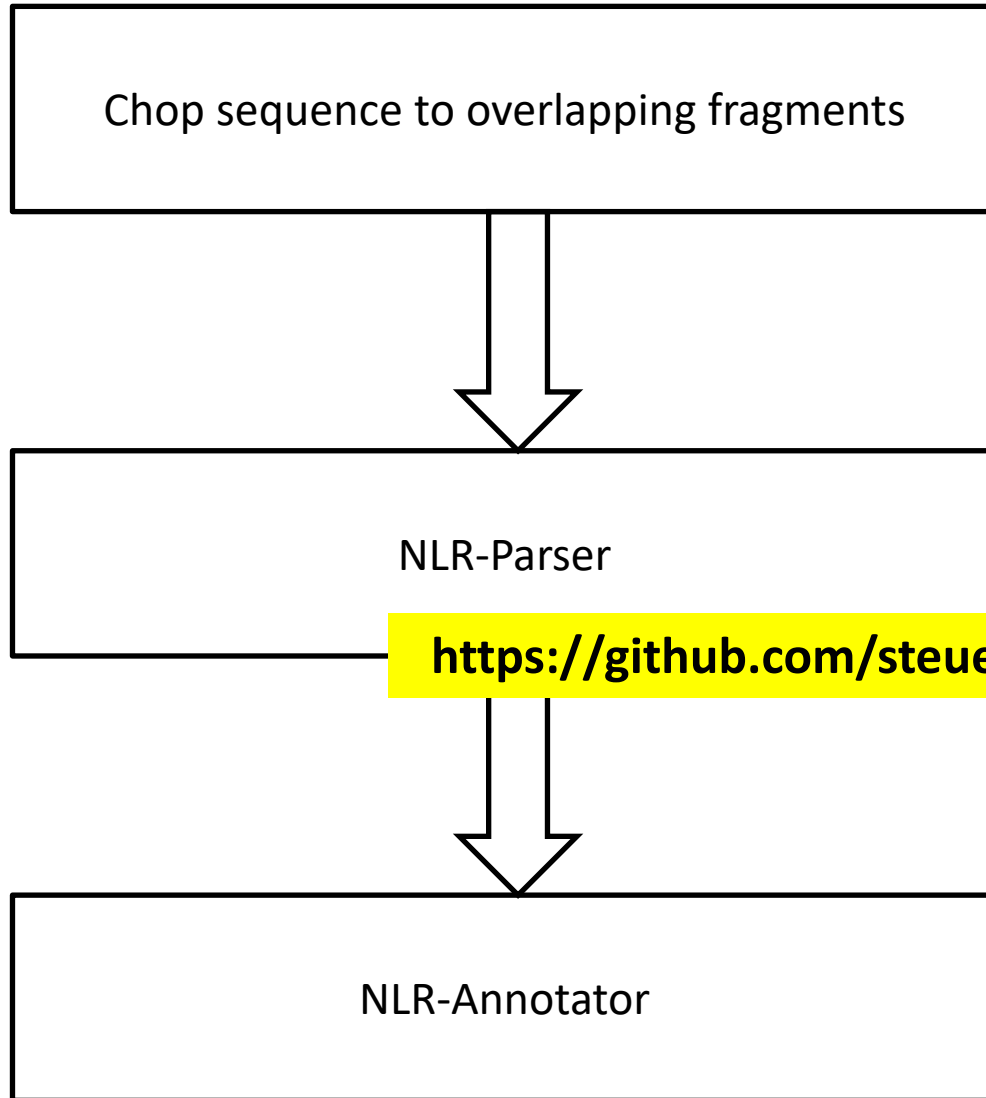
NLR-Annotator

- Annotation of NLR loci in genomic sequence
- An NLR locus is associated with a single NLR
- An NLR locus might be overlapping with a gene

https://github.com/steuernb/NLR-Annotator

NLR-Annotator on other genomes

Species	Source	NLRs
Coffee	http://coffee-genome.org	**932**
Maize	http://ensembl.gramene.org	**154**
Papaya	Phytozome	**50**
Cucumber	Phytozome	**71**
Soybean	Phytozome	**527**
Brachypodium	Phytozome	**342**

NLR-Annotator on wheat assemblies

Assembly	NLRs total	NLRs complete
IWGSC 2014	12441	1050
Meraculous	15315	1249
w2rap	3251	2465
Masurca	3568	2731
IWGSC refseq v1.0	3400	2580

Complete NLRs have at least 3 consecutive NB-ARC motifs,
the motif1/p-loop and at least one LRR.
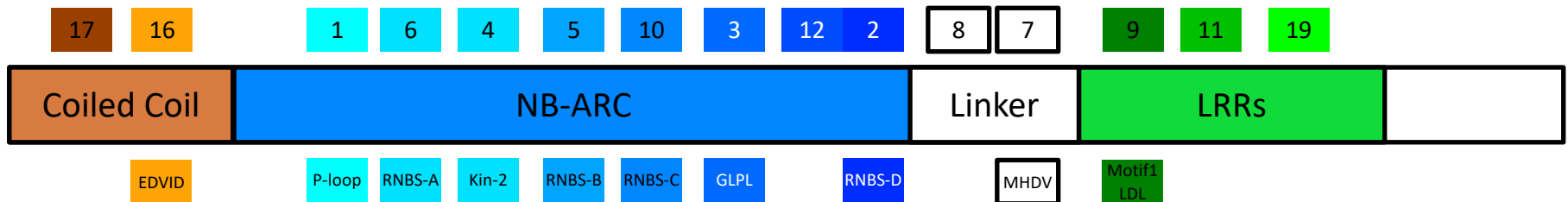
Integrating *R* gene markers with NLR loci

- Literature research for *Sr* and *Lr* gene

- Blast markers to refseq

- Evaluate distance and NLR loci in between

- Highlights
  - *Sr6*: 1NLR
  - *Lr13*: 3NLR
  - *Lr16*: 5NLR
  - *Lr49*: 3NLR
  - *Lr65*: 2NLR
  - *Lr70*: 5NLR

John Innes Centre
Unlocking Nature's Diversity

2Blades

Integrating *R* gene markers with NLR loci

NLR loci vs. genes

NLR-Annotator ⟷ IWGSC refseq annotation
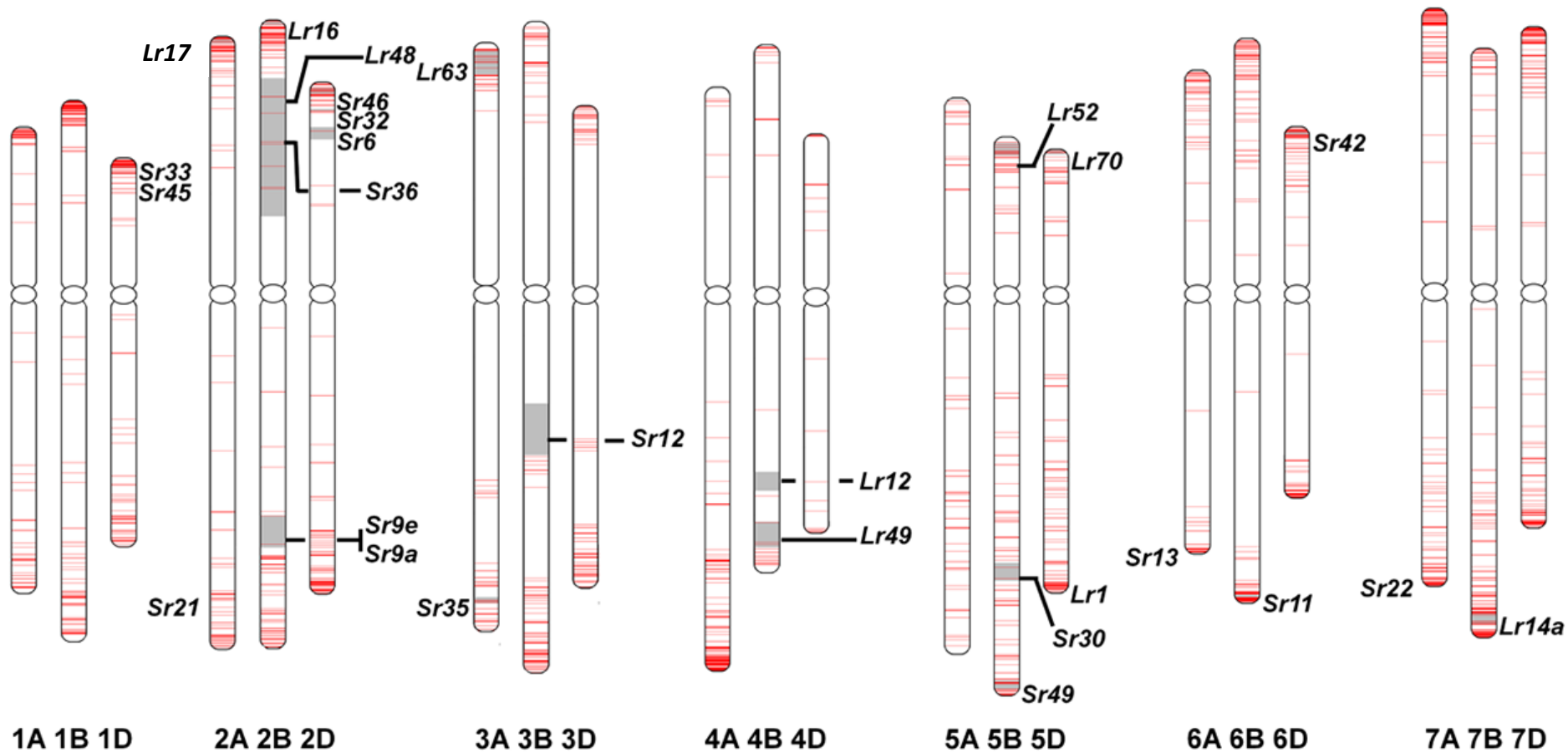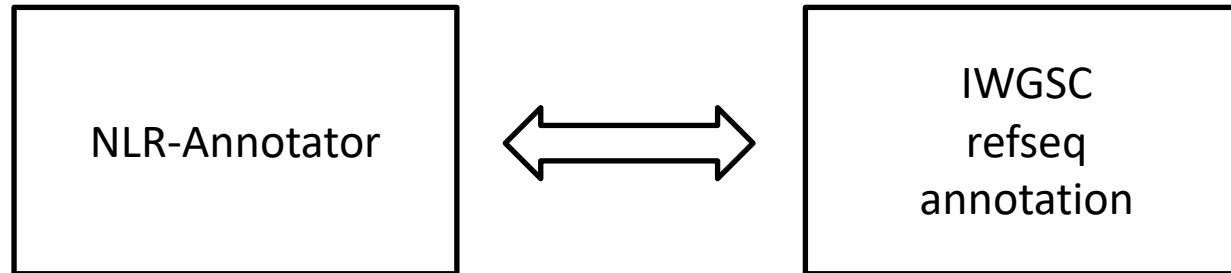
- Of 3,400 NLR loci 2,914 overlap with genes
- 632 of those overlap with more than one gene

- 3506 genes overlap with an NLR locus
- 198 genes overlap with more than one NLR locus

1 NLR overlapping with several genes

1 NLR overlapping with several genes
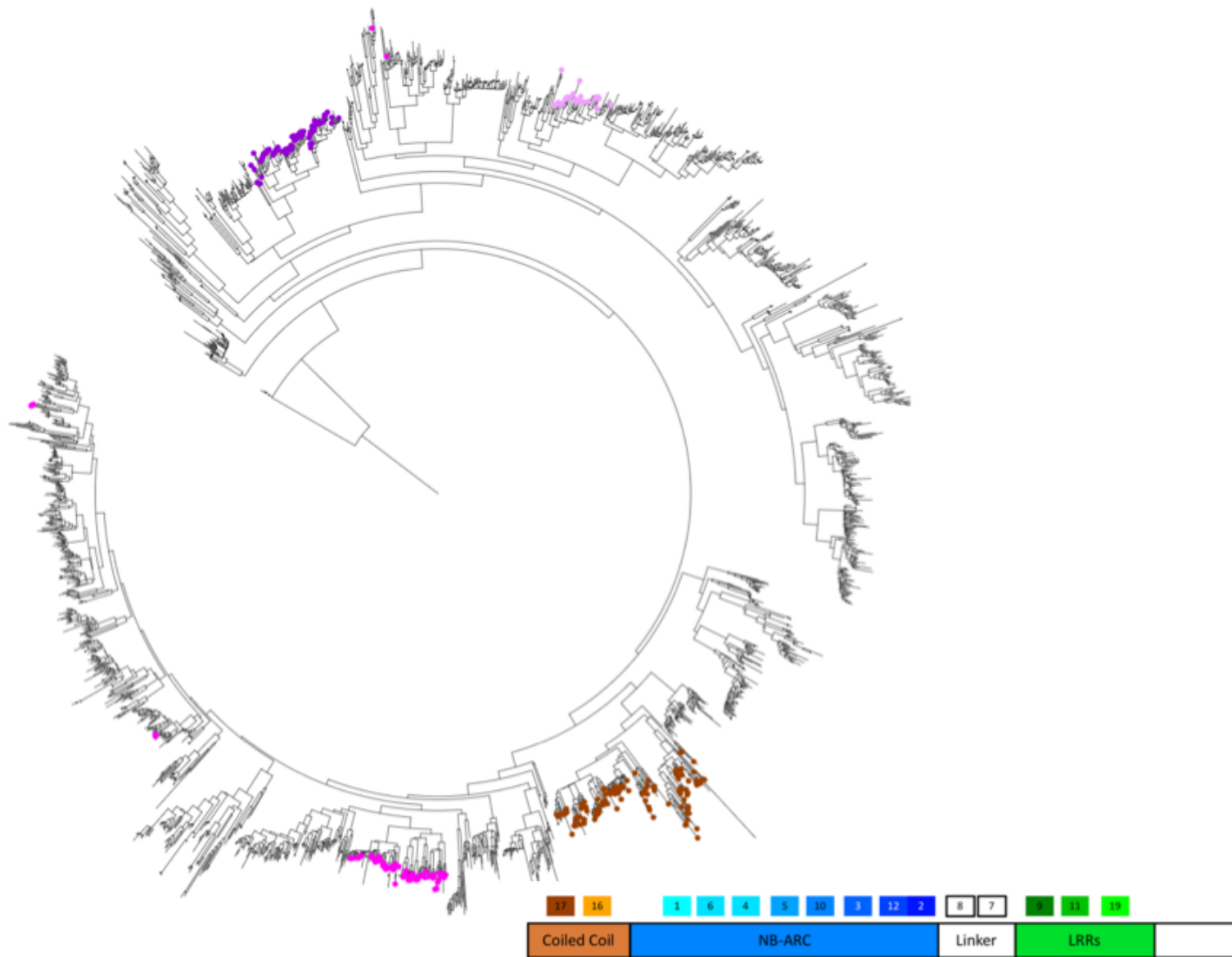

Pm2 homologue

NLR gene annotation in Chinese Spring

- The current annotation v. 1.0 is not perfect

- Manual annotation required

- Reference genomes are constantly being improved

- Too much effort for little gain

- We use available data for analysis, well knowing it is a subset

Phylogenetic Tree of NLR genes
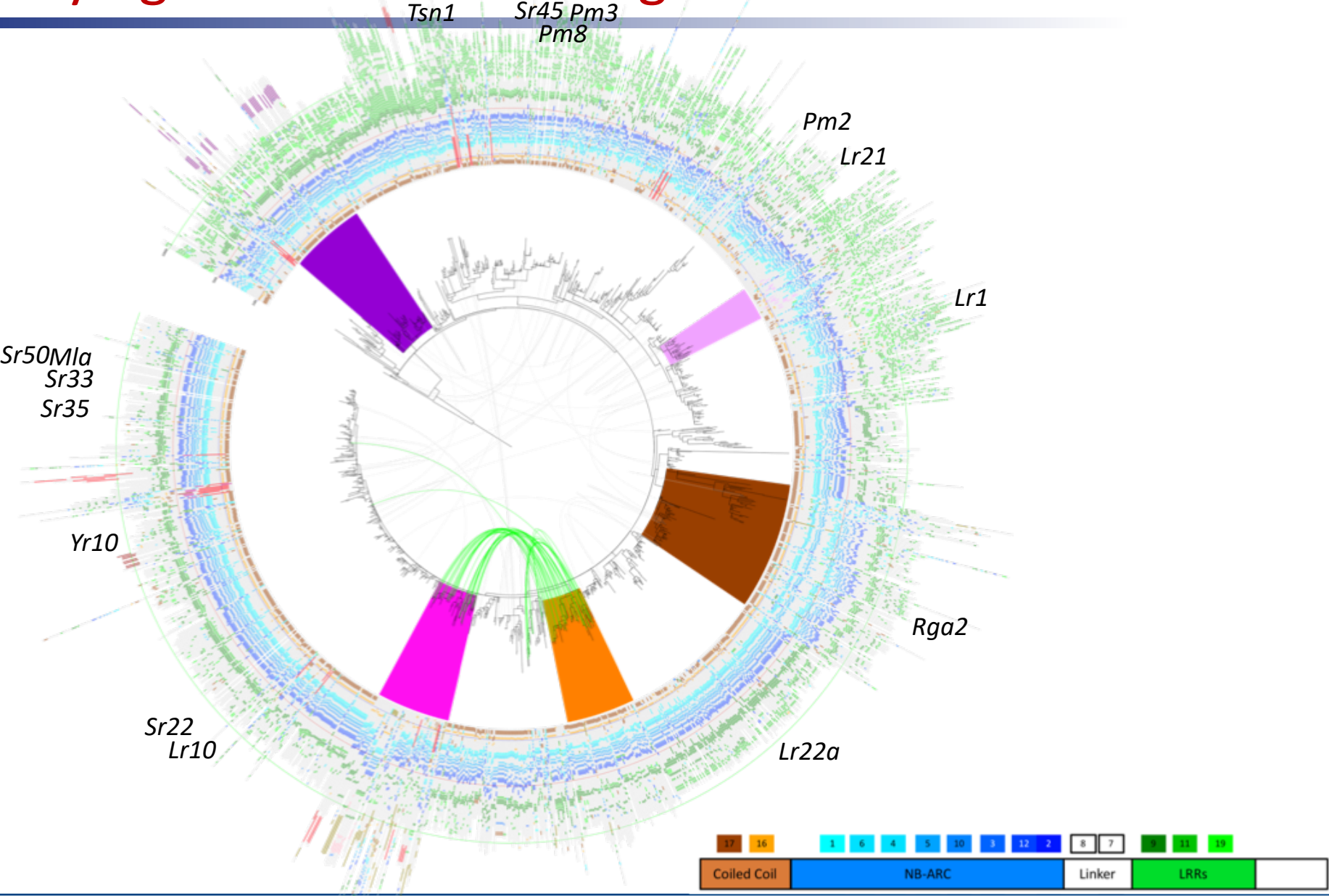
Phylogenetic studies on loci

Summary

- NLR-Annotator annotates NLRs in genome assemblies

- Candidate gene approaches

- Limitations in RefSeq v1.0 annotation

- Phylogenetic characterization

Acknowledgements

Brande Wulff
Kamil Witek
Jonathan Jones
Simon Krattinger
David Swarbreck
Ksenia Krasileva
Erin Baggs

2Blades

BBSRC
bioscience for the future

John Innes Centre
Unlocking Nature's Diversity

I W G S C