MEICOM Marie Curie ITN 2018 ESR Progress Summary

ESR Name: Chiara Di Dio	Supervisor: Dr. James Higgins

Workpackage Title:

Investigating synaptonemal complex morphogenesis in polyploid wheat

Research aims and progress for the period:

Background

Nowadays factors such as population growth, urban development, mass migration and climate change necessitate sustained improvements in food production, storage and distribution. As regards, crop breeding might play a rising role in terms of food security and sustainability.

Bread wheat (*Triticum aestivum*) is the largest crop in the UK, grown over 2M hectares, adding over £1.6Bn to the UK economy, with a value almost ten times that for processed wheat-derived products (http://esa.un.org/unpd/wpp/Documentation/publications.htm). It accounts for 20 per cent of the calories and protein consumed by humans and is also an important source of vitamins and micronutrients. Despite substantial increases during the green revolution, yields are now susceptible to decline due to extreme weather patterns. By 2050, the world's population is expected to rise to 8.9 billion leading to considerable pressure on resources, such as water availability, food and affordable fuel (https://www.ipcc.ch/publications). In addition, since 1950, the planet has been warming by 0.13°C per decade and this is expected to rise (to 0.2°C per decade) over the next two to three decades (Lobell *et al*. 2011). This will have a direct effect on agriculture such that a 1°C rise is expected to reduce yields by ~10% Lukaszewski *et al.* 2012). With an ever increasing demand for wheat by a growing world population and unpredictable abiotic factors, there is a pressing need to improve wheat varieties by utilizing biological knowledge to assist traditional plant breeding.

Currently, wheat breeding methods are primarily reliant on meiotic recombination to generate genetic variation through the formation of crossovers (COs) that produce new combinations of genes. Specifically, in wheat meiotic COs are restricted in number (1-3 per chromosome pair) and skewed in distribution to regions near the ends of chromosomes Higgins *et al.,* 2012). As a result an estimated 30-50% of genes reside in recombination 'cold' regions. This creates the problem of linkage-drag in the cold centromere-proximal and interstitial regions where undesirable variation cannot be separated from useful traits. In effect, these regions become inaccessible to researchers for genetic mapping of agronomical important traits. Thus, it is desirable to manipulate the frequency and distribution of COs in bread wheat to generate novel allelic combinations. Once the useful traits are recombined onto the same chromosomes it will also be advantageous to 'fix' these desirable traits by reducing the frequency of COs. In addition, such tools may be used to assist classical mapping by segregation of polymorphic markers associated with important agronomical traits. Hence, the challenge is to determine the factors that regulate CO frequency and localization in wheat and to devise strategies to overcome this problem.

Research strategy

The synaptonemal complex (SC) is a meiosis specific protein structure that is essential for the formation of wild-type levels of crossovers (COs) in sexually reproducing eukaryotes. It is composed of two axial elements and a central element. The axial elements comprise of cohesin proteins including SMC1, SMC3, REC8 and PDS5 as well as meiosis specific proteins ASY1 and ASY3. The only known protein in the central element in plants is ZYP1. Current evidence suggests that polymorphisms in SC genes lead to changes in the recombination landscape. CO formation is highly skewed towards the chromosome ends in bread wheat, thus limiting generation of novel allelic combinations during plant breeding. There is a major effort to try and unlock the 'cold' genomic regions of wheat so that the natural diversity can be exploited. In addition the number of COs per chromosome pair is low (usually 1-3) independent of chromosome size. Based on SC sequence conservation of functional domains, we hypothesise that organisms with larger chromosomes (and polyploids) have weaker alleles to maintain low levels of COs.

At this basis, the aim of our proposal is to identify and modify the factors controlling CO frequency and distribution during bread wheat meiosis via complementary genetic and cytogenetic approaches. This knowledge will also be used to manipulate CO formation using synthetic approaches and targeting recombination. We have recently shown that the skewed CO distribution in barley is due to a spatio-temporal asymmetry of meiotic progression (Higgins *et al.,* 2012, 2014). Although barley and bread wheat are both members of the Poaceae family, hexaploid bread wheat has experienced two polyploidization events and therefore may have evolved different strategies to maintain accurate CO control, necessitating direct study in this species. By achieving our objectives we will make a significant impact to the development and use of high throughput genotyping and phenotyping platforms and in doing so we will provide public and private breeders with the tools and technologies required to characterise and exploit genetic diversity to create novel wheat varieties.

Aims and objectives

Purpose of my project is to analyse mechanisms underlying chromosome pairing in polyploid wheat, focusing on the function of three meiotic genes: *ASY1, ASY3,* and *ZYP1*.

First, we propose to assess the expression of *ASY1, ASY3,* and *ZYP1* by determining homoeologous expression level and investigating regulatory sequences and splice variants of these genes.

Secondly, to further understand the role of these genes I will investigate their function using mutant TILLING lines for *ASY1, ASY3 and ZYP1* with a cytogenetic approach.

To obtain a complete scheme about the mechanism of chromosome pairing, I am planning to determine spatiotemporal phosphorylation patterns of proteins through mass spectrometry technique as a part of my secondment in Birmingham.

Secondments

Three secondments have been planned throughout the PhD: two-months-training to isolate the synaptonemal complex proteins from wheat meiocytes at University of Birmingham (UK) during the first year; one-month-training in super resolution microscopy techniques (SIM & STORM) at IPK (Germany) over the second year, and one-month-training regarding plant breeding techniques at KWS UK Ltd (UK) on third year.

Experimental design

In order to evaluate the level of gene expression, it has been necessary to determine the contribution of each homeologous gene. First of all, I tested whether all homeologs were expressed.

I selected two different hexaploid wheat cultivars [*Triticum aestivum* cv. Apogee and Cadenza (2n=6x=42)], grown under 16h days light cycle at constant temperature of 22 °C in greenhouse. Total RNA of wheat cv. Apogee and Cadenza was extracted from first leaf Zadoks' stage 11 (Zadoks *et al.,* 1974) leaves using Isolate II RNA Plant kit (Bioline). First-strand cDNA was synthesized using Tetro cDNA Synthesis Kit (Bioline). Gene specific primers with Q5 High-Fidelity DNA Polymerase (New England Biolabs) were employed to amplify the coding region of target genes. Lastly, PCR fragments were run onto 1% agarose gel, which have shown that all of homeologs were expressed.

In order to know in which proportion such meiotic genes were expressed, PCR amplicons were ligated into pDrive vector (Qiagen) and DH5α competent *E. coli* was selected as a strain for cloning the insert of interest. Recombinant clones were detected with blue-white screening, a common technique based on the principle of α-complementation of the β-galactosidase gene. Plasmid DNA was extracted through E.Z.N.A. Plasmid DNA Mini Kit (Omega Bio-Tec) and sequenced from Eurofins GATC Biotech Company. The full-length cDNA sequences were analysed using Ensembl Plant database (Kersey *et al.,* 2018) and Omega Clustal W DNA Multiple alignment (Sievers *et al.,* 2011).

Wheat genome assembly hosted on EnsemblPlants (http://plants.ensembl.org) is IWGSC (International Wheat Genome Sequencing Consortium) RefSeqv1.0, which uses the variety Chinese Spring.

To find the wheat orthologs of genes from other species on EnsemblPlant database, I used BLAST to query my sequence clone against *Triticum aestivum*. I downloaded the reference nucleotide sequences (Chinese Spring) and aligned them with my clones. I identified which subgenome each clone belongs to with Clustal W. It generates Neighbour-joining phylogentic trees as output from a multiple sequence alignment, using FASTA format as input.

Table 1. Cloned homeologs (A, B, D) to determine gene expression based on clones (%)

Cloned gene	Chromosome	A	B	D
ASY1	5	89 (n=17)	5 (n=1)	5 (n=1)
ASY3	2	0	100(n=7)	0
ASY3	5	14 (n=4)	42 (n=12)	14 (n=4)
ZYP1	2	75 (n=15)	20 (n=4)	5 (n=1)

Figure 1. Wheat RNA-Seq expression, from Wheat expVIP database

Table 1 shows the total clone number and relative proportion of homeologs per gene expressed in percentage (%). The outcomes have shown differential expression level among the genes of interest. Indeed, for *ASY1* and *ZYP1*, most of transcript derives from A subgenome; regards *ASY3*, I identified more than one copy per subgenome, one on chromosome 5 and another on chromosome 2. Overall, the most highly expressed was the 5B, followed by 2B, while I have not found any 2A and 2D clones (we hypothesised they are weakly expressed). At this point, I compared these data with Wheat expVIP database (http://www.wheat-expression.com/) (Figure 1) to test if there is any correspondence with my results. The outcomes were consistent only for *ASY1* and *ZYP1,* but not for *ASY3*. Therefore qPCR analysis are necessary to wider understand the regulatory activities of these genes.
To analyse the less expressed subgenomes, temperature gradient and colony PCR with subgenome specific primer were performed. Basically, I found the annealing temperature at which primers were specific for one subgenome.

Another relevant mechanism that is known to be involved in several regulatory aspect of gene expression is the alternative splicing.
Total RNA was extracted from wheat inflorescences (0.8–1.2mm anthers) using an Isolate II RNA Plant kit (Bioline). Splice variants, both in coding region (CDS) and in untranslated region (UTR), were detected by designing gene specific primers. The PCR amplicons were cloned into pDrive vector (Qiagen) and sequenced from Eurofins GATC Biotech Company. Cloned sequences were analysed using Ensembl Plant database, selecting transcript/splice variants tool and Omega Clustal W DNA Multiple alignment. Lastly, I checked on ExPASy Bioinformatics Resource Portal (Gasteiger *et al.,* 2003) whether my transcript would translate into the predicted protein.
Regarding untranslated regions, first-strand cDNA was synthesized with a GeneRacer kit (Invitrogen). The 5' and 3' ends of *ASY1, ASY3* and *ZYP1* cDNA were amplified by PCR in the presence of *Platinum* Taq DNA Polymerase High Fidelity (Thermo Fisher Scientific) according to the GeneRacer kit protocol. An aliquot of this reaction was used as a template in the second PCR with specific nested oligonucleotides previously designed.
The analysis was carried out both in ovaries and anthers to test whether there are any differences between female and male which could affect the gene expression. Only one splice variant in CDS has been found in Apogee CDS, while two different splice variants 335bp and 398bp have been identified in the *ASY1* 3'UTR. Short UTRs are highly expressed (95% of clones have short version). Therefore, results have not shown massive differences at sequence level among sex so far. We hypothesize that epigenetic modifications, such as methylation, could be involved in gene expression regulation. At this stage, I intend to repeat the same analysis on the other genes of interest.

Once homeologous gene dominance was identified, I planned to examine transcript function of mutant lines for *asy1, asy3 and zyp1.* This approach consists of three different steps:
-identification of TILLING (Targeting Induced Local Lesions In Genomes) mutant lines
-genotyping mutants by sequencing PCR products using sub-genome specific primers
-design crossing scheme to knock out single homeologous and analyse the mutants through cytological approach.

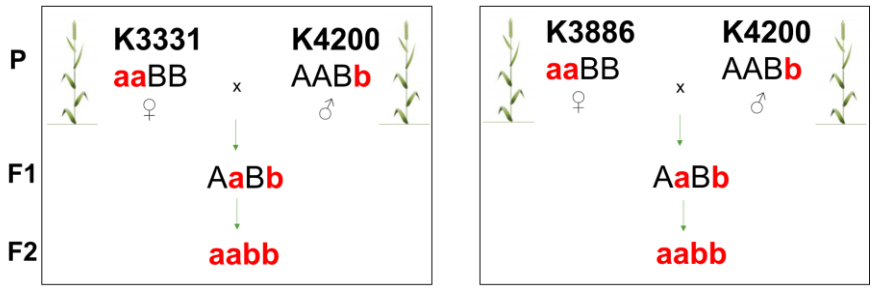
TILLING (http://www.wheat-tilling.com/) is a method of inducing point mutations in specific genes. To choose a mutant line, I used BLAST with the DNA Scaffold sequence of the gene of interest and selected the mutation "stop_gained" (premature termination codon), which potentially result in a loss of function of the gene. If there are no "stop_gained" mutations available, then a "splice_acceptor_variant" or "splice_donor_variant" could be used. These are mutations in the GT

or AG sites that are located at the start and end of introns, respectively. A mutation that results in loss of one of these sites could lead to incorrect splicing resulting in a non-functional protein. TILLING lines were analysed as follows: Genomic DNA was extracted from leaf Zadocks' stage 11 using DNA DNeasy Plant Mini Kit (Qiagen) and CDS were amplified using gene specific primers with MyTaq Red Mix Polymerase (Bioline). PCR fragments were purified using E.Z.N.A. Cycle Pure Kit (Omega Bio-tek). Samples were then sequenced with subgenome specific primers (GATC Company) and chromatogram peaks were analysed with Chromas Lite 2.1.1 software.
To confirm the mutations, RNA extraction from TILLING second inflorescence were performed, cDNA amplified by using gene specific primers and Q5 High-Fidelity DNA Polymerase (New England Biolabs); PCR products were then cloned and sequenced. Mutant clone CDS sequences were aligned up with a wild-type clone sequence in Omega Clustal W multiple alignment and the mutation were localised.

To test if any phenotype exists in the *asy1* (Figure 2) and *zyp1* (Figure 3) mutants, I selected two independent lines per gene having AaBB and AAbB genotypes firstly to generate single knock out in A subgenome, as it is the dominant homeologous in these genes. Seeds were ordered through the UK Germplasm Resource Unit (GRU) website called SeedStor. Eight seeds were sown and screened to confirm that they contain the mutation. Homozygous individuals for mutation were crossed, while heterozygous individuals were self-pollinated in parallel to create an F2 population segregating for the mutation. A wild-type line was also included as controls. Next, I will select homozygous progeny (aaBB and AAbb) from AaBb plants to complete knock outs of the genes and evaluate the effect of this alteration. I will repeat the procedure for *asy3*. Once I have generated the double mutants, I will analyse the phenotype cytologically. Meantime, I have immunolocalised the protein distribution by using specific antibodies provided in Higgins's lab.

Figure 2. Crossing scheme for asy1 knockouts in Kronos (4n)

Figure 3. Crossing scheme for zyp1 knockouts in Kronos (4n)



Summary

So far, I have identified subgenome dominance among the genes, two splice variants in 3'UTR *ASY1* gene and I have crossed *asy1, zyp1* tilling lines.
Next, I intend to perform qPCR analysis on homeologous expression of *ASY1, ASY3 and ZYP1*; complete UTRs analysis, cross and perform cytological analysis of all mutant lines.

Implication
Further understanding the dynamic of chromosome pairing and gene regulation activity in polyploid wheat, which is tightly connected to how recombination occurs, will benefit not only researchers and plant breeders, by providing a useful tool to generate novel cereal lines, but also the entire society for greater and sustainable reliability of the wheat crop.

Skills Training received:

- On October 2018 I attended a three-day course for Preparing to Teach in Higher Education at Leicester Learning Institute, University of Leicester, and received a certificate of attendance.

- From November to December 2018 I attended a 4-weeks English course EL7060 – Speaking section, at English Language Teaching Unit, University of Leicester, and I obtained a certificate of attendance at the end of course.

Meetings attended:
I held monthly meeting with my supervisor at the end of each month.

<u>Year 2018</u>

- May: I cloned the coding region of *ASY1, ASY3 and ZYP1* gene extracted from leaf

- June: I carried out homeologous expression analysis of genes by using Omega Clustal W

- July: I genotyped TILLING lines of *asy1a-b, asy3a-b* and *zyp1a-b*

- August: I complete coding region analysis, finding the less expressed subgenome by performing a colony PCR

- September: I cloned the coding region of *ASY1, ASY3 and ZYP1* gene extracted from spike

- October: I performed RACE PCR to localise UTRs of *ASY1, ASY3 and ZYP1*

- November: I immunolocalised ASY1, ASY3, ZYP1 and performed metaphase spread on Kronos wild-type; I crossed single mutant of *asy1* and *zyp1;* I attended English course – speaking section at University of Leicester

- December: I held my first PGR seminar at University of Leicester; I sequenced the coding region of TILLING lines and sown *asy1a-b* and *zyp1a-b* single knocked out plants

<u>Year 2019</u>

- January: I genotyped *asy1a-b* and *zyp1a-b* single knock out plants and sow *asy3* single mutant TILLING lines; currently I am running experiment to assess UTRs in less expressed subgenomes of genes of interest and next I will fix anthers to immunolocalise the meiotic proteins and made metaphase spread to assess the effect of mutation of each gene.

Outreach activity:
On 12th and 13th September I participate as volunteer to GENIE's Dynamic DNA at Department of Genetics, University of Leicester. The event deals with genetics outreach targeted for nine-year students and their teachers, and includes vary outreach programme, such as building a mini DNA model, extracting DNA from a banana, solving a crime using a DNA fingerprint and making a DNA bracelet.

References

- World Population Prospects http://esa.un.org/unpd/wpp/Documentation/publications.htm
- IPCC in Climate Change 2007 https://www.ipcc.ch/publications_and_data_reports.shtml
- Lobell *et al*. Science, 333: 616-620, (2011)
- Higgins *et al. Plant Cell*, 24: 4096-4109, (2012)
- Higgins et al. Plant Cell, 26:729-40 (2014)
- Zadoks *et al.,* Weed Research 1974 14:415-421
- Kersey *et al.,* Nucleic Acids Res. 2018 D802-D808. doi: 10.1093/nar/gkx1011 http://www.ensemblgenomes.org
- Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species Nucleic Acids Research 2018 46(D1) D802–D808
- Sievers *et al.,* Mol Syst Biol. 2011 Oct 11; 7:539
- Gasteiger *et al.,* ExPASy Nucleic Acids Res. 31:3784-3788 (2003)
- http://plants.ensembl.org
- http://www.wheat-expression.com/