Project Title:

Harvesting the genetic value of interspecific wheat introgressions

Report Title:

Progress towards harvesting the genetic value of interspecific wheat introgressions

Summary

A very short summary or abstract of your project and the main outcomes (200 word max)

Material generation: Two wheat lines, differing only for the target introgression region on chromosome 4A, were multiplied and phenotypically assessed at NIAB. Purity seeds from individual field-grown plants were transferred to our partners at the Centre for Plant Structural and Functional Genomics (CPSFG) for chromosome flow sorting.

Sequencing preparation: Toby Barber (TB) visited CPSFG to co-ordinate flow sorting, with the long arm of chromosome 4A extracted. Libraries were generated for transfer to Earlham Institute (EI) for sequencing. EI advised us of a 3-4 month delay before they could start the library preparation, with a further 8-10 week turnaround from this. It was therefore agreed that CPSFG would prepare the libraries (further in-kind contribution) and send directly to EI with an anticipated sequencing timeline of one month from receiving the libraries.

Additional outcomes: Whilst awaiting sequence data TB has taken the opportunity to prepare for the analysis by registering for an EBI training course on bioinformatic analysis of NGS data in April. Once the data is available we will annotate the genes contained within this region through a bioinformatics analysis. We have presented the project on several occasions to academic and agro-industry stakeholders at both internal and external events.

Report and outcomes

The report should be a short written description of what you accomplished and how, including an evaluation of where things didn't go to plan! Please include links to or copies of any outputs (papers, posters, data, code, hardware designs, photos, blogs, videos) - these do not need to be duplicated in the report. If you are attaching supplementary files, please refer to them in the text or add a list with a brief description at the end e.g. OpenPlantReport.pdf: Formatted version of full project report. FooSequences.fasta: Sequence file for DNA parts ACB123 and DEF345. Refer to the reporting guidelines for more information.

A Near-Isogenic line (NIL) pair which differs in the presence/ absence of the target interspecific introgression from *Triticum dicoccoides* on the long arm of chromosome 4A were identified. Their identities were confirmed via field phenotyping (recording key growth stages, plant height, flag leaf measurements and plot grain weight) to ensure that the correct lines were being grown and selected. 200 seeds of each of the NIL pair were transferred to the Centre for Plant Structural and Functional Genomics (CPSFG) in Olomouc, Czech Republic.

TB then visited the centre in early November 2017 to see first-hand the methods used to extract the meristems from roots tips in seedlings, and the process of flow sorting.

Whilst at CPSFG and after discussions in relation to forecast delays in library preparation at EI we decided to have the DNA libraries prepared by CPSFG. They have expertise in preparing libraries from a small sample size and were willing to make an additional in-kind contribution to the project. TB is in regular contact with both CPSFG and EI regarding details of timing and DNA concentrations and keeping both parties informed on timelines. The libraries have just been prepared at CPSFG and

once prepared they are stable for around one month. The prepared libraries will be transferred to EI for NGS and we are hopeful of receiving sequence data promptly.

The EI time delays have however given TB the opportunity to apply for and obtain a place on a course run at The European Bioinformatics Institute (EBI) 'Introduction to Next Generation Sequencing' in April. This will assist TB in the bioinformatic analysis and gene identification of the sequenced chromosome arms from the NIL pair.

We have presented the project at several meetings to date. In October TB presented the project to the Bioinformatics group at NIAB in Cambridge and via web-link to NIAB-EMR in Kent. The bioinformatics group consists of around 20 people who work on and have an active interest in bioinformatics and meet regularly. Some useful tips on how the analyses should be carried out were given along with offers of assistance from experienced members when the analysis begins.

TB also presented a poster on the project at the NIAB Poster Day in early February. This is an annual event which brings together all of NIAB with different teams and departments (over 200 people) showcasing their work and projects (attached is a PDF of this poster: TBarber_OpenPlant_poster).

On 6th February team member Alison Bentley presented at the Hutchinsons Arable Conference to ~250 agronomists about the importance of diversity in wheat breeding. This included the 'Robigus' pedigree figure underpinning our OpenPlant project and stimulated plenty of debate and feedback on the importance of understanding diversity bottlenecks (included in set of photos attached in: TBarber_OpenPlant_photos).

Expenditure
A summary of how you have spent the £4000 budget so far.

We have received an updated quote from EI (as the libraries are no longer being prepared by them), with a cost saving of £457 on the Net price. The current NGS cost £4,041.79 (a reduction from £4,498.79). The flow sorting and library preparation have been successfully performed by CPFSG as an in-kind contribution. TBs travel and stay in the Czech Republic was funded from his student training travel budget.

Item	Total
Pre-made library QC x 1	
HiSeq2500_RR_lane (250PE) x 1	£4,041.79
Data transfer	
TOTAL	£4,041.79
VAT (If applicable)	£808.36

Follow on Plans
A short description of your plans for follow-on work with a breakdown of how you will spend the additional £1000 (if requested) and any remaining funding from the initial £4000. Please include timings as it is expected that all funds will be spent within six months of this report, after which point a brief report from the follow-on activities and return of the remaining funding will be requested.

We wish to apply for the £1000 follow-on funding.

Due to the delay in the sequencing of the flow sorted chromosomes the project is not yet complete but it is continuing as planned and we have made significant progress. We propose to use the follow-on funding as follows:

- **EBI bioinformatics training course** for TB in April (£560). This will be just in time for the data to be available to start analysis (provided no further delays in receiving NGS data).

- **Project dissemination**: TB has applied to present a poster and 'flash presentation' at the Monogram conference at the end of April at JIC (www.monogram.ac.uk). Monogram is a UK network of researchers and industry stakeholders with an active interest in small grain cereal and grasses and this would be an ideal place to present our OpenPlant project and talk to academia and industry about this project. The total cost is estimated as £270 (conference registration £150; poster printing £60; overnight accommodation £60).

- **Sequencing**: We would also like to use the remaining £170 towards some of the additional cost of the sequencing which was over the originally granted £4000 (Net: £4,041.79, VAT: £808.36, Gross: £4,850.15) with the remaining scheduled to come from TB's PhD operating budget. This saving will allow TB to use his existing budget for further validation experiments.

Proposed expenditure of additional £1000 follow on fund.

Expenditure	Cost
Introduction to Next Generation Sequencing course at EBI	£560
Monogram Conference registration	£150
Monogram accommodation	£60
Monogram poster	£60
Sequencing at EI	£170
total	£1000

Changes to team

Please include here a note if there have been any changes to the team so that we can make sure the information on the www.biomaker.org website is up to date. If additional team members have joined, please provide a photo together with their name, job role and affiliation.

NA