New Phytologist

Research

Identification of rye B chromosome-associated peptides by mass spectrometry

Wei Ma[1,2]*, ZhaoJun Liu[3,2,4]*, Sebastian Beier[2], Andreas Houben[2] and Sebastien Carpentier[5,6,7]

[1]Science Center for Future Foods, Jiangnan University, 1800 Lihu Road, Wuxi, Jiangsu 214122, China; [2]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstrasse 3 Stadt Seeland 06466, Germany; [3]Microelement Research Center/Key Laboratory of Arable Land Conservation (Middle and Lower Reaches of Yangtze River), Ministry of Agriculture, Huazhong Agricultural University, Wuhan 430070, China; [4]School of Life Sciences Life, Science Center Weihenstephan, Crop Physiology, Technical University Munich, Alte Akademie 12, Freising 85354, Germany; [5]Department of Biosystems, KU Leuven, Willem Decroylaan 42, 2455-3001 Leuven, Belgium; [6]SYBIOMA, KULeuven, Herestraat 49, Leuven 3000, Belgium; [7]Genetic Resources, Bioversity International, Willem Decroylaan 42, 2455-3001 Leuven, Belgium

Authors for correspondence:
Wei Ma
Email: mwei4518@hotmail.com

Sebastien Carpentier
Email: sebastien.carpentier@kuleuven.be

Received: *18 August 2020*
Accepted: *20 January 2021*

New Phytologist (2021) **230:** 2179–2185
doi: 10.1111/nph.17238

Key words: B chromosome, glycine-rich RNA binding protein, mass spectrometry, proteome, polymorphisms, *Secale cereale*.

Summary

- B chromosomes (Bs) are supernumerary dispensable components of the standard genome (A chromosomes, As) that have been found in many eukaryotes. So far, it is unkown whether the B-derived transcripts translate to proteins or if the host proteome is changed due to the presence of Bs.
- Comparative mass spectrometry was performed using the protein samples isolated from shoots of rye plants with and without Bs. We aimed to identify B-associated peptides and analyzed the effects of Bs on the total proteome.
- Our comparative proteome analysis demonstrates that the presence of rye Bs affects the total proteome including different biological function processes. We found 319 of 16 776 quantified features in at least three out of five +B plants but not in 0B plants; 31 of 319 features were identified as B-associated peptide features. According to our data mining, one B-specific protein fragment showed similarity to a glycine-rich RNA binding protein which differed from its A-paralogue by two amino acid insertions.
- Our result represents a milestone in B chromosome research, because this is the first report to demonstrate the existence of Bs changing the proteome of the host.

Introduction

B chromosomes (Bs) are supernumerary dispensable components of the standard genome (A chromosomes, As) that have been found in many eukaryotes, ranging from fungi to plants and animals including mammals (Camacho *et al.*, 2000; Jones *et al.*, 2008; Douglas & Birchler, 2017; Martins & Ahmad, 2019). They may vary in size, structure and chromatin properties among different species. Despite the diversity of Bs, they share some common features. They are not essential and are dispensable, they do not pair or recombine with As at meiosis, and they often show chromosome drive (Jones & Rees, 1982).

The presence of Bs leads to only mild or slight phenotypic changes if their number is low. Conversely, excessive numbers of Bs can cause phenotypic effects and may reduce the fertility and fitness of the host. The contributions of single Bs to these phenotypes are usually cumulative, with the severity of the effects increasing with the number of Bs (reviewed in Jones & Rees, 1982; Jones, 1995; Bougourd & Jones, 1997; Carlson, 2009). For example, the viability and fertility of rye plants are impaired when the number of Bs is high (reviewed in Jones & Puertas, 1993).

Although Bs are not essential, they are not necessarily inert. Bs were found to cause antibiotic resistance and pathogenicity in the fungus *Nectria haematococca* (Coleman *et al.*, 2009). It has been proposed that Bs accumulate mutations and structural rearrangements more rapidly over time than A chromosomes, and, as a result, carry genes relevant to host–pathogen interactions (reviewed in Bertazzoni *et al.*, 2018). In cichlids, Bs may play a role in the sex determination mechanism (Yoshida *et al.*, 2011). With the development of high throughput sequencing technologies, the rapid growth of genomic sequence data provides new insights into the organization of Bs and confirms that they carry genic sequences in a number of different species (reviewed in Banaei-Moghaddam *et al.*, 2015; Ruban *et al.*, 2017; Dalla *et al.*, 2019). In the cichlid *Astatotilapia latifasciata*, whilst the majority of the identified

*These authors contributed equally to this work.

B-genes are fragmented, some B-located genes were largely intact and three of them have been confirmed *in silico* (Valente *et al.*, 2014). The B-located *CAP-G* gene has been found in two Spanish populations of grasshoppers (Navarro-Dominguez *et al.*, 2017). In rye, it has been demonstrated that Bs contain *c.* 5000 putative B-located genic sequences (Martis *et al.*, 2012), three of the B-located genes have been confirmed intact and resembled the A-located paralogues experimentally (Ma *et al.*, 2017). With the emergence of next generation sequencing (NGS) technology, B chromosome-located genic sequences have been identified in many species. However, the number of intact genes remains low, as, due to the high fraction of repetitive elements of Bs, the *de novo* assembly from the short reads generated by NGS is very challenging.

Numerous B-located protein coding genes were discovered to contribute to the host transcriptome. In rye, thousands of B-derived transcripts with a reading frame were identified (Banaei-Moghaddam *et al.*, 2013; Ma *et al.*, 2017). The presence of B chromosome genes can alter the expression of A-located paralogue genes. For example, the presence of B chromosomes in the plants *Scillaautumnalis* (Rejón *et al.*, 1980; Oliver *et al.*, 1982) and *Allium schoennoprasum* (Plowman & Bougourd, 1994) have been shown to influence the expression of A-located genes for esterase and endosperm protein, respectively. Likewise, the presence of Bs can impact the expression of nuclear organizer regions (NORs) and the A chromosomes in the grasshopper *E. plorans* (Cabrero *et al.*, 1999). The Bs of *Crepis capillaris* encode transcriptionally active 45S rDNA variants (Leach *et al.*, 2005). In rye (Carchilan *et al.*, 2009; Banaei-Moghaddam *et al.*, 2013; Ma *et al.*, 2017), grasshopper (Navarro-Dominguez *et al.*, 2019) and maize (Huang *et al.*, 2016), pseudogene-like fragments on Bs were found to affect the transcription of A chromosome-located genes in a dosage-dependent manner. In addition, chromosome regions harbouring loci affecting B chromosome behaviour and transmission to progeny were identified (Carlson, 2007; Han *et al.*, 2007; Endo *et al.*, 2008). However, whether the B-derived transcripts finally translate to proteins *in vivo*, as well as the role they may play in the total proteome, are unknown.

Here, we report the comparative proteome analysis of rye plants with and without B chromosomes. We demonstrate that the presence of rye Bs affects the total proteome, including various biological function processes. We found 319 of 16 776 quantified peptide features in at least three of five +B plants but not in 0B plants; 31 of 319 peptide features were identified with confidence as B-associated peptide features (*P*-value < 0.05). According to our data mining, one B-specific protein fragment showed similarity to a glycine-rich RNA binding protein which differed from its presumed A-paralogue by two amino acid insertions. These data demonstrate that B chromosomes are able to contribute and influence the host proteome and likely translate their own transcripts to peptides. Hence, physiological effects, associated with the presence of Bs, may be explained partly by the presence of B-associated proteins.

Materials and Methods

Plant material and cultivation

Seeds of a rye (*Secale cereale* L.) self-fertile inbred line 7415 with (+) and without (0) B chromosomes (Bs) (Jimenez *et al.*, 1994), were germinated on wet filter papers. After germination, the seedlings were transferred to soil and cultivated for 20 d under long-day conditions (16 h : 8 h, light (22°C) : dark (16°C). Shoots of 0B and +B plants were harvested and stored at − 80°C for proteome analysis, each with five biological replicates.

Genotyping

To identify +B and 0B plant chromosomes, genomic DNA was extracted from rye leaves by a DNAeasy plant mini kit (Qiagen). Forward primer 5′-AGGAGGATCTTTTGTCCGCA-3′ and reverse primer 5′-ATCGTCAACCAGCACCAACT-3′ were used to amplify A- and B-derived *ScSHOC1* fragments (Ma *et al.*, 2017) in order to distinguish 0B and +B plants because of the single polymorphic sites. Twenty-five microlitres of PCR reaction mixture contained: 100 ng genomic DNA, 10 μM of each forward and reverse primers, 5 mM of each deoxynucleotide triphosphates, 2.5 μl 10× PCR reaction buffer and 1 unit of Taq polymerase (Qiagen). The cycling protocol was: 94°C for 3 min; 35 cycles at 94°C for 40 s, 57°C for 40 s, 2 min elongation at 72°C, and 72°C final elongation for 10 min. Then, the PCR products were checked for nonspecific amplification on a 1% agarose gel. The *Nsi*I enzyme for *ScSHOC1* was used for genotyping 0B and +B plants (Ma *et al.*, 2017). Twenty microlitres of the restriction enzyme digestion reaction included: 10 μl PCR mixture, 2 μl 10× buffer and 10 units of *Nsi*I (Thermo Fischer Scientific, Waltham, MA, USA). The reaction was incubated for 5 h at 37°C. The digestion products were checked after 2% agarose gel electrophoresis.

Protein extraction and data acquisition through mass spectrometry

Ground shoot samples were subjected to protein extraction according to the phenol extraction/ammonium acetate precipitation (Carpentier *et al.*, 2005) and adapted for gel free proteomics (Buts *et al.*, 2014). After protein quantification (2-D Quant Kit; GE Healthcare Europe, Freiburg, Germany), 20 μg of proteins were briefly digested with trypsin (Trypsin Protease, MS grade, Thermo Scientific, Merelbeke, Belgium) followed by purification with Pierce C18 Spin Columns (Thermo Scientific). The digested samples (0.5 μg/5 μl) were separated in an Ultimate 3000 UPLC system (Thermo Scientific) and injected in a Q Exactive Orbitrap mass spectrometer (Thermo Scientific) as described (van Wesemael *et al.*, 2018). The Q Exactive Orbitrap mass spectrometer (Thermo Scientific) was operated in positive ion mode with a nanospray voltage of 1.5 kV and a source temperature of 250°C. ProteoMass LTQ/FT-Hybrid ESI Pos. Mode Cal Mix (MS CAL5-1EA SUPELCO, Sigma-Aldrich) was used

as an external calibrant. The instrument was operated in data-dependent acquisition (DDA) mode with a survey MS scan at a resolution of 70 000 (fw hm at *m/z* 200) for the mass range of m/z 400–1600 for precursor ions, followed by MS/MS scans of the top 10 most intense peaks with +2, +3, +4, and +5 charged ions above a threshold ion count of 16 000 at 17 500 resolution using normalized collision energy of 25 eV with an isolation window of 3.0 *m/z* and dynamic exclusion of 10 s. All data was acquired with Xcalibur 3.0.63 software (Thermo Scientific). For protein identification, we used Mascot v.2.2.06 (Matrix Science) against rye 0B and +B genomic and RNA seq dataset (Martis *et al.*, 2012; Ma *et al.*, 2017), and the usual contaminants for MS (76 220 accessions). The parameters used for the identification were: parent mass tolerance of 10 ppm, fragment tolerance of 0.02 Da, oxidation of M as variable modification, carbamidomethyl C as fixed modification and up to one missed cleavage was allowed for trypsin. For identification and alignment, all raw data were converted into mgf-files using Progenesis (v.4.1; Nonlinear Dynamics, Newcastle, UK). To calculate the false discovery rate (FDR) in the identification and judge the protein inference, the sample-specific mgf files were loaded into Scaffold 3.6.5. The proteins were quantified in Progenesis based on all matching features. Statistical values indicated in Supporting Information Table S4 were calculated by Progenesis. Protein quantifications were exported from Progenesis and subjected to additional statistical analysis (Statistica 13; Tibco Software Inc., Palo Alto, CA, USA). Protein abundances were investigated in relation to the genotype via the PLS-DA NIPALS algorithm and *t*-test (Statistica 13).

Gene ontology term assignment

Sequences identified as being significantly different in 0B and +B plants were further characterized using gene ontology (GO) terms (Ashburner *et al.*, 2000) (Table S1). For the purpose of this assignment InterProScan5 (Jones *et al.*, 2014) (v.5.26-65.0) was applied to the dataset. Fasta files were split before the analysis into chunks of 250 sequences each with shell commands. Sequences were then scanned by InterProScan5 in parallel with GNU parallel (Tange, 2018) against these applications: Tigrfam (Haft *et al.*, 2012), Pirsf (Wu *et al.*, 2004), ProDom (Bru *et al.*, 2005), Smart (Letunic *et al.*, 2012), Prosite Profiles (Sigrist *et al.*, 2012), Prosite Patterns (Sigrist *et al.*, 2012), PfamA (Bateman *et al.*, 2002), Prints (Attwood *et al.*, 2012), Superfamily (de Lima Morais *et al.*, 2010). GO term assignments were extracted and consolidated between these applications with shell commands.

Results and Discussion

Presence of B chromosomes influences the host proteome

In order to evaluate whether the presence of B chromosomes influences the host proteome, a comparative MS analysis was performed using the protein samples isolated from shoots of five individual rye plants, with and without Bs, at the same developmental stage (20 d after germination). A total of 16 776 peptide features were generated from the quantitative proteomics analysis from rye 0B and +B plants (Table S2). According to the partial least squares (PLS) analysis, the proteome was significantly correlated to the presence/absence of Bs (Fig. 1a; Table S3), indicating the effect of B chromosome-associated peptide features for the shoot proteome. To further determine which biological processes were affected by the presence of Bs, we analyzed the differentially abundant proteins in 0B and +B plants. Protein sequences were scanned with InterProScan5 to detect conserved domains. Also, the GO terms were extracted (Table S1). Proteins with shared ontologies were visualized in Cytoscape. Manual searching of interaction maps for the proteins revealed nine interesting biological processes, which are illustrated in Fig. 1(b–j).

The annotation of differential proteins (Fig. 1b–j) is in agreement with our previous GO enrichment analysis for B chromosome-enriched transcripts (Ma *et al.*, 2017). The identified proteins involved in photosynthetic, glycolytic and oxidation processes, also are in line with our previous observation that transcriptionally active B chromosomes contain DNA fragments that were derived from chloroplasts and mitochondria (Martis *et al.*, 2012; Ruban *et al.*, 2014; Ma *et al.*, 2017). One example we identified in this study was a 'negative' regulator of metabolic processes (carbon fixation), from the protein encoded by LN626639, indicating the possible causes of detrimental effects of too many Bs and fitting with the reduced transcription in +B plants observed previously (Ma *et al.*, 2017). The upregulated nucleosome assembly proteins observed in +B plants point towards a role of Bs in chromatin assembly. In short, Bs have an influence on the proteome and are involved in different biological function processes, such as photosynthesis, glycolysis, oxidation, and metabolic, ribosomal and nucleosome assembly.

The impact of the proteins only detected in +B plants

In order to identify proteins which were encoded only by the B chromosome or induced by the presence of Bs, we searched for peptides that were detected with confidence in +B plants only. There are 30 peptide features that have been detected in all five +B samples and never in 0B samples. This pattern is very significant according to the Fisher's exact test ($P < 0.01$ after the Bonferroni correction for multiple testing) (Dunn, 1961). Three hundred and nineteen peptide features were detected in at least three of five +B plants and not detected in 0B plants; of these 319, 31 features returned a confident hit against our database and had a P-value < 0.05 (Table S4). The number of identified B-associated peptide features was very low for the following reasons: (1) the sequence information about B-located genes is very limited; (2) B-derived peptides could share the same sequences as their A-derived paralogues; and (3) the gene annotation of the rye genome is currently modest.

Our previous studies showed that A- and B-derived paralogous transcripts show polymorphisms, including deletions, insertions and single nucleotide polymorphisms (SNPs) (Banaei-Moghaddam *et al.*, 2013; Ma *et al.*, 2017). Therefore, we assumed that A-

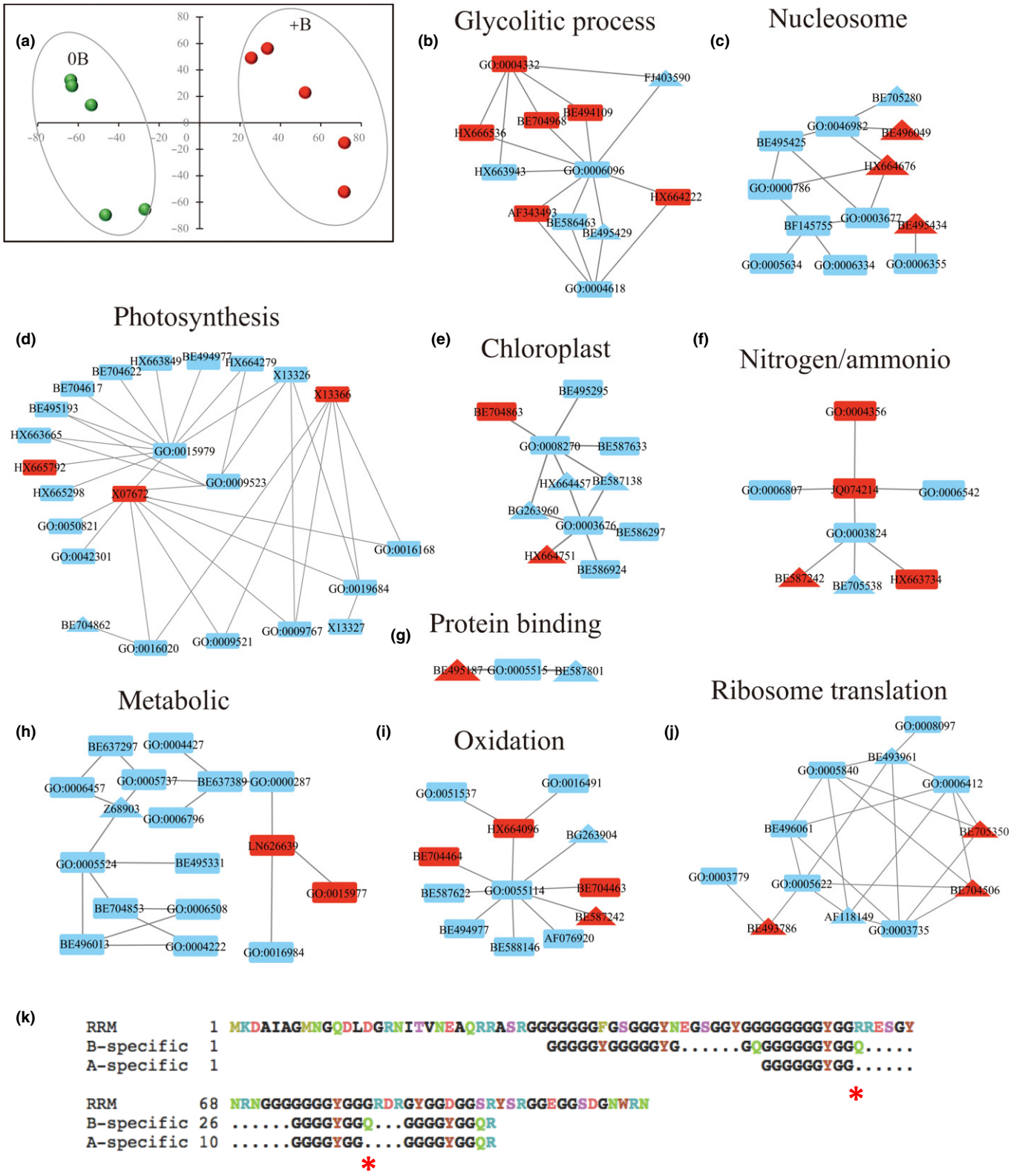

Fig. 1 Comparative proteome analysis for rye 0B and +B plants. (a) Partial least squares (PLS) analysis of proteome profiles of +B (red dot) and 0B (green dot) plants. Five plants of each type were used. (b–j) *t*-test significant proteins were BLASTed and annotated to the gene ontology, and proteins with shared ontologies were scanned with INTERPROSCAN 5 to assign gene ontology (GO) terms and proteins visualized in CYTOSCAPE. The up- and downregulated proteins in the presence of Bs were indicated in red triangle and rectangle, respectively. (k) Alignment of the RNA recognition motif (RRM) superfamily in *Glycine max* (soybean) with rye A- and B-specific peptides. Red asterisks mark the polymorphic sites between A- and B-specific peptides.

and B-derived paralogous peptide features could be distinguished by polymorphic sites. Among these 31 B-encoded peptide features, one peptide with the sequence GGGGGYGGGGGYGGQ GGGGGGYGGQGGGGYGGQGGGGYGGQR (BE587164|BE587164.1_1; Table S4) showed homology to the peptide feature with sequence GGGGGGYGGGGGGGYGGGGGG YGGQR (BE586297|BE586297.1_7; Table S2) appearing in both 0B and +B samples (Fig. 1k). Comparison of these two peptides revealed two insertions (Fig. 1k). The functional annotation of both peptide features is a glycine-rich RNA binding protein (Fig. 1k). From our proteome analysis, we identified five additional proteins (BE494459, BE586297, BE587164, HX663568 and HX665852) with peptides having glycine-rich regions (Table S2). Three of them are significantly more abundant in +B plants ($P < 0.1$: BE587164, HX663568 and HX665852).

What are the functional consequences of an additional glycine-rich RNA binding protein being expressed by the B chromosome of rye? RNA-binding proteins with an RNA recognition motif (RRM)-type domain play an important role in developmental and environmental responses (Gomez-Porras *et al.*, 2011). Therefore, physiological effects associated with the presence of Bs may be explained partially by the activity of B-derived RNA binding proteins. Also, essential processes of gene expression at the post-transcriptional level are regulated mainly by proteins containing well-defined sequence motifs involved in RNA binding. The RNA recognition domain is one of the most widely spread motifs (Lorković & Barta, 2002). As with an increasing number of Bs, the total transcripts encoded by rye As and Bs do not change in a linear manner (Ma *et al.*, 2017). We hypothesize that an additional B chromosome-encoded RNA recognition domain gene and a greater abundance of several paralogues are involved in the regulation of gene expression in a dosage-sensitive manner.

In summary, the most noteworthy result of our analysis is that the existence of B chromosomes can alter the composition of the rye proteome. Moreover, *c.* 1.9% (319 of 16 776) of the quantified features were found in at least three of five +B plants but not in 0B plants; 31 of 319 features returned a hit against our database and were considered as B-associated peptide features. This is likely due to the translation of the B-derived transcripts or due to the presence of Bs inducing the expression of A-paralogues. We speculate that B-specific proteins might have the same function as their A-paralogues. It is likely that B-encoded peptides behave like proteins derived from duplicated or pseudogenized genes. In the future, a sequence annotated rye B chromosome will help us to confirm more B chromosome-derived proteins and to better understand the biological and evolutionary implications of B chromosomes.

Acknowledgements

This work was supported by the China CSC scholarship, DFG Germany (grant no. HO 1779/30-1) and European Cooperation in Science and Technology (COST) Action FA1306 (grant no. COST-STSM-ECOST-STSM-FA1306-230117-082320), National Natural Science Foundation of China (grant no. 32002122) and China Postdoctoral Science Foundation (grant no. 2020M682444). We thank Patty Liao and Sarah Houben for the linguistic assistance during the preparation of this manuscript. We thank Yong Jiang and Armin Meister for the mathematical analysis during the preparation of this manuscript.

Author contributions

WM, ZJL and AH designed the research; WM, ZJL and SC performed experiments; WM, ZJL, SB and SC analyzed the data; and WM, ZJL, SB, SC and AH wrote the article. WM and ZJL contributed equally to this work.

ORCID

Sebastian Beier https://orcid.org/0000-0002-2177-8781
Sebastien Carpentier https://orcid.org/0000-0002-7389-6405
Andreas Houben https://orcid.org/0000-0003-3419-239X
ZhaoJun Liu https://orcid.org/0000-0003-3686-9245
Wei Ma https://orcid.org/0000-0001-7089-7093

Data availability

Raw sequencing data from RNA-seq including 'Anther + B transcripts' (Anth_4B_Trinity), 'Root & shoot + B transcripts' (TrSc_4B) and '0B transcripts' (Trinity_Assembly_0B) are available at EMBL/ENA (accession no. PRJEB12520) (Ma *et al.*, 2017). The '+B WGS contigs' sequences are available in the European Nucleotide Archive database, http://www.ebi.ac.uk/ena/ (accession no. ERP001061) (Martis *et al.*, 2012). Sequence assemblies are available as BLAST databases under https://webblast.ipk-gatersleben.de/ryeselect/.

References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT *et al.* 2000. Gene ontology: tool for the unification of biology. *Nature Genetics* **25**: 25–29.

Attwood TK, Coletta A, Muirhead G, Pavlopoulou A, Philippou PB, Popov I, Romá-Mateo C, Theodosiou A, Mitchell AL. 2012. *The PRINTS database: a fine-grained protein sequence annotation and analysis resource—its status in 2012.* (Database). [WWW document] URL www.bioinf.manchester.ac.uk/dbbrowser/PRINTS/.

Banaei-Moghaddam AM, Martis MM, Macas J, Gundlach H, Himmelbach A, Altschmied L, Mayer KF, Houben A. 2015. Genes on B chromosomes: Old questions revisited with new tools. *Biochimica et Biophysica Acta–Gene Regulatory Mechanisms* **1849**: 64–70.

Banaei-Moghaddam AM, Meier K, Karimi-Ashtiyani R, Houben A. 2013. Formation and expression of pseudogenes on the B chromosome of rye. *Plant Cell* **25**: 2536–2544.

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer EL. 2002. The Pfam protein families database. *Nucleic Acids Research* **30**: 276–280.

Bertazzoni S, Williams AH, Jones DA, Syme RA, Tan KC, Hane JK. 2018. Accessories make the outfit: accessory chromosomes and other dispensable DNA regions in plant-pathogenic fungi. *Molecular Plant–Microbe Interactions* **31**: 779–788.

Bougourd SM, Jones RN. 1997. B chromosomes: a physiological enigma. *New Phytologist* **137**: 43–54.

Bru C, Courcelle E, Carrère S, Beausse Y, Dalmar S, Kahn D. 2005. The ProDom database of protein domain families: more emphasis on 3D. *Nucleic Acids Research* **33**: D212–D215.

Buts K, Michielssens S, Hertog M, Hayakawa E, Cordewener J, America A, Nicolai B, Carpentier S. 2014. Improving the identification rate of data independent label-free quantitative proteomics experiments on non-model crops: a case study on apple fruit. *Journal of Proteomics* **105**: 31–45.

Cabrero J, López-León MD, Bakkali M, Camacho JPM. 1999. Common origin of B chromosome variants in the grasshopper *Eyprepocnemis plorans*. *Heredity* **83**: 435.

Camacho JPM, Sharbel TF, Beukeboom LW. 2000. B-chromosome evolution. *Philosophical Transactions: Biological Sciences* **355**: 163–178.

Carchilan M, Kumke K, Mikolajewski S, Houben A. 2009. Rye B chromosomes are weakly transcribed and might alter the transcriptional activity of A chromosome sequences. *Chromosoma* **118**: 607–616.

Carlson W. 2009. The B chromosome of maize. In: Bennetzen JL, Hake S, eds. *Handbook of maize*. New York, NY, USA: Springer, 459–480.

Carlson WR. 2007. Locating a site on the maize B chromosome that controls preferential fertilization. *Genome* **50**: 578–587.

Carpentier SC, Witters E, Laukens K, Deckers P, Swennen R, Panis B. 2005. Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics* **5**: 2497–2507.

Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S *et al.* 2009. The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. *PLoS Genetics* **5**: e1000618.

Dalla BE, Akbari OS, Ferree PM. 2019. Sequence expression of supernumerary B chromosomes: function or fluff? *Genes* **10**: 123.

Douglas RN, James AB. 2017. B chromosomes. In: Bhat TA, Wani AA, eds. *Chromosome structure and aberrations*. New York, NY, USA: Springer, 13–39.

Dunn OJ. 1961. Multiple comparisons among means. *Journal of the American Statistical Association* **56**: 52–64.

Endo TR, Nasuda S, Jones N, Dou Q, Akahori A, Wakimoto M, Tanaka H, Niwa K, Tsujimoto H. 2008. Dissection of rye B chromosomes, and nondisjunction properties of the dissected segments in a common wheat background. *Genes & Genetic Systems* **83**: 23–30.

Gomez-Porras J, Lewinski M, Riaño-Pachón DM, Staiger D. 2011. Molecular evolution of RRM-containing proteins and glycine-rich RNA-binding proteins in plants. *Nature Precedings* **1**: 1.

Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E. 2012. TIGRFAMs and genome properties in 2013. *Nucleic Acids Research* **41**: D387–D395.

Han F, Lamb JC, Yu W, Gao Z, Birchler JA. 2007. Centromere function and. nondisjunction are independent components of the maize B chromosome accumulation mechanism. *Plant Cell* **19**: 524–533.

Huang W, Du Y, Zhao X, Jin W. 2016. B chromosome contains active genes and impacts the transcription of A chromosomes in maize (*Zea mays* L.). *BMC Plant Biology* **16**: 1.

Jimenez M, Romera E, Puertas M, Jones R. 1994. B-chromosomes in inbred lines of rye (*Secale cereale* L.). *Genetica* **92**: 149–154.

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G *et al.* 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**: 1236–1240.

Jones RN. 1995. Tansley review no 85. B chromosomes in plants. *New Phytologist* **131**: 411–434.

Jones RN, Puertas MJ. 1993. The B-chromosomes of rye (*Secale cereale* L.). In: Dhir KK, Sareen TS, eds. *Frontiers in plant science research*. Delhi: Bhagwati Enterprises, 81–112.

Jones RN, Rees H. 1982. *B chromosomes, 1st edn.* London, UK: Academic Press.

Jones RN, Viegas W, Houben A. 2008. A century of B chromosomes in plants: so what? *Annals of Botany* **101**: 767–775.

Leach CR, Houben A, Field B, Pistrick K, Demidov D, Timmis JN. 2005. Molecular evidence for transcription of genes on a B chromosome in *Crepis capillaris*. *Genetics* **171**: 269–278.

Letunic I, Doerks T, Bork P. 2012. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Research* **40**: D302–D305.

de Lima Morais DA, Fang H, Rackham OJ, Wilson D, Pethica R, Chothia C, Gough J. 2010. SUPERFAMILY 1.75 including a domain-centric gene ontology method. *Nucleic Acids Research* **39**: D427–D434.

Lorković ZJ, Barta A. 2002. Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. *Nucleic Acids Research* **30**: 623–635.

Ma W, Gabriel TS, Martis MM, Gursinsky T, Schubert V, Vrána J, Doležel J, Grundlach H, Altschmied L, Scholz U *et al.* 2017. Rye B chromosomes encode a functional Argonaute-like protein with *in vitro* slicer activities similar to its A chromosome paralog. *New Phytologist* **213**: 916–928.

Martins C, Ahmad SF. 2019. The modern view of B chromosomes under the impact of high scale omics analyses. *Cells* **8**: 156.

Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, Scholz U, Gundlach H, Wicker T, Simkova H *et al.* 2012. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences, USA* **109**: 13343–13346.

Navarro-Domínguez B, Martín-Peciña M, Ruiz-Ruano FJ, Cabrero J, Corral JM, López-León MD, Sharbel TF, Camacho JPM. 2019. Gene expression changes elicited by a parasitic B chromosome in the grasshopper *Eyprepocnemis plorans* are consistent with its phenotypic effects. *Chromosoma* **128**: 53–67.

Navarro-Domínguez B, Ruiz-Ruano FJ, Camacho JPM, Carbrero J, López-León MD. 2017. Transcription of a B chromosome *CAP-G* pseudogene does not influence normal condensin complex genes in a grasshopper. *Scientific Reports* **7**: 17650.

Oliver JL, Posse F, Martinez-Zapater JM, Enriquez AM, Ruiz-Rejon M. 1982. B-chromosomes and E-1 isozyme activity in mosaic bulbs of *Scilla autumnalis* (Liliaceae). *Chromosoma* **85**: 399–403.

Plowman AB, Bougourd SM. 1994. Selectively advantageous effects of B chromosomes on germination behaviour in *Allium schoenoprasum* L. *Heredity* **72**: 587.

Rejón MR, Posse F, Oliver JL. 1980. The B chromosome system of *Scilla autumnalis* (Liliaceae): effects at the isozyme level. *Chromosoma* **79**: 341–348.

Ruban A, Fuchs J, Marques A, Schubert V, Soloviev A, Raskina O, Badaeva E, Houben A. 2014. B chromosomes of *Aegilops speltoides* are enriched in organelle genome-derived sequences. *PLoS ONE* **9**: e90214.

Ruban A, Schmutzer T, Scholz U, Houben A. 2017. How next-generation sequencing has aided our understanding of the sequence composition and origin of B chromosomes. *Genes* **8**: 294.

Sigrist CJ, De Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L, Xenarios I. 2012. New and continuing developments at PROSITE. *Nucleic Acids Research* **41**: D344–D347.

Tange O. 2018 GNU parallel 2018, March 2018. doi: 10.5281/zenodo.1146014

Valente GT, Conte MA, Fantinatti BE, Cabral-de-Mello DC, Carvalho RF, Vicari MR, Kocher TD, Martins C. 2014. Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Molecular Biology and Evolution* **31**: 2061–2072.

van Wesemael J, Hueber Y, Kissel E, Campos N, Swennen R, Carpentier S. 2018. Homeolog expression analysis in an allotriploid non-model crop via integration of transcriptomics and proteomics. *Scientific Reports* **8**: 1353.

Wu CH, Nikolskaya A, Huang H, Yeh LSL, Natale DA, Vinayaka CR, Hu ZZ, Mazumder R, Kumar S, Kourtesis P *et al.* 2004. PIRSF: family classification system at the protein information resource. *Nucleic Acids Research* **32**: D112–D114.

Yoshida K, Terai Y, Mizoiri S, Aibara M, Nishihara H, Watanabe M, Kuroiwa A, Hirai H, Hirai Y, Matsuda Y *et al.* 2011. B chromosomes have a functional effect on female sex determination in lake victoria cichlid fishes. *PLoS Genetics* **7**: e1002203.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Table S1 List of peptide features with shared ontologies which showed different expression in 0B and +B plants.

Table S2 List of generated spectrums from the quantitative proteomics analysis from rye 0B and +B plants.

Table S3 Proteome profiles of 0B and +B plants for the PLS analysis.

Table S4 List of peptide features were found in at least three of five +B plants but not detected in 0B plants.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.