Final Report

Application of Association Mapping and Genomic Sequencing to Starch and Glycaemic Index in Potato

Ref: 11140010

Reporting Period: September 2013 - September 2016

Report Authors: Mark Taylor, Sanjeev Kumar Sharma, Laurence Ducreux, Glenn Bryan

Date report submitted: February 2017

Report No.: 2017/1

Additional copies of this report and a list of other publications can be obtained from:

Publications

AHDB
Stoneleigh Park
Kenilworth
Warwickshire
CV8 2TL

Tel: 02476 692051
Fax: 02476 478902
E-mail: Potatoes.Publications@ahdb.org.uk

Our reports, and lists of publications, are also available at potatoes.ahdb.org.uk

CONTENTS

1. SUMMARY

***Aim**:* This project aimed to develop a high throughput method for assessing starch digestibility from a panel of 300 potato varieties. Using an association genetics approach, a further aim was to link starch digestibility characteristics to gene markers providing a tool for predicting starch digestibility.

Methodology: The panel of 300 varieties was grown in replicated field trials over two growing seasons and starch was extracted and purified from tuber samples. An *in vitro* digestibility method for predicting the glycaemic index (GI) from the 300 genotypes was developed to give a robust and high throughput method for starch characterisation. An association mapping approach was applied to the starch digestibility data to detect genomic regions associated with GI traits in potato.

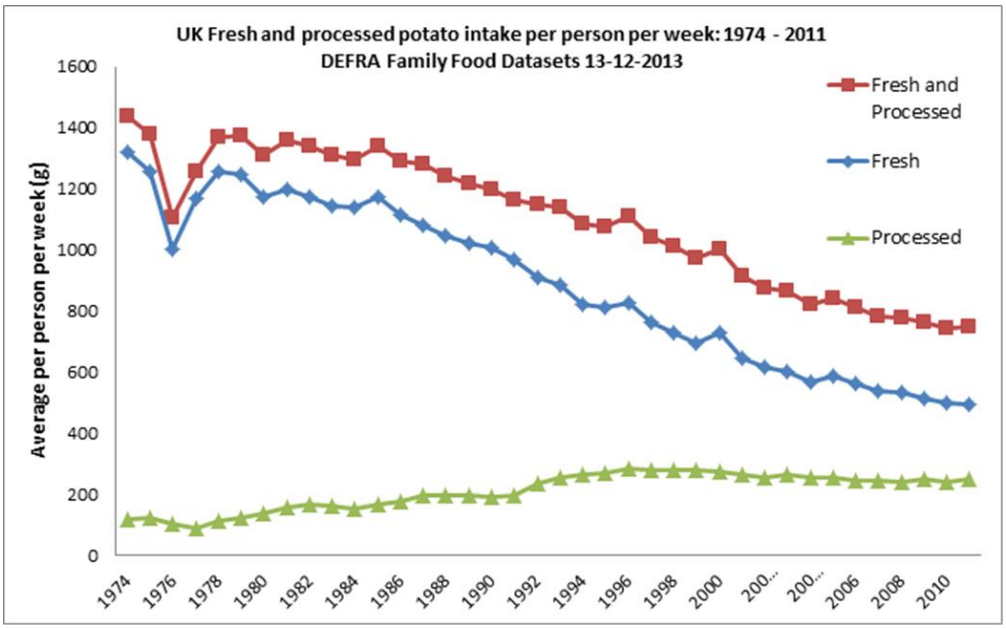
***Key findings**:* The starch digestibility characteristics of samples from the 300 genotypes used in the study were determined. Trials were conducted over two growing seasons, using a replicated block design. Consistent variation in starch properties between the genotypes was observed in the year 1 trial. The starch digestibility characteristics were used in an association mapping approach using the SolCAP SNP marker data available for the association panel (ca. 8300 SNPs) and 44700 SNP markers available from a genotype-by.-sequencing approach. Significant associations with starch digestibility were detected using the first trial. For the second trial the data were not well correlated between the two blocks used in the trial or with the first trial.

Practical recommendations: A reproducible and reliable starch digestibility assay has been developed. The assay requires 100 mg of starch and measures glucose release over 60 minutes. Compared with previous methods, this is a high throughput assay that can be used to assess many samples simultaneously. The starch digestibility data from the first trial indicated consistent associations between markers and traits, providing an indication of the genetic architecture of the starch digestibility trait. Data from the second year trial were not consistent between replicates within the trial or with the first trial. Reasons for these discrepancies are suggested. Nevertheless the first trial data are promising and provide the potential for validating the markers identified in follow on work.

2. INTRODUCTION

Despite potatoes ranking as the second most prevalent carbohydrate food consumed in the UK, (AHDB 2013), a report from the Department of Environment, Food and Rural Affairs (DEFRA) concluded that the consumption of fresh potatoes and processed potato products had decreased by ca. 40% over the past 30 years, (see Figure 1), (DEFRA 2013). Possible reasons for the decline in the consumption of potatoes have been discussed by Riley (2010) who states that working patterns, price, demands for convenience foods and increased food choices may contribute to this decline. Other negative factors in potato acceptability with consumers are general nutritional issues.

Figure 1: UK Fresh and Processed Potato Intake per Person per Week: 1974-2011



Source: DEFRA 2013

The developed world is facing an obesity crisis. Metabolic syndrome occurs from having a combination of the following medical conditions: obesity; high blood pressure; high blood glucose levels and high cholesterol levels. It is a disease which affects energy storage and utilization and increases the risk of developing diseases such as diabetes and heart disease (Alberti et al 2005). Diabetes is caused by uncontrolled blood glucose concentrations after eating a carbohydrate meal. Insulin, a hormone produced in the pancreas, is used to enable glucose to enter cells in the body as well ensuring that blood glucose levels are kept at an appropriate level (International

Diabetes Federation (IDF) 2014). Diabetes occurs when insulin is not being produced at all (type 1) or when the body is unable to use insulin correctly (type 2) (Mishra et al 2012). Type 2 diabetes is the most common type of diabetes affecting over 90% of individuals who have the disease worldwide and is becoming more apparent in children, (WHO 2013). The glycaemic index was developed as a means to classify foods based upon their blood glucose response after carbohydrate consumption and is defined by Jenkins et al (1987) as "the increase in area under the blood glucose response curve after consuming a food portion containing 50g available carbohydrate as a percentage of the response to a 50g dose of glucose". Thus low GI diets are recommended to individuals that suffer from diabetes as a tool to modulate the disease. Foods are ranked into 3 categories:

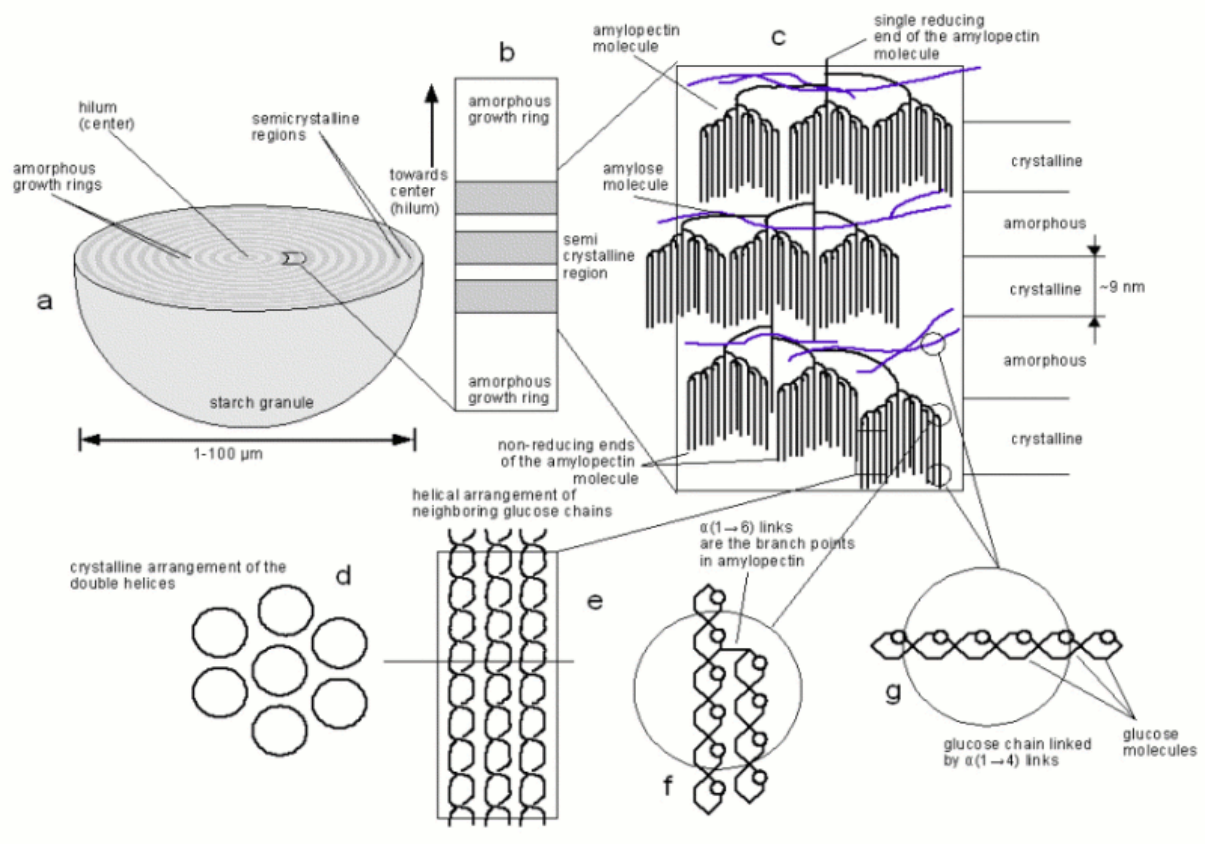
- High (GI value >=70)
- Medium (GI value 56-69)
- Low (GI value <=55)

The GI of potato is generally high but extends over a wide range (56-104) (Ek et al 2012). Study to study variability is a problem despite there being an International Organization for Standardization (ISO) approved method for determining the GI of foodstuffs. There is also inherent variation between genotypes for example the potato variety Marfona has a GI value of 56 and Maris Piper has a GI value of 85 (Henry et al 2005). Intrinsic factors such as starch characteristics e.g. amylose: amylopectin ratio; starch granule size; starch phosphate content, tissue structure e.g. cell walls, can affect the GI of potatoes (Singh et al 2010), along with food processing techniques, textural and rheological characteristics of the food and the presence of other nutrients (Henry et al 2006; Monro et al 2009 and Kinnear et al 2011). As a result of the current obesity crisis the desire to lose weight has led consumers to opt for low carbohydrate diets, such as the Atkins diet, and low GI diets. In turn this has led consumers to perceive potatoes to be bad for their health as they are a large source of carbohydrate and some are of a high GI value. As a consequence dietary intake of carbohydrate has decreased, suggesting that this also could have contributed to the decline in consumption of potatoes (Riley 2010).

Starch is the most common storage carbohydrate in plants and is the biggest source of glycaemic carbohydrate in the human diet, (Singh et al 2010). The starch granule is present in plant amyloplasts and is a glucan polymer consisting of amylose (approx.

25%) and amylopectin (approx. 75%), and is comprised of semi-crystalline and amorphous layers, (Figure 2), (Ahuja et al 2013).

Figure 2: Starch granule structure



Source: Waigh et al 1997

In vitro starch digestibility methods have become a popular tool and a proxy method for obtaining an indication of GI without the use of human volunteers (Englyst et al 1992). The method involves a two hour shaking incubation (37°C) with digestive enzymes to mimic digestion that takes place in the small intestine (Englyst et al 1992). From this method Englyst and others concluded that there were three major types of starch:

- Rapidly Digestible Starch (RDS): glucose release after 20 minutes
- Slowly Digestible Starch (SDS): glucose release between 20 and 120 minutes
- Resistant Starch (RS): starch which has remained undigested after 120 minutes

SDS and RS provide a number of health benefits. SDS can offer the possibility of modulated glucose delivery to the body, along with fullness and satiety (Haub 2013; Zhang and Hamaker 2013). RS escapes digestion in the small intestine and undergoes fermentation in the large intestine by gut microbiota (Flint 2013). This offers the possibility of reduced post-prandial blood glucose levels, improved insulin sensitivity and the prevention of colorectal cancer (Maningat and Seib 2013). Therefore, the ability to control or predict starch digestion and glucose absorption is of increased importance in order to modulate diseases such as diabetes and obesity.

Association mapping (AM) is a method of mapping quantitative trait loci (QTL), regions of DNA that contain genes which contribute to a specific trait, in order to link phenotypes to genotypes (Zhu et al 2008). Originally used to study disease in human populations in the form of Genome-Wide Association Studies (GWAS), it is now becoming important in plant research with the potential to accelerate into plant breeding. This is because AM can be used to dissect trait variation across much broader germplasm populations than conventional genetic mapping of biparental crosses. In comparison to traditional linkage analysis, it offers significant advantages such as increased mapping resolution, reduced research time, and greater allele number (D'hoop et al 2014). In this project the aim was to apply an association mapping approach to the starch digestibility dataset to detect genomic regions associated with GI traits in potato.

Key project objectives are to:

- Phenotype ~300 varieties from the association panel over two years for factors relating to GI in potato
- Validate *in vitro* digestibility methods for predicting the GI from the 304 genotypes by adapting current digestibility methods (Englyst et al 1992-1999)
- Use association mapping to detect genomic regions associated with GI traits in potato, identify gene markers and potentially candidate genes associated with GI in potato

3. MATERIALS AND METHODS

Field Trials

The potato association panel containing of 300 tetraploid genotypes, was grown under conventional agronomic conditions in replicate field plots at Balruddery Farm, Dundee in mid-April 2013 and harvested in October 2013 and in the same period in 2014.

Starch Extraction from Potato Tubers

The starch extraction method was adapted from Muazu et al (2011). 8 tubers per genotype were selected, washed carefully and patted dry with tissue paper. Opposite eighths were cut from each tuber and the eighths diced into cubes. 400mL of distilled water was placed into a blender and approx. 10mg of sodium metabisulphite was added - this is an antioxidant to ensure that the tubers do not oxidise and turn brown. The diced eighths were then placed into the blender and blended for 30 seconds. The mixture was passed through two layers of muslin, placed over a 1L beaker secured in place with an elastic band, before removing the muslin from the beaker and gently squeezing any remaining liquid into the beaker. Solid was removed from the muslin and placed back into the blender along with 200mL of water and approximately 10 mg of antioxidant and blended again and filtered through two new layers of muslin. This extraction was repeated once more with 200mL of water and antioxidant, but passed through a single later of muslin instead of two layers. The remaining solid and muslin was discarded at this stage.

The starch was allowed to sediment for ca. 1 hour until thick white starch was visible at the bottom of the beaker. Liquid was poured off and the starch sediment resuspended in 500mL of distilled water before sieving through two layers of muslin into a clean beaker, ensuring that all starch is removed from the beaker. The starch was then allowed to sediment again for ca. 1 hour. Liquid was poured off and the starch granules were left to dry, uncovered, at room temperature overnight. Starch was transferred into a clean petri dish and incubated at 30°C for 48hours, breaking up the lumps with a clean spatula every two hours until completely dried. Once dried, starch was transferred into labelled 50mL tubes (in duplicate for each genotype) and stored in racks in boxes at room temperature.

This procedure was repeated for duplicate samples from each of the 300 genotypes.

Starch washing procedure

For accurate assessments of starch digestion, further purification of starch was required to remove residual sugars associated with the crude starch preparation. Washing in hot 80% ethanol proved to be an effective means of removing sugar contaminants. The detailed protocol is given below.

1- Weigh 1.5g of crude starch extract into 50ml tube
2- Add 30ml of 80% ethanol and vortex
3- Place in an 80°C water bath for 1 hour
4- Place sample on ice for 5 min to stop ethanol extraction
5- Centrifuge for 20 min at 4000rpm at 4^0C
6- Remove supernatant
7- Add 30ml 80% ethanol and vortex
8- Place in an 80°C water bath for 30 min
9- Place sample on ice for 5 min
10- Centrifuge for 20 min at 4000rpm at 4^0C
11- Remove supernatant
12- Wash the pellet twice with 30ml sterile distilled water, centrifuging for 10min at 4000rpm (4°C) between each wash.
13- Remove supernatant
14- Pour clean starch in a petri dish
15- Dry it at room temperature (usually 2 days)
16- Transfer clean dry starch into a 2ml Eppendorf tube

Determination of Dry Matter Content

The remaining opposite eighths from each genotype were diced into small cubes and placed into a 50mL tube on tared scales and the weight recorded. The tube containing potato was flash frozen in liquid nitrogen then put into the freezer at -20°C until ready to be freeze dried. The sample was put onto a freeze drier for one week, and then weighed into another 50mL tube on tared scales and the weight noted.

Starch Digestion Time Course Protocol

This method is an adaptation from Englyst et al (1992, 1996, 1999). In principle, an enzyme mix designed to mimic the major starch digestion activities in the human gut is added to a gelatinised and cooled starch sample. The glucose released from digestion

is measured over a 60 minute time course and from the degree of digestion at three time points (10, 20 and 60 minutes), the digestibility of the starch is estimated. The original method has been modified so that the assay can be performed in a high throughput manner with a smaller sample size.

Preparation of Enzyme mix (prepare on ice):

Pancreatin:

1. Weigh out 3.24 grams of Pancreatin (Sigma porcine pancreas 8 x USP specifications (ref P7545)) into a centrifuge tube.
2. Add 21.6mL of dH20 and vortex well. Add mini magnetic bar to the tube and mix on magnetic stirrer for 5 minutes.
3. Centrifuge the tube at 4000 rpm for ten minutes. Remove supernatant and place into a separate centrifuge tube. Discard the pellet.

Invertase:

Weigh out 12mg of Invertase (Sigma Grade VII >=300units/mg (ref I4504)) into 2mL Eppendorf tube and add 1.2mL of dH20. Mix slowly until dissolved.

Amyloglucosidase (AMG):

In a tube add 476µl of AMG solution (Sigma from *Aspergillus niger* >=300U/ml (ref A7095) into 524µl of dH20. Mix well with pipette.

Final Solution:

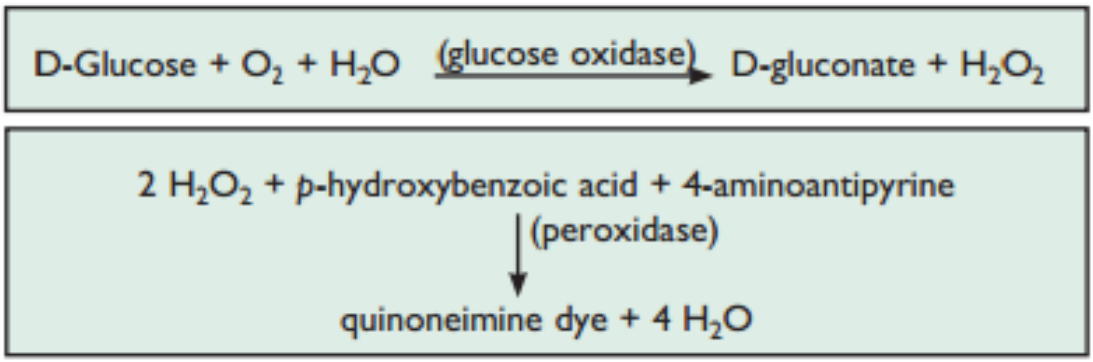
1. Into the pancreatic supernatant solution add 720µl of AMG and 1.08mL of Invertase and mix well. (20mL final solution)
2. Freeze in 5mL aliquots.
3. Before use thaw aliquot at room temperature then keep on ice throughout

Starch Digestion

1. Weigh out corrected weigh of starch (mg) into a 50mL tube e.g. 100mg
2. Place a stirrer bar in each tube and add 5 mL sodium acetate buffer.
3. Put 200mL water onto boil (in a beaker on a heated magnetic stirrer block) and cover with foil.
4. Once water is up to the boil place tubes in the beaker and turn stirrer switch on.
5. Boil samples for 10mins (timed with a timer)
6. Equilibrate to 37°C for 5 mins by placing in 37°C water bath for 5 mins (timed).
7. Add 1.25 mL enzyme mix to each tube. Screw caps on and gently shake back and forth to ensure contents are mixed.
8. Place tube into a shaking water bath (37°C, stroke speed setting 100)
9. Take 100µl aliquots into 1.9mL ethanol at the following time points: 10, 20, 40, 60, 80, 100, 120mins.
10. Centrifuge samples for 10mins at 4000rpm.
11. Released glucose was measured by means of the glucose oxidase-peroxidase (GOPOD) method. The GOPOD reaction process is shown in figure 9. This assay works on the principle that glucose oxidase (GO) oxidises glucose to produce

gluconate and hydrogen peroxide (H_20_2). In the presence of peroxidase (POD) H_20_2 is oxidised by 4-aminoantipyrine and phenol which produces a red/pink quinoeimine dye enabling samples to be colorimetrically visualised and measured at 510 nanometre wavelength. 8µl of supernatant directly into 600ul GOPOD reagent (pre-prepared in 1.5mL Eppendorf tubes) to measure glucose released. Shake to mix. 600µl of GOPOD reagent (glucose oxidase (12 000 U/L), peroxidase (650 U/L) and 4-aminoantipyrine, (0.4mM)) was added to each sample then samples placed into a water bath (50°C) for 20 minutes. 200µl aliquots of each sample were transferred to a 96 well plate and read at 510nm on a plate reader (Molecular Devices – SPECTRA MAX 190). A glucose standard curve, (0, 10, 30, 50, 100, 250 and 500µg/mL), was also included in this assay in order for glucose calculations to be made along with an enzyme only solution as a no substrate control of the experiment.

Figure 3: GOPOD reaction process Source: Megazyme International, 2011



Genetic Analysis of the Starch Digestibility Datasets

The association panel genotyping and characterization details are provided in Sharma et al. (in preparation). Briefly leaf genomic DNA from each clone was extracted (Qiagen DNeasy Plant Maxi Kit), quantified and normalized to a concentration of 30 ng/µL. The panel was genotyped using the Infinium 8k Potato SNP Array (Felcher et al. 2012; Hamilton et al. 2011), containing ca. 8300 SNPs according to the manufacturer's protocols. SNP genotypes were called using R package fit Tetra (Voorrips et al. 2011). Physical positions for all SNPs were derived from Sharma et al. (2013).

Genome Wide Association Scans (GWAS) for starch traits were performed using the mixed model procedure as described by Yu et al. (2006) implemented in the R package GWASpoly (Rosyara et al., 2016).To generate trait values for starch digestibility measurements (year 2013) Best Linear Unbiased Estimates (BLUEs) for 'time_10', 'time_20' and 'time_60' were computed using REML implemented in

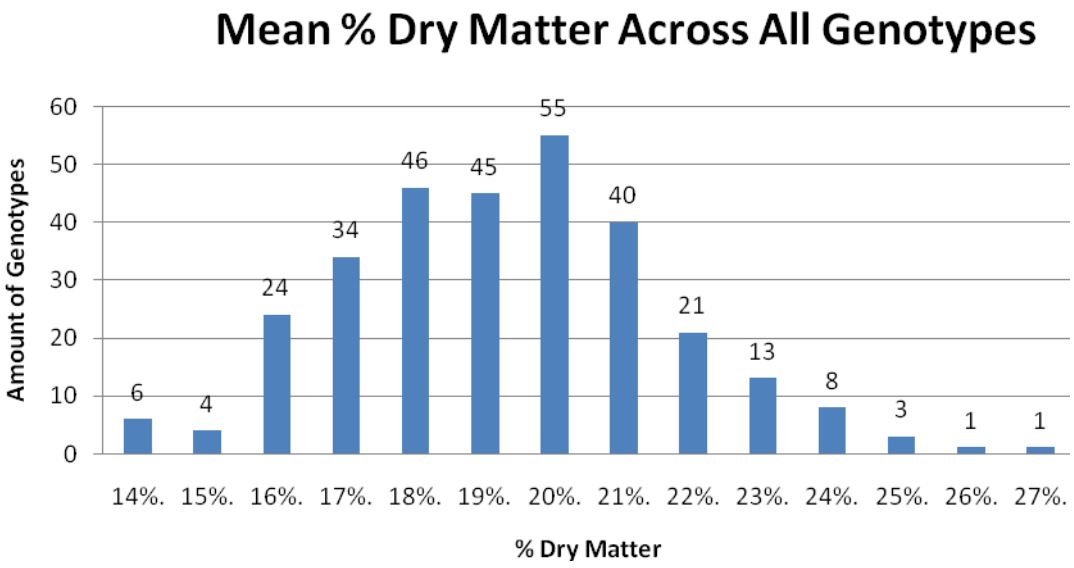
Genstat 15th edition (VSN International Limited, http://www.vsni.co.uk). The Bonferroni correction method was used for establishing a *p*-value detection threshold for statistical significance of marker-trait associations. GWAS analyses were performed on 290 tetraploid clones.

4. RESULTS

Determination of Dry Matter Content

In the course of the starch extractions from the 2013 harvest, tuber dry matter content was measured. Dry matter content ranged from 13%-29%, (full data set in appendix 1), across all genotypes with the mean ranging from 14%-27% dry matter content per genotype and 20% dry matter content being the most common among the 304 genotypes. The distribution of dry matter content across the genotypes is shown in Figure 4.

Figure 4: Mean Percentage Dry Matter Across all Genotypes



Starch Digestibility Assay

Assay development

An *in vitro* starch digestibility assay was developed, based on the assay of Englyst et al (1999) but adapted for use in a high throughput format. This was necessary due to the high number of genotypes that are required for association genetic analysis. The original method was scaled-down so that 100mg of sample rather than 500 mg was used. Additional it was found that digestion for 60 minutes was sufficient as further digestion after this was not significant. An important finding was that prior to digestion, the starch required a hot ethanol wash to remove residual sugars associated with the crude starch preparation. A comparison of results obtained using crude starch preparations (unclean) with those after ethanol washing is shown in Figure 5.

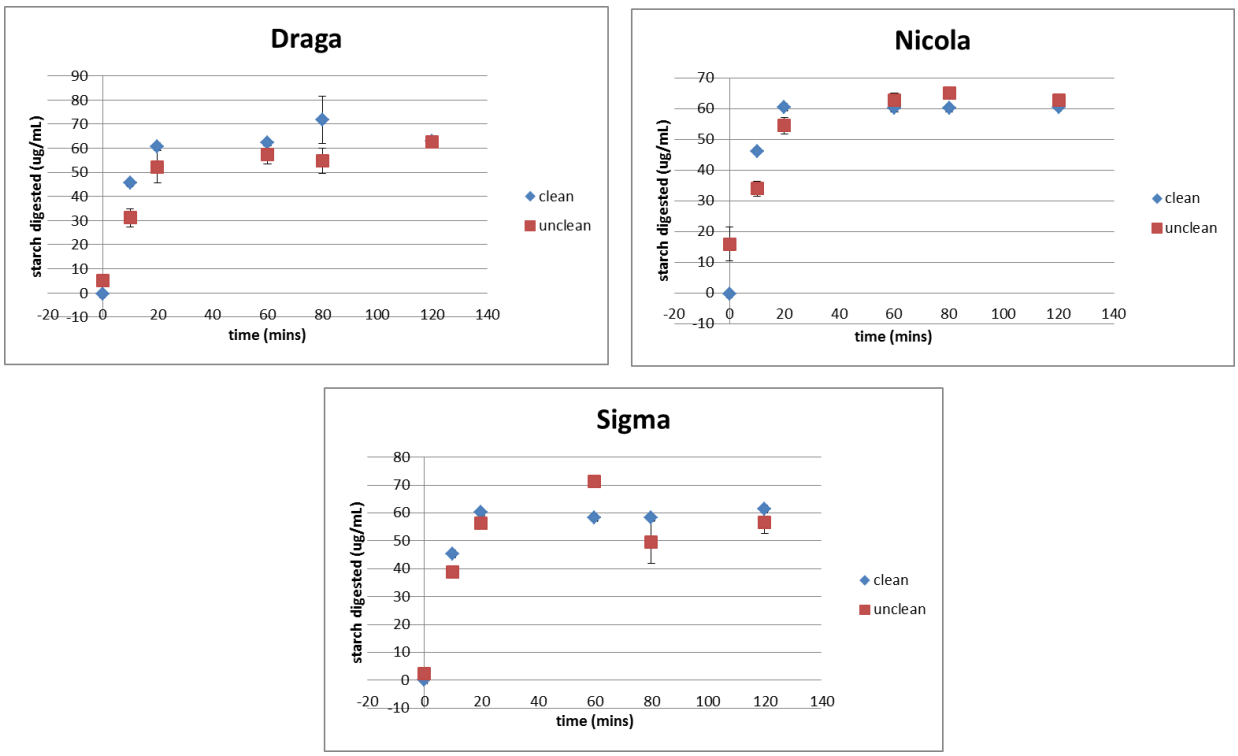
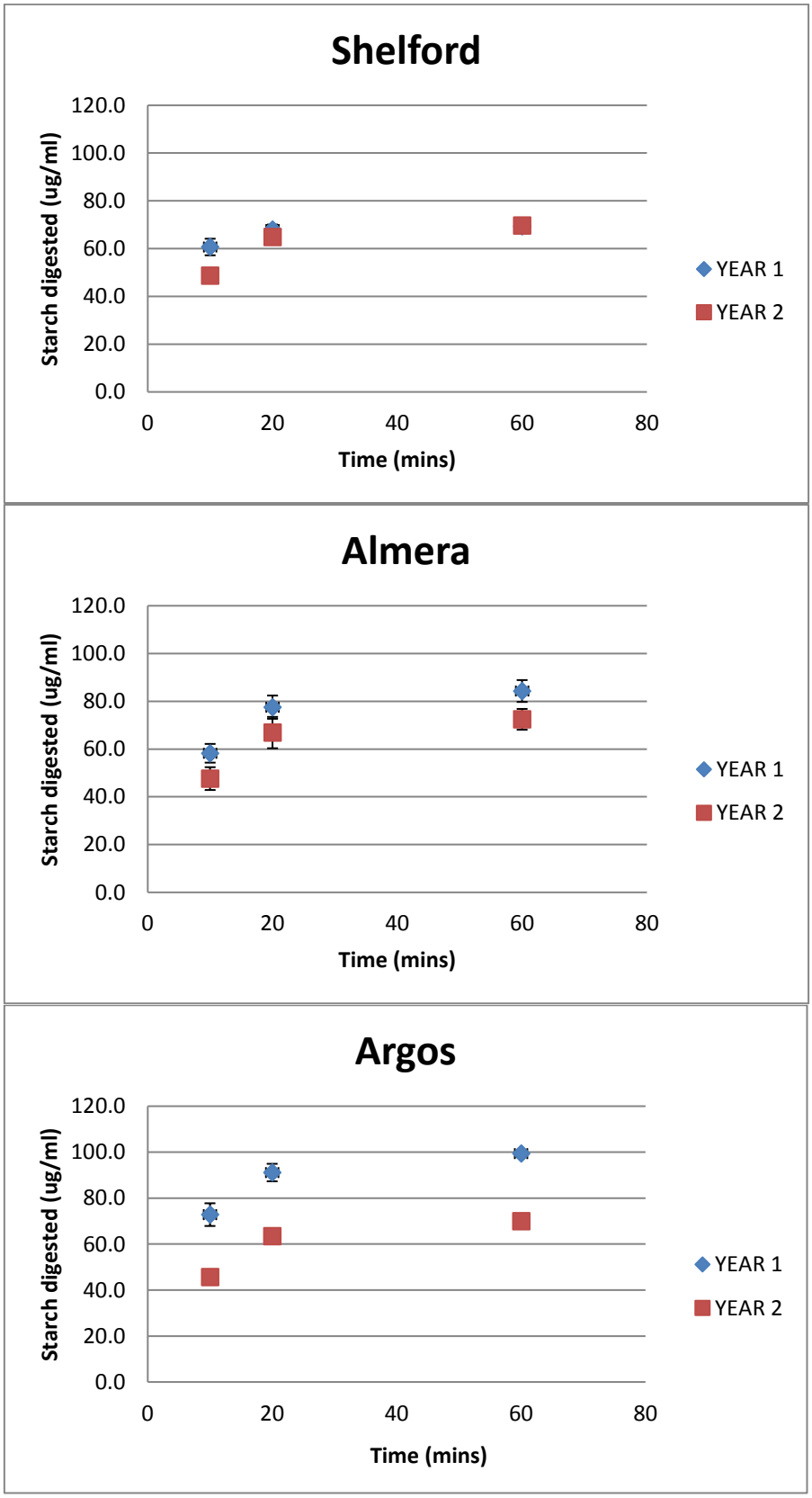


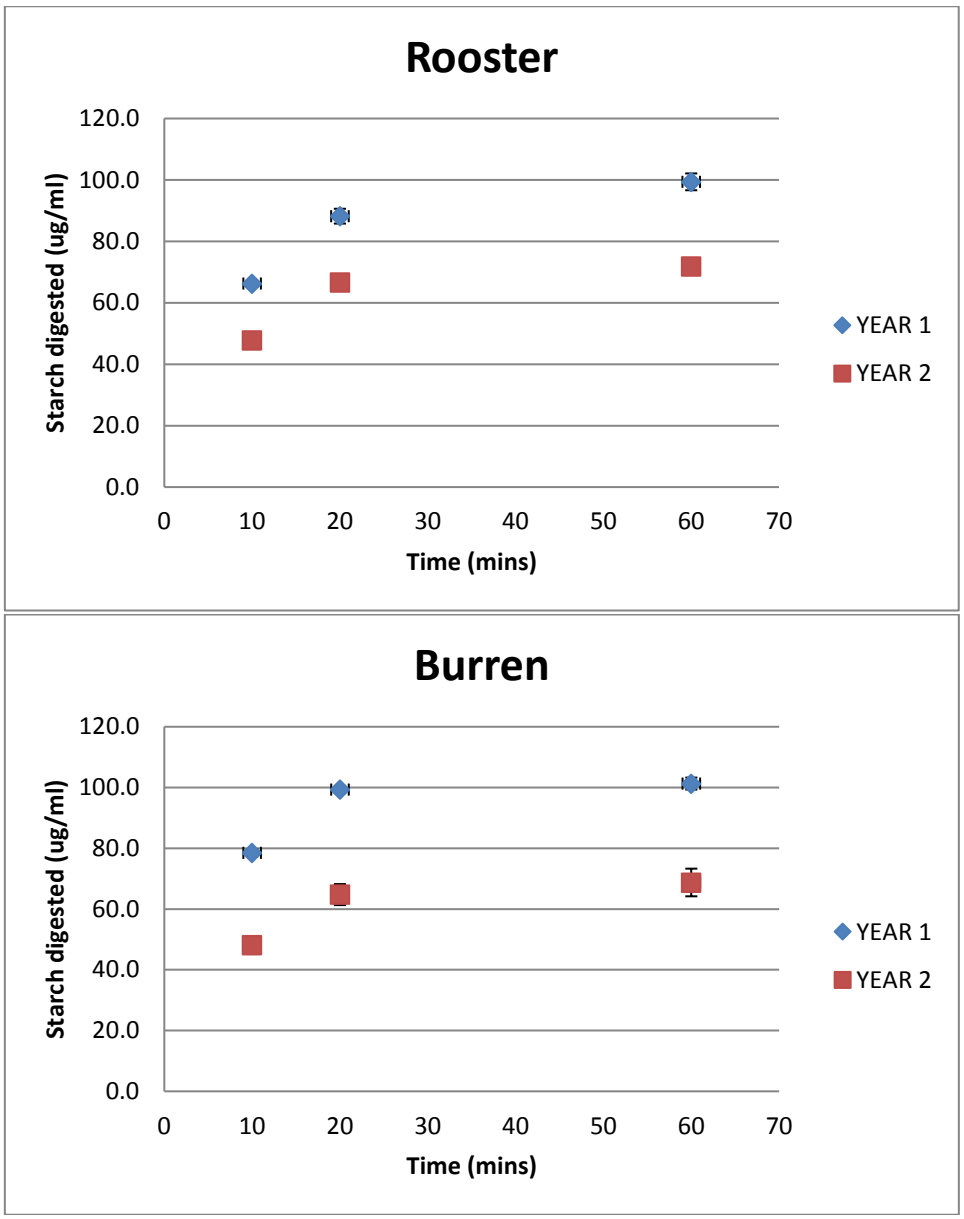
Figure 5 – comparison of glucose release profiles from starch samples prior to (unclean) and after (clean) washing.

Application of the starch digestion assay to samples from the association panel.

Having established a robust high throughput starch digestion method, the assay was then applied to samples from the association panel. The 300 genotypes were grown in two replicate plots in 2013 and 2014. Washed starch was prepared from these samples. The degree of starch digestion was assayed at three time points (10, 20 and 60 minutes, hereafter referred to as 'time_10', 'time_20' and 'time_60', respectively) using the optimised assay. Examples for the glucose release profiles from five cultivars are shown in Figure 6. Based on the 2013 field trial, varieties Shelford and Almera were selected as having a slower glucose release profile whereas Argos, Rooster and Burren exhibited a rapid glucose release profile.

Figure 6 – Examples of glucose release profiles from five cultivars



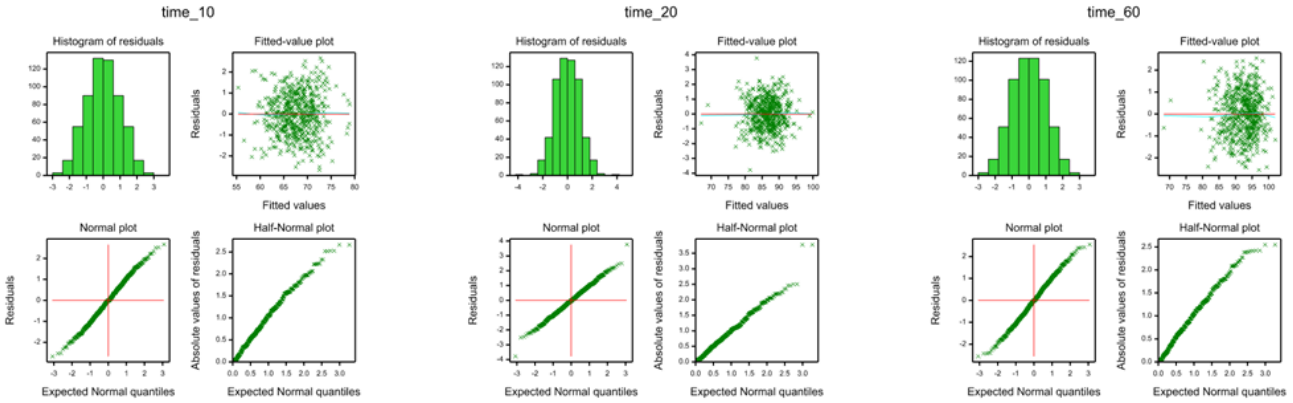


Starch digestion data analysis and GWAS

Best Linear Unbiased Estimates (BLUEs) for 'time_10', 'time_20' and 'time_60' were computed for data obtained from the 2013 and 2014 trials (Appendix 2) using REML. In general the digestibility values for the 2014 (year 2) trial were lower than those observed in 2013 (year 1). Using standard potato starch, purchased from Sigma, reproducible values for starch digestion were obtained during the analysis of both the 2013 and 2014 trial, although batch to batch variations in the enzyme preparations used for analysis were noted. For this reason, large batches of digestion enzyme mix were prepared and used to analyse the starch from each season. For the 2013 data, the residual plots (Figure 7) from REML analysis were inspected for presence of any outliers and suitability of the data for performing GWAS. Three genotypes in 'time_20'

data showed significant deviation from rest of the panel. These were further inspected for any data recording errors and/or mix-ups and, as both replicates for these three genotypes behaved in a similar manner, the obtained values were considered genuine and retained in the analysis. The error term (residuals) appeared to be normally distributed for all three time-point starch digestibility traits, thus, suggesting the suitability of the trait data for the GWAS analysis.

Figure 7 Residual plots for all starch time-point traits from REML analyses 2013 (year1) data.



For performing GWAS on each time-point for the starch digestibility data, four different genetic models as described by Rosyara et al. (2016) were tested, namely (1) additive, (2) Simplex dominant, (3) Duplex dominant, and (4) General. Significant marker-trait associations (MTAs) for 'time_20' and 'time_60' were detected ('General' genetic model) on chromosomes 1, 4 and 6. Sporadic MTAs were also observed on other chromosomes as listed in Table 1. Figures 8, 9 and 10 illustrate the GWAS results obtained from these analyses using the 8300 SolCAP markers and 44700 SNP markers obtained using a genotyping-by-sequencing approach (Bryan and Sharma unpublished). All significant marker-trait-associations are presented in Table 1.

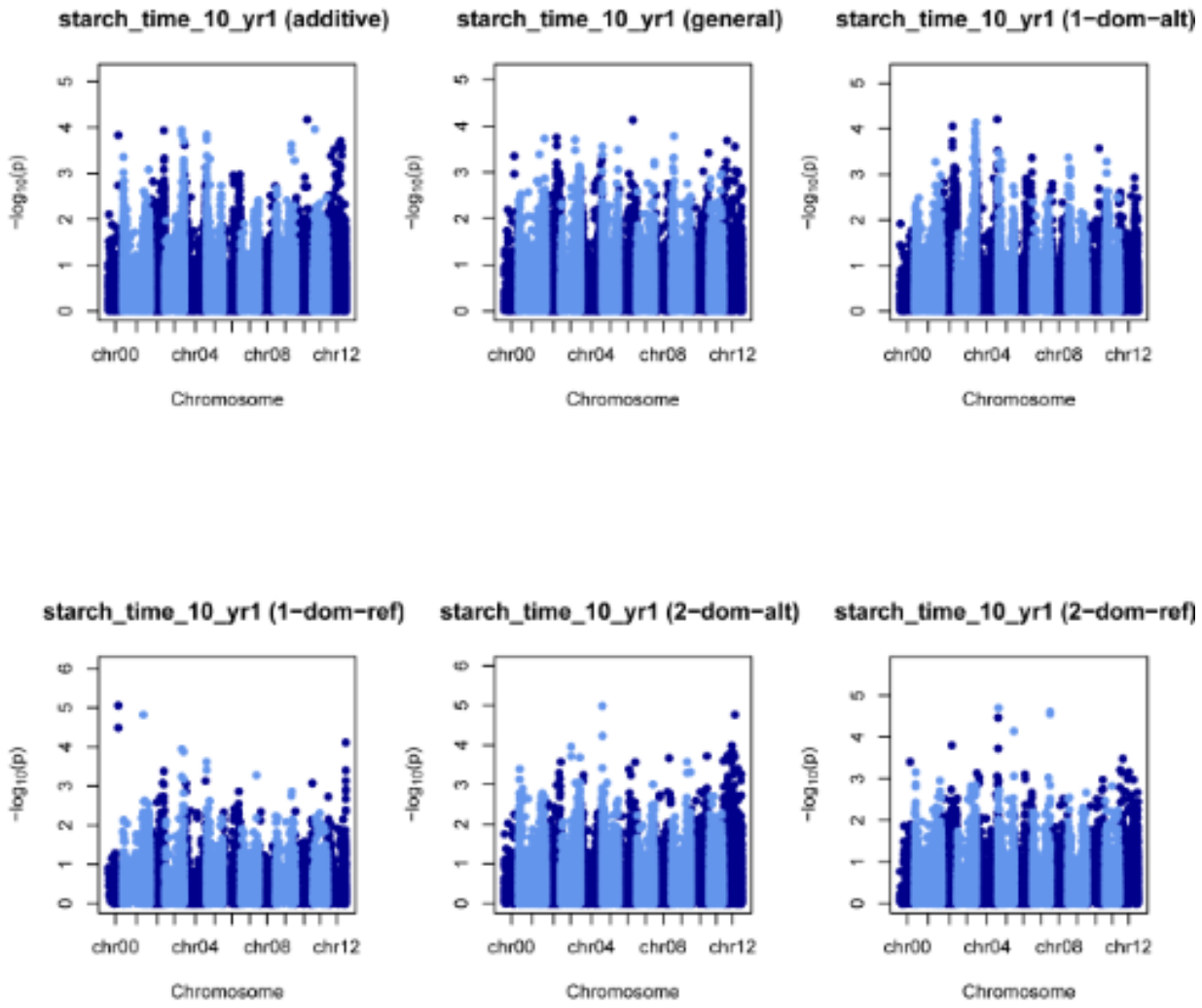


Figure 8: Manhattan plots for starch 'time_10' trait using different GWAS models..

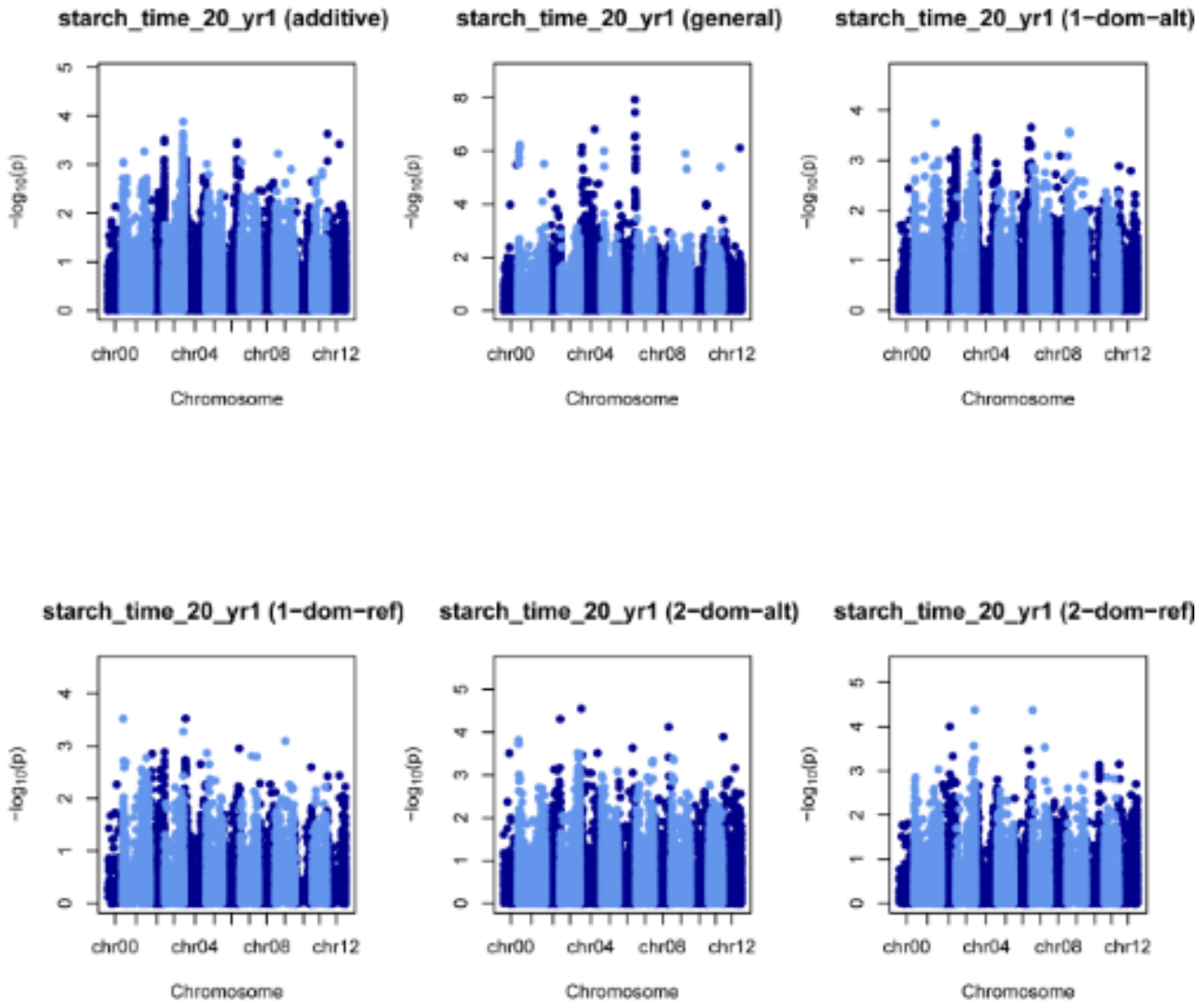


Figure 9: Manhattan plots for starch 'time_20' trait using different GWAS models.

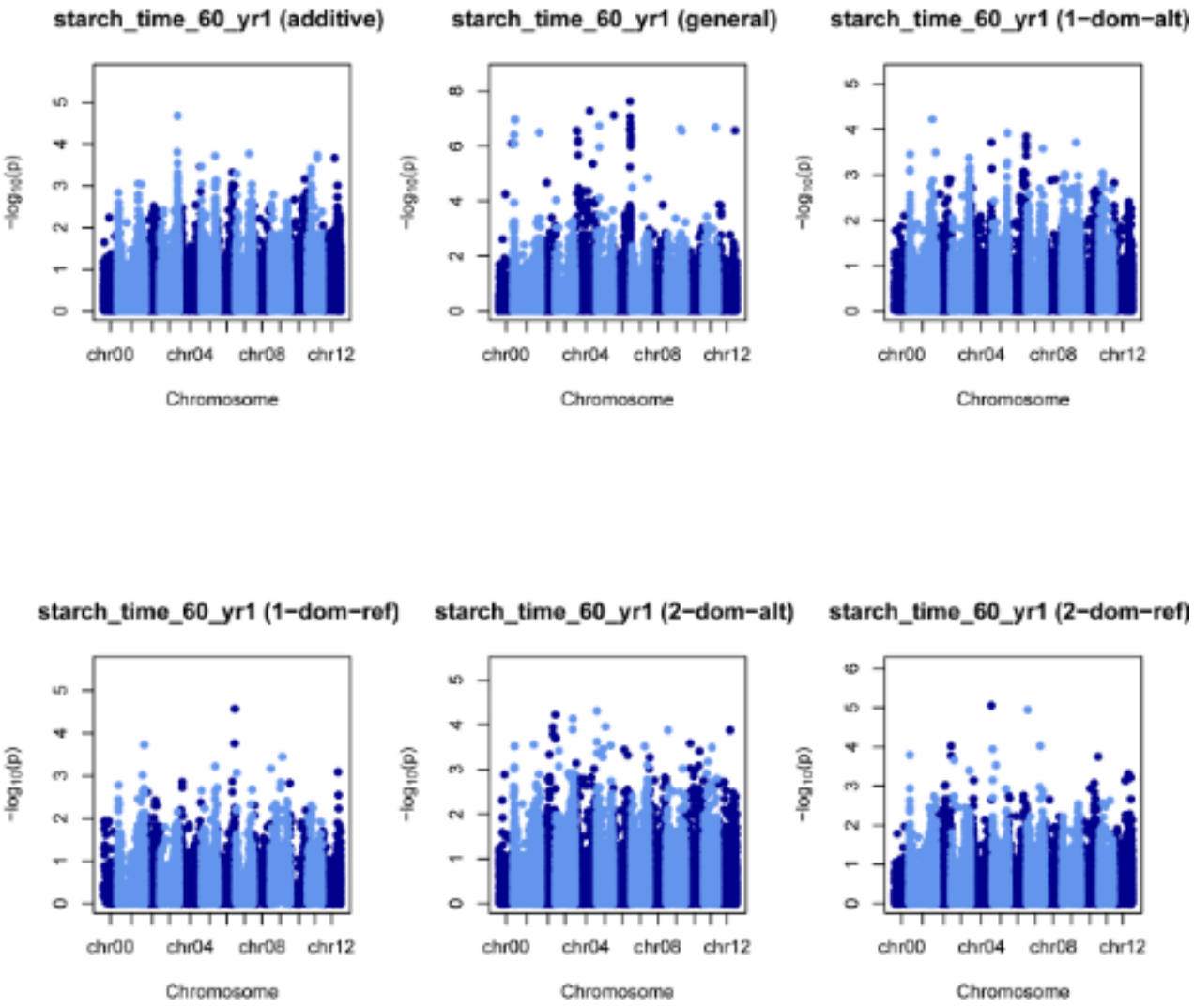


Figure 10: Manhattan plots for starch 'time_60' trait using different GWAS models.

Comparison of the starch digestibility values between the two year 1 field blocks indicated a high degree of replication. (Figure 10)

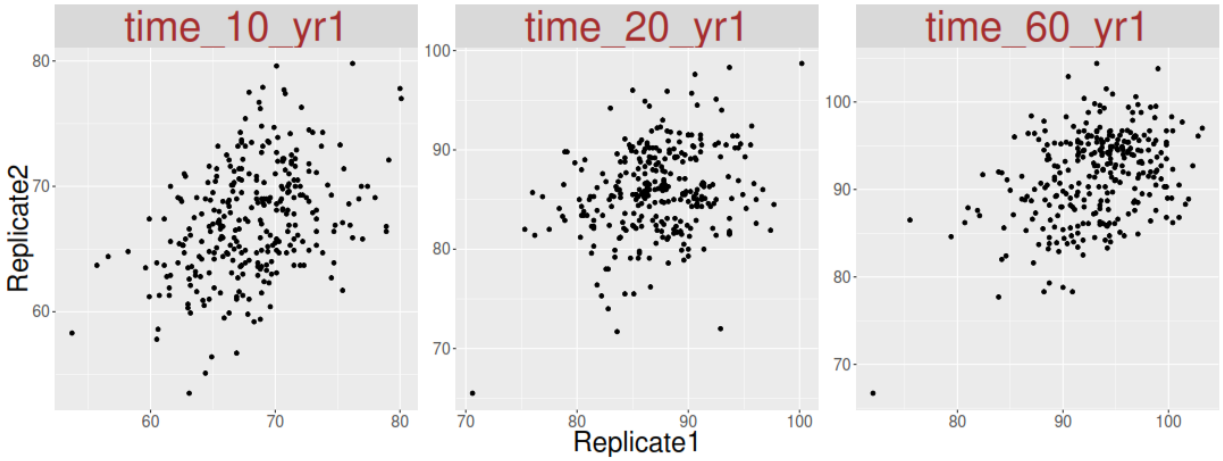


Figure 11 – Comparison of the starch digestibility data between the two field replicates for the year 1 trial. Each data point represents an association panel genotype.

In contrast there was only a weak replication between the year 2 field blocks (Figure 12).

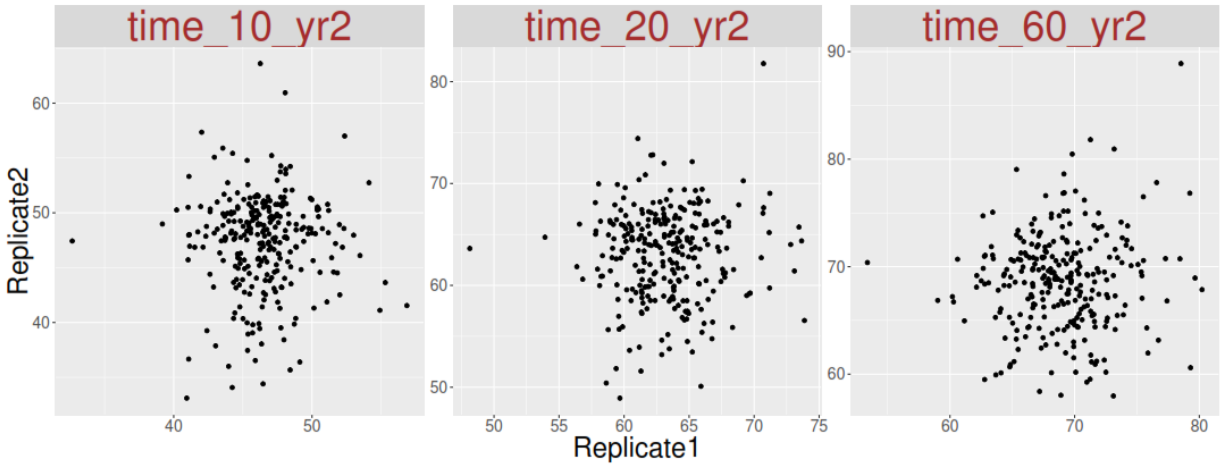


Figure 12 – Comparison of the starch digestibility data between the two field replicates for the year 2 trial. Each data point represents an association panel genotype.

Starch digestibility data were also compared between the field trials conducted in 2013 and 2014, using mean values for each genotype at each time point. There was no significant correlation between the data obtained for each year (Figure 13).

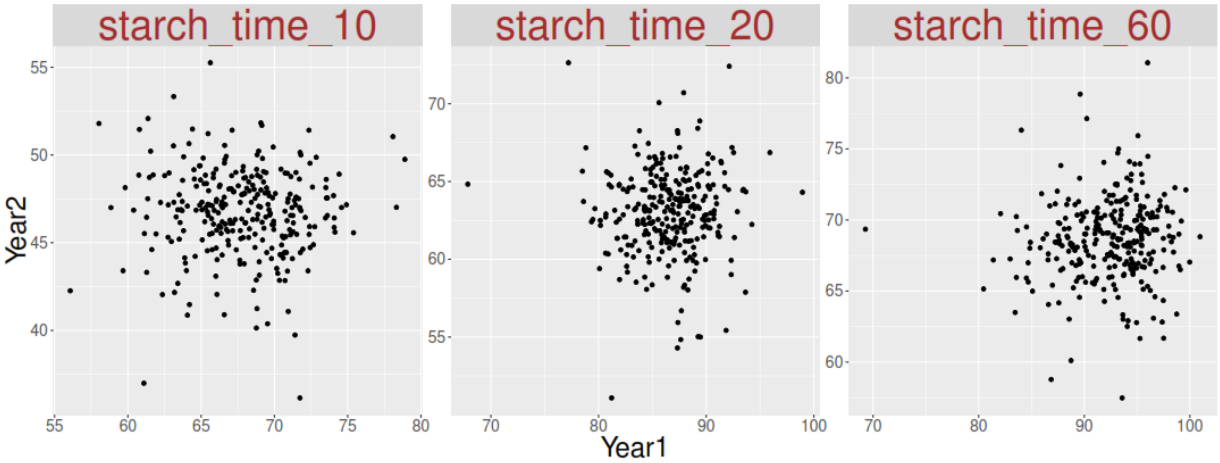


Figure 13 – Comparison of starch digestibility data for 2013 (year1) and 2014 (year2) at 10, 20 and 60 minutes digestion.

Despite the lack of correlations between the replicates and year 1 data, the mean year 2 values were analysed further for marker trait associations.

For the 2014 (year 2) data, the residual plots (Figure 14) from REML analysis were inspected for presence of any outliers and suitability of the data for performing GWAS. The error term (residuals) appeared to be normally distributed for all three time-point starch digestibility traits, thus, suggesting the suitability of the trait data for the GWAS analysis.

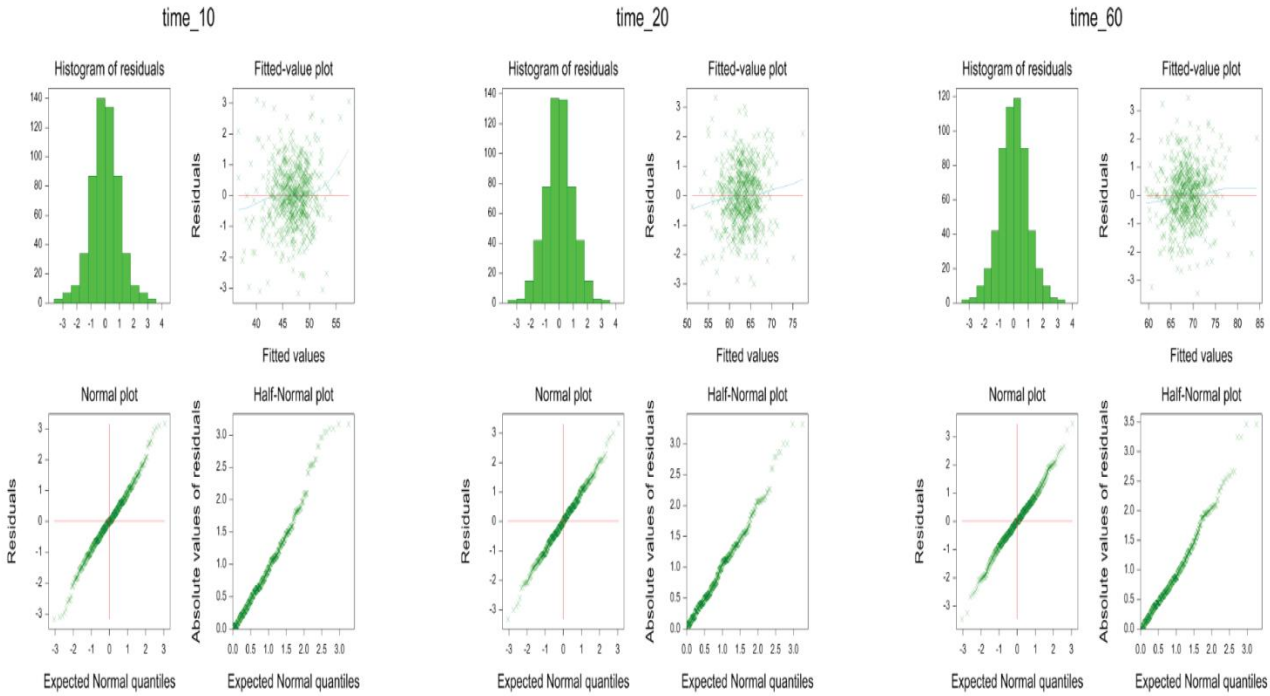


Figure 14 Residual plots for all starch time-point traits from REML analyses 2014 (year2) data.

Figures 15, 16 and 17 illustrate the 2014 (year 2) GWAS results obtained from these analyses using the 8300 SolCAP markers and 44700 SNP markers obtained using a genotyping-by-sequencing approach (Bryan and Sharma unpublished). In contrast to the 2013 results, the only major marker trait association was located on Chromosome 11 at time point 60 minutes (Figure 17).

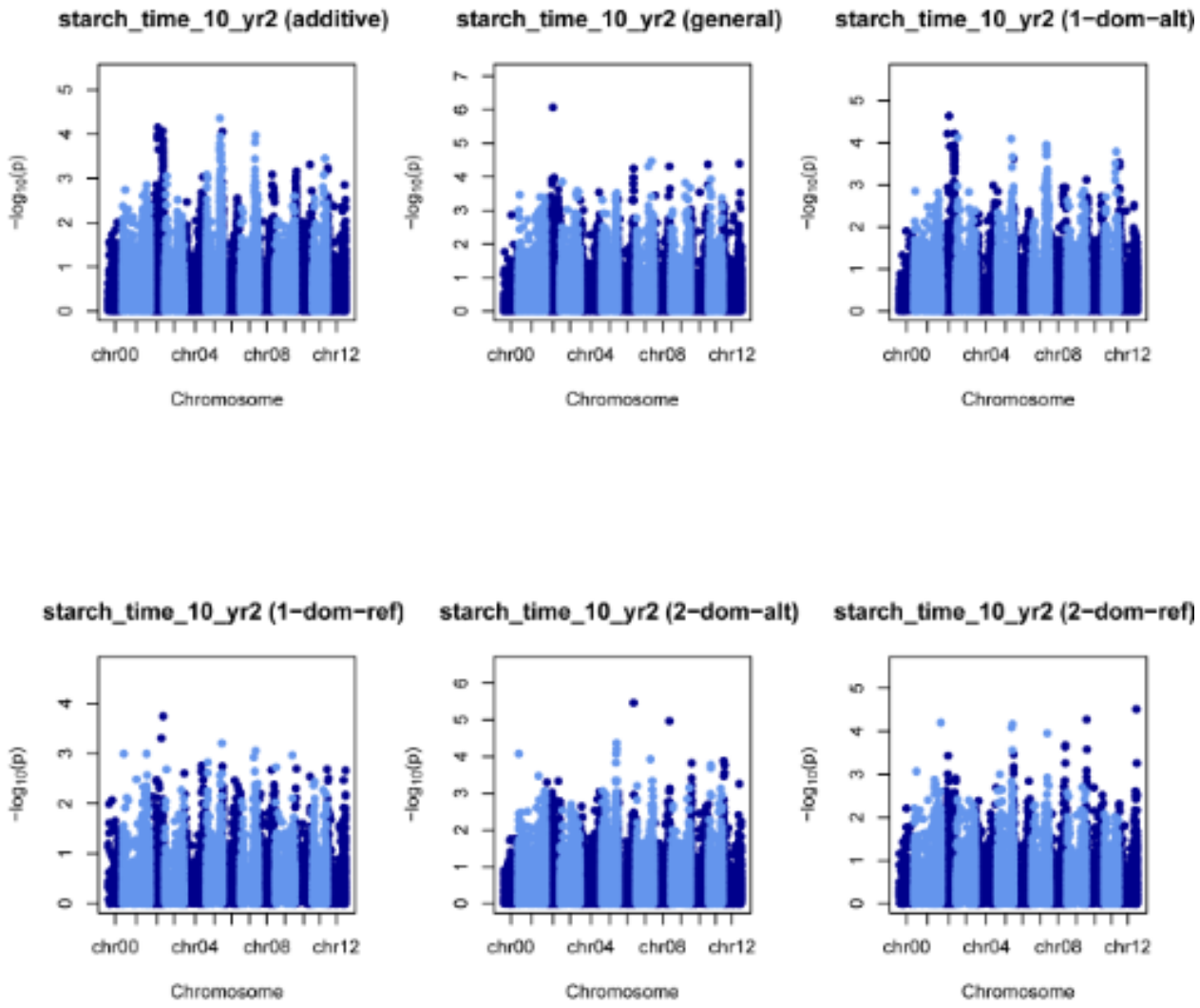


Figure 15: Manhattan plots for starch 'time_10' trait using different GWAS models

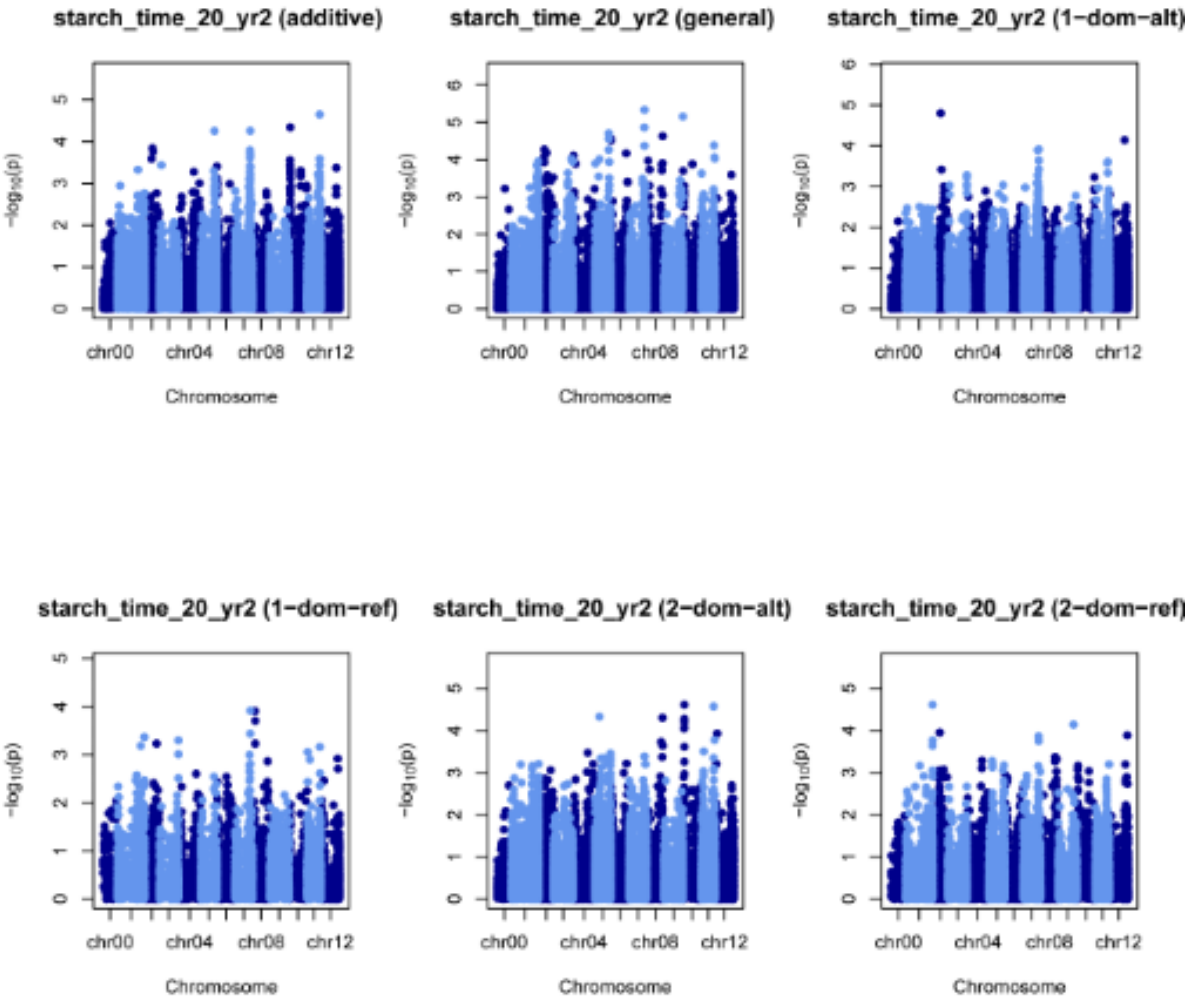


Figure 16: Manhattan plots for starch 'time_20' trait using different GWAS models

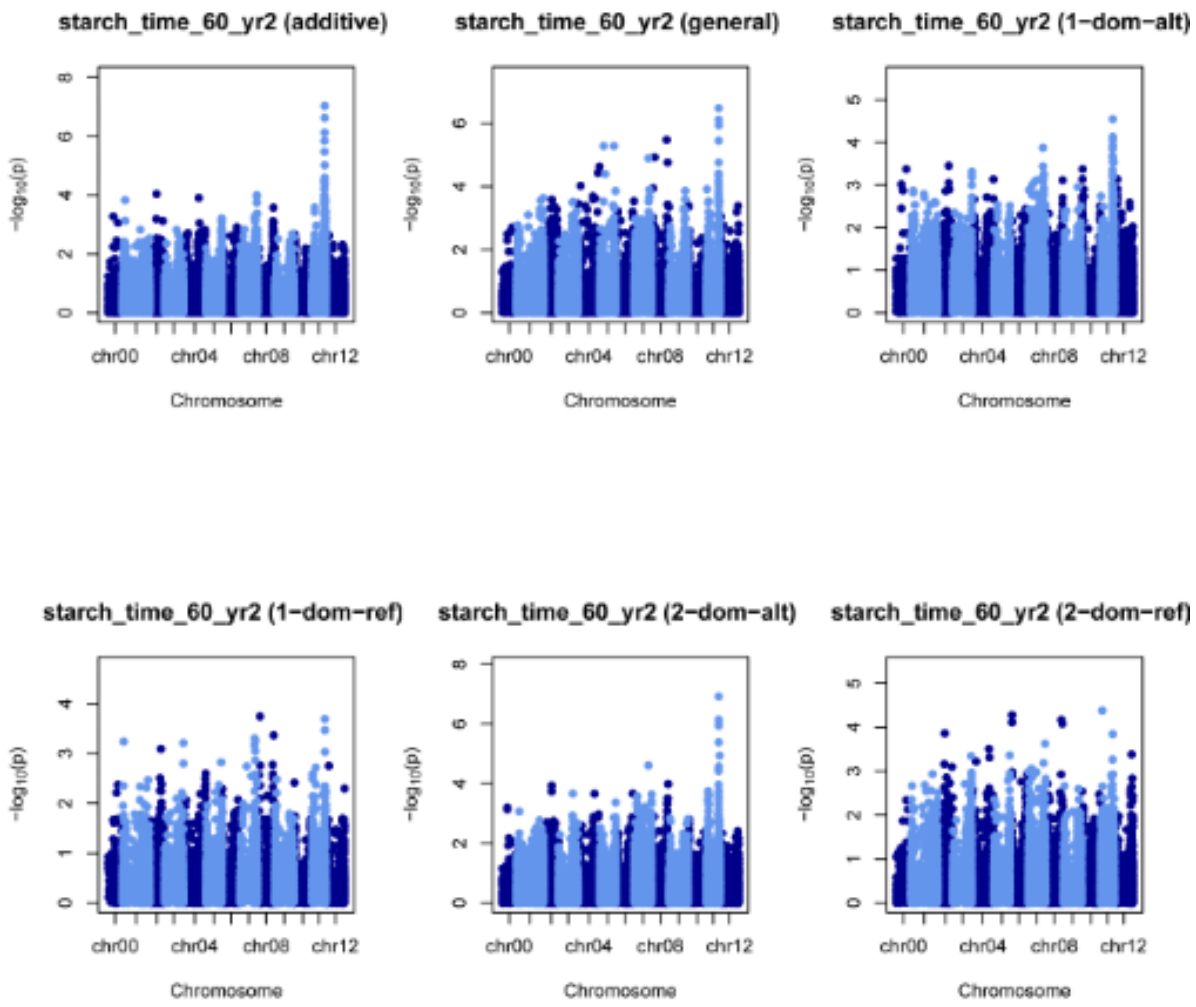


Figure 17: Manhattan plots for starch 'time_60' trait using different GWAS models

Trait	Model	p value threshold	Significant QTL/MTA marker	Chromo-some	Position (bp)	Observed p value
starch_time_20_yr1	general	5.99	St_pooled_SNPs_1186	1	5161801	6.15
starch_time_20_yr1	general	5.99	St_pooled_SNPs_1187	1	5161804	6.24
starch_time_20_yr1	general	5.99	St_pooled_SNPs_1188	1	5161816	6.18
starch_time_20_yr1	general	5.99	St_pooled_SNPs_17698	4	7966239	6.15
starch_time_20_yr1	general	5.99	St_pooled_SNPs_18767	4	48928241	6.82
starch_time_20_yr1	general	5.99	St_pooled_SNPs_22280	5	7239939	6.01
starch_time_20_yr1	general	5.99	St_pooled_SNPs_28257	6	54743002	6.53
starch_time_20_yr1	general	5.99	St_pooled_SNPs_28303	6	55180502	7.93
starch_time_20_yr1	general	5.99	St_pooled_SNPs_28328	6	55225166	7.45
starch_time_20_yr1	general	5.99	St_pooled_SNPs_28440	6	56346819	6.57
starch_time_20_yr1	general	5.99	St_pooled_SNPs_28732	6	57828705	6.11
starch_time_20_yr1	general	5.99	St_pooled_SNPs_49175	12	55947683	6.12
starch_time_60_yr1	general	5.99	St_pooled_SNPs_395	0	43061361	6.09
starch_time_60_yr1	general	5.99	St_pooled_SNPs_397	0	43061365	6.09
starch_time_60_yr1	general	5.99	St_pooled_SNPs_401	0	43061395	6.09
starch_time_60_yr1	general	5.99	St_pooled_SNPs_782	1	2557844	6.42
starch_time_60_yr1	general	5.99	St_pooled_SNPs_793	1	2579803	6.1
starch_time_60_yr1	general	5.99	St_pooled_SNPs_938	1	3350058	6.08
starch_time_60_yr1	general	5.99	St_pooled_SNPs_1186	1	5161801	6.98
starch_time_60_yr1	general	5.99	St_pooled_SNPs_1187	1	5161804	6.95
starch_time_60_yr1	general	5.99	St_pooled_SNPs_1188	1	5161816	6.96
starch_time_60_yr1	general	5.99	St_pooled_SNPs_5997	1	84384128	6.5
starch_time_60_yr1	general	5.99	St_pooled_SNPs_17558	4	6585624	6.57
starch_time_60_yr1	general	5.99	St_pooled_SNPs_17698	4	7966239	6.53
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18036	4	10993507	6.16
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18047	4	11108108	6.15
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18065	4	11182328	6.16
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18066	4	11182333	6.16
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18072	4	11208179	6.13
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18094	4	11386445	6.21
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18095	4	11388802	6.19
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18096	4	11388829	6.19
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18097	4	11388850	6.19
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18104	4	11654109	6.19
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18105	4	11654120	6.21
starch_time_60_yr1	general	5.99	St_pooled_SNPs_18767	4	48928241	7.29
starch_time_60_yr1	general	5.99	St_pooled_SNPs_22280	5	7239939	6.74
starch_time_60_yr1	general	5.99	St_pooled_SNPs_25169	6	2722578	7.13
starch_time_60_yr1	general	5.99	St_pooled_SNPs_25170	6	2722599	7.13
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28257	6	54743002	6.61
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28303	6	55180502	7.63
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28328	6	55225166	7.07
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28440	6	56346819	6.84
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28515	6	56820109	6.04
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28599	6	57219521	6.06
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28618	6	57244831	6
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28657	6	57443580	6.18
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28728	6	57820640	6.29
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28732	6	57828705	6.65
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28736	6	57863499	6.18
starch_time_60_yr1	general	5.99	St_pooled_SNPs_28737	6	57863524	6.4
starch_time_60_yr1	general	5.99	St_pooled_SNPs_38402	9	46179219	6.64
starch_time_60_yr1	general	5.99	St_pooled_SNPs_38779	9	50367631	6.55
starch_time_60_yr1	general	5.99	St_pooled_SNPs_45460	11	37905607	6.68
starch_time_60_yr1	general	5.99	St_pooled_SNPs_49175	12	55947683	6.57
starch_time_10_yr2	general	5.99	St_pooled_SNPs_7997	2	24553720	6.07
starch_time_60_yr2	general	5.99	St_pooled_SNPs_45881	11	40657191	6.12
starch_time_60_yr2	general	5.99	St_pooled_SNPs_45907	11	40760889	6.48
starch_time_60_yr2	additive	6	St_pooled_SNPs_45881	11	40657191	7.03
starch_time_60_yr2	additive	6	St_pooled_SNPs_45907	11	40760889	6.13
starch_time_60_yr2	additive	6	St_pooled_SNPs_45920	11	40815043	6.63
starch_time_60_yr2	2-dom-alt	5.9	St_pooled_SNPs_45881	11	40657191	6.93
starch_time_60_yr2	2-dom-alt	5.9	St_pooled_SNPs_45899	11	40739239	5.95
starch_time_60_yr2	2-dom-alt	5.9	St_pooled_SNPs_45907	11	40760889	6.15
starch_time_60_yr2	2-dom-alt	5.9	St_pooled_SNPs_45920	11	40815043	6.01

Table 1: Significant marker trait associations

5. DISCUSSION

Starch digestion assay

A reliable and reproducible assay has been developed for starch digestion that can be used for high throughput assessment of starch digestibility. We found that it was necessary to wash crude starch samples in 80% hot ethanol in order to remove residual sugars associated with the starch granules prior to conducting the assay. The washing step overcame earlier problems with reproducibility in the assays. Although the assay was reproducible, some problems were encountered with the absolute values of starch digestibility obtained with different batches of enzymes used. To reduce this source of variation it is recommended to prepare large batches of enzyme mix and store aliquots at -20^{0}C until required.

Starch digestibility values

The starch digestibility values indicate considerable variation in this trait between the 300 genotypes tested. For the year 1 (2013) trial for example, when the data are sorted based on the percentage of starch digested after 60 minutes, values range from 69.33% for the variety Shelford to near complete digestion for varieties such as Burren, Argos and Rooster. We expect that this reflects the percentage of starch resistant to digestion, indicating that those with the least amount of starch digested would have tubers with the lowest GI. The variety Almera, marketed on its low GI, has a value of 84.4% of starch digested after 60 minutes. This is by no means the lowest value tested in the 300 varieties, and it is of interest that the variety Shelford (69.33%) gave a particularly low value. The year 2 (2014) trial data were poorly replicated between experimental blocks and with the year 1 data although individual samples gave reproducible results when assayed on different occasions. In this trial the percentage of starch digested after 60 minutes, values range from 57.5% to 81.1%. The lack of reproducibility between field replicates indicates either a large block effect or trivial handling and labelling errors occurred. The good block replication observed in the year 1 dataset coupled with the finding that Almera starch was relatively slowly digested in the year 1 trial but not the year 2 trial leads us to believe the year 1 dataset is more reliable.

Genetic Analysis

Not surprisingly, given the lack of consistency of the digestibility data, the genetic analysis failed to identify consistent markers for starch digestibility over the two seasons. Nevertheless the GWAS scan using ca. 8300 SolCAP markers in combination with 44700 additional SNP markers indicated several strong marker trait associations in the year 1 data, with three markers on chromosome 1, 4 and 6 being particularly strongly associated with the starch digestibility trait. Genes within 1 Mbp of the best scoring marker are listed in Appendix 3. A marker trait association was also identified from the year 2 analysis on chromosome 11 and genes located close to the best scoring SNP are also listed in Appendix 3.

Conclusions

A robust and high throughput assay for assessing starch digestion *in vitro* has been developed and applied to replicate samples from 300 genotypes from 2 field trials conducted in 2013 and 2014. The 2013 data showed good replication between the experimental blocks however this was not seen in the 2014 data and the 2014 data and 2013 datasets were poorly correlated. We assume the 2013 dataset provides better quality data however we cannot exclude a strong environmental effect of the starch digestibility trait. Genetic analysis of year 1 starch samples shows a strong association of starch digestibility with SNP markers in three regions located on Chromosomes 1, 4 and 6. Further work would require another field trial to validate the marker trait associations observed.

6. REFERENCES

AHDB (2013). *Potato Consumer Report.* Available: http://www.potato.org.uk/sites/default/files/%5Bcurrent-page%3Aarg%3A%3F%5D/PCL%20Potato%20Consumer%20september%202013.pdf. [Last accessed 30th April 2014].

Ahuja, G et al. (2013). Starch biosynthesis in relation to resistant starch. In: Shi, Y and Maningat, C.C. *Resistant starch: Sources, Applications and Health Benefits*. London: John Wiley and Sons Ltd. 1-22.

Alberti, K.G.M.M et al. (2005). The metabolic syndrome—a new worldwide definition. *The Lancet*. 366 (9491), 1059–1062.

DEFRA. (2013). *Family Food 2012: UK household food and drink purchases and derived nutrient intakes.* Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/265245/familyfood-2012statsnotice-12dec13.pdf. [Last accessed 30th April 2014].

D'hoop, B.B et al. (2014). Identification of agronomically important QTL in tetraploid potato cultivars using a market-trait association analysis. *Theoretical and Applied Genetics*. 127 (3), 731-748.

Ek, K.L et al. (2012). Glycaemic effect of potatoes. *Food Chemistry*. 133 (4), 1230–1240.

Ek, K.L et al. (2013). Discovery of a low glycemic index potato and relationship with starch digestion *in vitro*. British Journal of Nutrition. 1 (1), 1-7.

Englyst H.N., Kingman S.M. and Cummings J.H. (1992).*Classification and measurement of nutritionally important starch fractions*. European Journal of Clinical Nutrition. 46 (2), 33-50.

Englyst H.N., Kingman S.M., Hudson G.J. and Cummings J.H. (1996).*Measurement of resistant starch in vitro and in vivo. British Journal of Nutrition*. 75, 749-755.

Englyst K.N., Englyst H.N., Hudson G.J., Cole T.J. and Cummings J.H. (1999).*Rapidly available glucose in foods: an in vitro measurement that reflects the glycaemic response.* American Journal of Clinical Nutrition. 69, 448-454.

Felcher, K.J., Coombs, J.J., Massa, A.N., Hansey, C.N., Hamilton, J.P., Veilleux, R.E., Buell, C.R. and Douches, D.S. (2012*). Integration of two diploid potato linkage maps with the potato genome sequence*. PLoS One, *7*(4), p.e36347.

Flint, H.J. (2013). The microbiology of resistant starch fermentation in the human large intestine: A host of unanswered questions. In: Shi, Y and Maningat, C.C *Resistant Starch: Sources, Applications and Health Benefits*. London: John Wiley and Sons Ltd. 251-265.

Hamilton, J.P., Hansey, C.N., Whitty, B.R., Stoffel, K., Massa, A.N., Van Deynze, A., De Jong, W.S., Douches, D.S. and Buell, C.R., 2011. *Single nucleotide polymorphism discovery in elite North American potato germplasm*. BMC genomics, *12*(1), p.1.
Haub, M.D. (2013). Resistant starch on glycaemia and satiety in humans. In: Shi, Y and Maningat, C.C *Resistant Starch: Sources, Applications and Health Benefits*. London: John Wiley and Sons Ltd. 207-214.

Henry, C.J.K. (2005). Glycaemic index values for commercially available potatoes in Great Britain. *British Journal of Nutrition*. 94 (1), 917–921.

Henry, C.J.K et al. (2006). The impact of the addition of toppings/fillings on the glycaemic response to commonly consumed carbohydrate foods. *European Journal of Clinical Nutrition*. 60 (6), 763-769.

International V (2014) GenStat Reference Manual (Release 17), Part 1 Summary. 17 edn. VSN International, Hemel Hempstead, UK
DF. (2014). *About Insulin.* Available: http://www.idf.org/about-insulin-0. [Last accessed 30th April 2014].

Jenkins, D.J.A, et al. (1987). The effect of starch-protein interaction in wheat on the glycaemic response and rate of *in vitro* digestion. *American Journal of Clinical Nutrition*. 45 (1), 946-951.
Kinnear, T et al. (2011). Effect of preparation method on the glycaemic index of novel potato clones. *Food and Function*. 2 (1), 438-444.

Maningat, C.C and Seib, P.A. (2013). RS4-type resistant starch: Chemistry, functionality and health benefits. In: Shi, Y and Maningat, C.C *Resistant Starch: Sources, Applications and Health Benefits*. London: John Wiley and Sons Ltd. 111-130.

Megazyme International.(2011). *D-GLUCOSE ASSAY PROCEDURE (GOPOD-FORMAT).* Available: http://secure.megazyme.com/files/BOOKLET/K-GLUC_1107_DATA.pdf. [Last accessed 14th May 2014].

Mishra, S et al. (2012). Food structure and carbohydrate digestibility. In: Chang, C *Carbohydrates - Comprehensive Studies on Glycobiology and Glycotechnology*. Winchester: InTech. 289-316.

Monro, J et al. (2009). Potato genotype differences in nutritionally distinct starch fractions after cooking, and cooking plus storing cool. *Journal of Food Composition and Analysis*. 22 (6), 539–545.

Muazu, J et al. (2011). Extraction and characterization of Kaffir Potato Starch: A potential. *Journal of Natural Product and Plant Resources*. 1 (2), 41-49.
Nielsen, S.S. (2010). Phenol-sulphuric Acid Assay for Total Carbohydrates. In: Nielsen, S.S *Food Analysis Laboratory Manual*. 2nd ed. New York: Springer. 47-53
Riley, H. (2010). Potato consumption in the UK – why is ‘meat and two veg’ no longer the traditional British meal? *Nutrition Bulletin*. 35 (4), 320-331.

Rosyara U.R., De Jong W.S., Douches D.S., Endelman J.B. (2016) *Software for Genome-Wide Association Studies in Autopolyploids and Its Application to Potato*. The Plant Genome 9 (2): doi:10.3835/plantgenome2015.08.0073

Singh, J et al. (2010). Starch digestibility in food matrix: a review. *Food Science and Technology*. 21 (1), 168-180.

Sharma, S.K., Bolser, D., de Boer, J., Sønderkær, M., Amoros, W., Carboni, M.F., D'Ambrosio, J.M., de la Cruz, G., Di Genova, A., Douches, D.S. and Eguiluz, M. (2013). *Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps*. G3: Genes| Genomes| Genetics, *3*(11), pp.2031-2047.

Voorrips, R.E., Gort, G. and Vosman, B. (2011). *Genotype calling in tetraploid species from bi-allelic marker data using mixture models.* BMC bioinformatics, *12*(1), p.1.

Waigh, T.A. et al (1997) "Analysis of the native structure of starch granules with X-ray microfocus diffraction." Macromolecules*.* 30 (13), 3813-3820.
WHO.(2013). *Diabetes.* Available: http://www.who.int/mediacentre/factsheets/fs312/en/. [Last accessed 30th April 2014].
Xu, X. et al. (2011). Genome sequence and analysis of the tuber crop potato. *Nature*. 475 (1), 189–195.

Yu, J., Pressoir, G., Briggs, W.H., Bi, I.V., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B. and Kresovich, S.(2006). *A unified mixed-model method for association mapping that accounts for multiple levels of relatedness*. Nature genetics, *38*(2), pp.203-208.

Zhang, G and Hamaker.B.R. (2013).Slowly digestible starch and health benefits. In: Shi, Y and Maningat, C.C *Resistant Starch: Sources, Applications and Health Benefits*. London: John Wiley and Sons Ltd. 111-130.
Zhu, C et al. (2008). Status and Prospects of Association Mapping in Plants. *The Plant Genome*. 1 (1), 5-20.

7. APPENDICES

Appendix1 – Tuber dry matter values for year 1 samples
Appendix 2 – Best Linear Unbiased Estimates (BLUEs) for ‘time_10’, ‘time_20’ and ‘time_60’ digestion data
Appendix 3 – Candidate gene lists within 1Mbp of the most strongly associated markers

8. ACKNOWLEDGEMENTS

The financial support from AHDB and Greenvale AP is gratefully acknowledged

The scientific team for this project was:
Lara Hilley, Laurence Ducreux, Sanjeev Kumar Sharma, Glenn Bryan and Mark Taylor, James Hutton Institute

Harry Flint and Silvia Gratz Rowett Institute of Nutritional Health.