GWAS reveals the genetic complexity of fructan accumulation patterns in barley grain

Andrea Matros[1]*, Kelly Houston[2], Matthew R. Tucker[3], Miriam Schreiber[2], Bettina Berger[4], Matthew K. Aubert[3], Laura G. Wilkinson[3], Katja Witzel[5], Robbie Waugh[2,3], Udo Seiffert[6], Rachel A. Burton[1]

[1]ARC Centre of Excellence in Plant Energy Biology, School of Agriculture, Food and Wine, University of Adelaide, Adelaide, South Australia, Australia;
[2]Cell and Molecular Sciences, The James Hutton Institute, Dundee, Scotland, UK;
[3]School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Urrbrae, SA, Australia
[4]Australian Plant Phenomics Facility, The Plant Accelerator, School of Agriculture, Food and Wine, University of Adelaide, Adelaide, South Australia, Australia;
[5]Leibniz Institute of Vegetable and Ornamental Crops, Großbeeren, Brandenburg, Germany
[6]Biosystems Engineering, Fraunhofer IFF, Magdeburg, Saxony-Anhalt, Germany

Running title: GWAS for fructan profiles in two-row spring barley grain

E-Mail addresses:
andrea.matros@adelaide.edu.au
kelly.houston@hutton.ac.uk
matthew.tucker@adelaide.edu.au
miriam.schreiber@hutton.ac.uk
bettina.berger@adelaide.edu.au
matthew.aubert8@gmail.com
laura.wilkinson@jic.ac.uk
witzel@igzev.de
robbie.waugh@hutton.ac.uk; robbie.waugh@adelaide.edu.au
udo.seiffert@iff.fraunhofer.de
rachel.burton@adelaide.edu.au

***Corresponding author:**

Andrea Matros

ARC Centre of Excellence in Plant Energy Biology, School of Agriculture, Food and Wine, University of Adelaide, Adelaide, South Australia, Australia

andrea.matros@adelaide.edu.au

Date of submission: 29/06/2020

Figures number: 6

Tables number: 4

Supplementary data: 15 (5 figures, 8 tables, one Suppl. Method, one Suppl. Result)

Words count: 6435 (from introduction until end of discussion)

GWAS reveals the genetic complexity of fructan accumulation patterns in barley grain

Highlight:

Grain fructan profiles in barley are more complex than previously expected and variations in a diversity panel relate to a genomic region where fructan biosynthesis genes cluster.

Abstract

We profiled the grain oligosaccharide content of 154 two-row spring barley genotypes and quantified 27 compounds, mainly fructans, that exhibited differential abundance. Clustering revealed two major profile groups where the 'high' set contained greater amounts of sugar monomers, sucrose and overall fructans, but lower fructosylraffinose. GWAS identified a significant association for the variability of two fructan types; neoseries-DP7 and inulin-DP9 which showed increased strength when a compound-ratio GWAS was applied. Gene models within this region included five fructan biosynthesis genes, of which three (*fructan:fructan 1-fructosyltransferase*, s*ucrose:sucrose 1-fructosyltransferase*, and s*ucrose:fructan 6-fructosyltransferase)* have already been described. The remaining two, *6(G)-fructosyltransferase* and *vacuolar invertase1* have not previously been linked to fructan biosynthesis in barley and showed expression patterns distinct from those of the other three genes, including exclusive expression of *6(G)-fructosyltransferase* in outer grain tissues at the storage phase. From exome capture data several SNPs related to inulin- and neoseries-type fructan variability were identified in *fructan:fructan 1-fructosyltransferase* and *6(G)-fructosyltransferase* genes Co-expression analyses uncovered potential regulators of fructan biosynthesis including transcription factors. Our results provide evidence for the distinct biosynthesis of neoseries-type fructans during barley grain maturation plus new gene candidates likely involved in the differential biosynthesis of the various fructan types.

Keywords: fructans, barley, grain, neoseries, 6G-FFT, oligosaccharides, GWAS, ratio-GWAS, expression analysis

Abbreviations:

1-FFT:	fructan:fructan 1-fructosyltransferase
1-SST:	sucrose:sucrose 1-fructosyltransferase
6-SFT:	sucrose:fructan 6-fructosyltransferase
6G-FFT:	6(G)-fructosyltransferase
DAP:	days after pollination
DP:	degree of polymerisation
DM:	dry matter
ELSD:	evaporative light scattering detection
FDR:	false discovery rate
FOS:	fructooligosaccharides
FPKM:	fragments per kilobase, per million mapped reads
GWA:	genome wide association
GWAS:	genome wide association study
HAI:	hours after imbibition
HPAEC–PAD:	high pH anion exchange chromatography with pulsed amperometric detection
HPLC:	high performance liquid chromatography
KP:	kestopentaose
KT:	kestotetraose
LC:	liquid chromatography
LD:	linkage disequilibrium
LOD:	logarithm of odds
MAF:	minimum allele frequency
MS:	mass spectrometry
NG:	Neural Gas
NS:	neoseries-type fructan
P:	probability value
PEG:	polyethylene glycol
QTL:	quantitative trait loci
RFO:	raffinose family oligosaccharides
RT:	retention time
SNP:	single nucleotide polymorphisms
SPE:	solid phase extraction

TFA:	trifluoroacetic acid
TPM:	transcripts per million
VI-1:	vacuolar invertase1

Introduction

Starch, fructans and (1,3; 1,4)-β-glucans represent the major plant reserve carbohydrates (Vijn and Smeekens, 1999; Burton and Fincher, 2009). Among them, fructan biosynthesis has evolved polyphyletically in about 15% of higher plants, including species of the orders Asterales, Buxales, Asparagales and Poales (Hendry and Wallace, 1993; Cairns *et al.*, 2000; Van den Ende, 2013). In cereals, fructans accumulate in all plant organs (Pollock and Cairns, 1991).

Fructans consist of repeating fructose residues linked to a sucrose unit. The classification relates to the position of the sucrose, the linkage-type between the fructose residues (i.e. ß(2,1), inulin; ß(2,6), levan; or containing both ß(2,1) and ß(2,6)-d-fructosyl units referred to as graminan-type) and the chain lengths (Cochrane, 2000; Matros *et al.*, 2019). Fructans can form oligomers with a degree of polymerization (DP) of 3-9 or polymers with a DP ≥10. Here, fructans is used to indicate either fructooligosaccharides (FOS) or fructan polymers. Fructans are typically discussed in the literature without differentiation of the DP, but since they have become more important in a dietary context (Dwivedi *et al.*, 2014; Verspreet *et al.*, 2015b; Liu *et al.*, 2017) more attention has recently been paid to the role of fructans according to their DP level.

All types of fructans are known to occur in the Poaceae (Carpita *et al.*, 1991; Pollock and Cairns, 1991; Bonnett *et al.*, 1997). However, *Triticum*, *Secale* and *Hordeum* are believed to mainly contain branched-type fructans (graminan-type) whereas the Poeae tribe mostly comprises levan-type fructans (Bonnett *et al.*, 1997; Huynh *et al.*, 2008a). Recently, the presence of graminan- and neoseries-type fructans was reported in wheat (Verspreet *et al.*, 2015c). Neoseries-type fructans, in contrast to other fructan-types, are characterised by an internal glucose unit (Matros *et al.*, 2019). Additional structural variations are likely to occur between different plant organs.

New developments in fructan analysis based on mass spectrometry (MS) detection revealed the fine structure of cereal grain fructans with DP3-5 (Verspreet *et al.*, 2017). Variations in fructan composition pattern and abundance were observed in oat, barley, rye, spelt and wheat

flour, suggesting a putative link between accumulation of certain fructan types and cereal phylogeny (Verspreet, *et al.*, 2017).

Reports of the beneficial health effects of fructans (Verspreet *et al.*, 2015b; Liu *et al.*, 2017; Anrade *et al.*, 2019) have prompted screens for variation in their natural abundance and composition and biotechnological approaches to increase FOS content in classical non-fructan cereals, such as maize (Dwivedi *et al.*, 2014). However, most studies on grain fructan content still focus on wheat (Huynh *et al.*, 2008a and b; Veenstra *et al.*, 2017; Veenstra *et al.*, 2019). Investigation of two doubled haploid (DH) populations (Berkut x Krichauff and Sokoll x Krichauff) revealed several quantitative trait loci (QTL) for high fructan content in wheat grain (Huynh *et al.*, 2008b). Winter wheat grain fructan content was found to be significantly influenced by either the genotype or the environment as well as by genotype × environment interactions (Veenstra *et al.*, 2019). Fructan content in developing barley grain was compared between seven genotypes, demonstrating peak accumulation between 6 and 17 days after pollination (DAP) (De Arcangelis *et al.*, 2019) as previously reported (Peukert *et al.*, 2014). Notably, a comparative mapping approach involving wheat and barley revealed clusters of genes encoding fructan biosynthesis enzymes (Huynh *et al.*, 2012) on 7AS in wheat and 7HL in barley. These clusters included *sucrose:sucrose 1-fructosyltransferase* (*1-SST*), *fructan:fructan 1-fructosyltransferase* (*1-FFT*), *sucrose:fructan 6-fructosyltransferase* (*6-SFT*), and several vacuolar invertases. Similar gene structures and physical positions of these clusters of functionally related genes in both genomes indicate that they may have evolved in parallel and that the genes within a cluster may be linked functionally in controlling fructan accumulation.

Due to its increasing potential as a health-promoting functional cereal, there is considerable interest in identifying factors that influence barley grain quality (Meints *et al.*, 2016; Langridge and Waugh, 2019). Here we report on an analysis of natural variation in fructan content and composition across a diversity panel of two-row spring barley. We identified significant associations between fructan composition/content and fructan biosynthesis genes. We obtained support for the involvement of some of these in underpinning the observed variation from transcriptomic analysis. Additionally, potential regulators of fructan biosynthesis were assigned by co-expression analyses.

Materials and methods

Plant material

We used 154 two-row spring barley genotypes sourced from The James Hutton Institute, complemented by three Australian elite barley varieties and the wheat line Piccolo as checks (Table S1). The germplasm was selected for minimum population structure while maintaining as much genomic diversity as possible based on principle components analysis of a much larger set of genotypes (>800). Three plants per genotype (biological replicates) were grown in a randomised main-unit design in a glasshouse compartment in a mix of clay-loam and cocopeat (50:50 v/v) and day/night temperatures of 22˚C/15˚C between July and December 2014 in The Plant Accelerator, Adelaide, Australia. Mature grains were harvested and stored until oligosaccharide analysis. For each sample, five grains were ground together to a fine powder using a PowerLyzerTM24 Homogenizer (QIAGEN) and used for oligosaccharide analysis immediately.

Oligosaccharide extraction and profiling

A 'mixed sample' was assembled composed of equal amounts from each individual sample (154 genotypes x three biological replicates) to capture systematic shifts during extraction and measurement. Soluble sugars were extracted following a method adapted from Verspreet *et al.* (2012) by incubation in 80% ethanol at 85°C for 30 min followed by Milli-Q water at 85°C for 30 min on a mixer (700 rpm) in a final dilution of 1:40 (w/v, mg/µl), and supernatants combined. Extracts were diluted with water to 1:1000 (w/v, mg/µl) and 25 µl per sample analysed by high pH anion exchange chromatography with pulsed amperometric detection (HPAEC–PAD) on a Dionex ICS-5000 system using a DionexCarboPAC^{TM}PA-20 column (3 x 150 mm) with a guard column (3 x 50 mm) kept at 30°C and operated at a flow rate of 0.5 ml min^{-1}. The eluents used were (A) 0.1 M sodium hydroxide and (B) 0.1 M sodium hydroxide with 1 M sodium acetate. The gradient used was: 0% (B) from 0-2 min, 20% (B) from 2-35 min, 100% (B) from 35-36.5 min, 0% (B) from 37.5-38.5 min. Detector temperature was maintained at 20°C, data collection was at 2 Hz and the Gold Standard PAD waveform (std. quad. potential) was used.

Data acquisition, processing, and peak integration were performed using the Chromeleon™ version 7.1.3.2425 software (Thermo Scientific). Compounds were annotated based on available analytical standards. Glucose, fructose, sucrose, raffinose, 1-kestose, maltose, maltodextrin, nystose and mixtures of inulin from chicory (DP2-60) and levan from *Erwinia*

herbicola were purchased from Sigma-Aldrich, while 1,1,1-kestopentaose was obtained from Megazyme. Additional inulin and neoseries-type fructans were isolated from onions and barley grain and analysed by mass spectrometry (MS). Fructan-related chromatographic peaks were identified based on fructanase digestion and mild acid hydrolysis (Supplementary Methods). A total of 27 peaks were annotated (Table S2).

Metabolic data analyses

Peak area entry means and variances with respective standard deviations were calculated in Excel 2007 (Microsoft) from the 'mixed sample'-normalised integrated peak area values of the individual biological replicates for each two-row spring barley line and the check lines (Table S3). Data from at least three biological replicates were available for 143 lines. For ten lines (Agenda, Alliot, Appaloosa, Cellar, Drought, Goldie, Scarlett, Tankard, Tartan, Turnberry), data from two biological replicates were available. The mean values for the two lines with just one entry (Calgary and Saana) were replaced by the only available data. Bonferroni outlier test was performed and pair-wise correlations between the abundances of the 27 metabolites were revealed by applying the average linkage clustering method, based on Pearson correlation coefficients implemented in the MVApp (Julkowska *et al.*, 2019; http://mvapp.kaust.edu.sa/MVApp/). The metabolite abundances were analysed with the software package MATLAB (The MathWorks, Inc.) with a log-logistic distribution applied. The Neural Gas (NG) algorithm, implemented in MATLAB was applied for cluster analysis following Kaspar-Schoenefeld *et al.* (2016) and Peukert *et al.* (2016). Analyses were performed for the biological replicates individually and the number of NG clusters was set to four.

GWAS

GWAS was carried out by combing the phenotypic data for the 154 barley accessions with genotypic data generated using the Barley 50K iSelect genotyping platform (Bayer *et al.*, 2017). We focused on two-row spring barley accessions to reduce the confounding effects of population structure (Comadran *et al.*, 2012) that could have been introduced by including other row types and growth habits (Darrier *et al.*, 2019). Prior to analysis any single nucleotide polymorphism (SNP) with a minimum allele frequency (MAF) of < 0.05 was removed which left 24,925 polymorphic markers for our analysis. Marker-trait association analysis was carried out using R 2.15.3 (http://www.R-project.org) and performed with a compressed mixed linear model (Zhang *et al.*, 2010) implemented in the GAPIT R package

(Lipka *et al.*, 2012). Linkage disequilibrium (LD) was calculated across the genome between pairs of markers using a sliding window of 500 markers and a threshold of $R^2 < 0.2$ using Tassel v 5 (Bradbury *et al.*, 2007) to identify local blocks of LD, facilitating a more precise delimitation of quantitative trait loci (QTL) regions. We anchored regions of the genome containing markers that had passed the Benjamini-Hochberg threshold ($p < 0.05$) as implemented in GAPIT to the barley physical map (Mascher *et al.*, 2017) using marker positions provided in Bayer *et al.* (2017) and then expanded this region using local LD derived from genome wide LD analysis as described above. Putative QTL represented by less than 5 SNPs with -log10(p) values < 3 were not considered to be robust given the marker density and extensive LD present in the barley genome (Mascher *et al.*, 2017). The SNP with the highest LOD score was used to represent a significant QTL. We investigated significantly associated regions using BARLEX (https://apex.ipk-gatersleben.de/apex/f?p=284:39) to identify putative candidate genes. Gene annotations refer to entries in the UniProt database (https://www.uniprot.org/uniprot/, June 2019). Unknown genes were searched against the non-redundant entries for plants (June 2019) in the NCBI database using the BLASTX 2.9.0+ software (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

For compounds showing an association that passes the false discovery rate (FDR) calculated in GAPIT, ratios between these and all other compounds quantified were generated. The ratios were log transformed and then used to carry out further GWAS. The 'p-gain', defined as the ratio of the lowest p-value of the two individual metabolites and the p-value of the metabolite ratio (Petersen *et al.*, 2012) was then calculated. A critical value for the p-gain was derived using $B/(2*\alpha)$, where α is the level of significance (0.05) and B the number of tested metabolite pairs. As we tested fifty-two pairs of compounds our critical value threshold was 5.2×10^2.

Publicly available exome capture datasets (Mascher *et al.*, 2017) were used to identify potential causal polymorphisms in candidate genes. We only considered non-synonymous SNPs with less than 10% missing data across the set of germplasm to be informative.

Gene transcript expression analyses of various developmental stages and tissues

Transcript abundance of genes of interest were measured in whole germinated grain (mean data from genotypes Navigator and Admiral) and isolated Navigator grain tissues from 0 to 96 hours after imbibition (hai). Aleurone tissues were divided into approximately thirds, with the proximal aleurone closest to the embryo (Betts *et al.*, 2019).

Data for seedling tissues (germinated embryo, root, and shoot) were obtained from the Expression Atlas organ dataset (https://www.ebi.ac.uk/gxa/home).
Data from epidermal strips (4 weeks after sowing, W4), roots (W4), inflorescences, rachis (W5), inflorescences, lemma, inflorescences, lodicule, dissected inflorescences, palea (W6), inflorescence (10 mm), and internode, as well as for senescing leaf were obtained from BARLEX (https://apex.ipk-gatersleben.de/apex/f?p=284:10) (Colmsee *et al.*, 2015).
A developing anther dataset was obtained from Barakate *et al.* (2020) covering four anther stages (premeiosis, leptotene/zygotene, metaphase I to tetrad, pachytene/diplotene) and two meiocyte stages (leptotene/zygotene, pachytene/diplotene). Raw expression data were mapped against the transcriptome of barley (https://webblast.ipk-gatersleben.de/barley_ibsc/downloads/; merging high-confidence and low-confidence transcripts as well as isoforms) using Salmon v14.0 (Patro *et al.*, 2007).
RNA-sequencing data were obtained from developing pistils at Waddington (W) stages W8, W8.5, W9, W9.5 and W10 (Wilkinson *et al*., 2019) and are shown as a mean value for five genotypes including Golden Promise (1x replicate per stage; Aubert *et al*., 2018), Salka, Wren, Forum and Gant (2x replicates per stage). In addition, RNA-sequencing data from individual pistil tissues including the nucellus, integuments, ovary wall, embryo sac, egg apparatus and central cell, antipodal cells, and chalaza were analysed from the Sloop genotype.
Gene transcript expression data for whole developing grain (from 7 to 20 DAP) minus the embryo were generated by RNA-sequencing and are shown as a mean value from 6 genotypes including Sloop (1x replicate per timepoint; Aubert *et al*., 2018), Alabama, Pewter, Extract, Taphouse, and Hopper (1x replicate per timepoint), while isolated developing grain tissues of interest including the pericarp, aleurone, sub-aleurone, and starchy endosperm were generated from medial sections at 7 to 25 DAP for the genotype Sloop (1x replicate per timepoint).

Correlation analyses of gene transcript expression

Correlations among transcript abundance of fructan metabolism genes with other gene models from the QTL interval detected in the GWAS were evaluated for each of the RNA-sequencing datasets listed above, individually. Pair-wise correlations between the gene transcript expression levels were revealed by applying the average linkage clustering method, based on Pearson correlation coefficients implemented in the MVApp (Julkowska *et al.*, 2019).

Results

Grain oligosaccharide profiling revealed the abundance of fructans

HPAEC-PAD chromatograms of non-structural soluble carbohydrates from mature barley grain allowed for separation of monosaccharides, disaccharides, and oligosaccharides with a DP <15 (Table 1, Figure S1). Among the latter, we identified two raffinose family oligosaccharides (RFO) and three maltose-type oligosaccharides. Most compounds were found to be fructans (Table S2), including levan-, inulin-, graminan, and NS-inulin-types. A high abundance of fructans with DP3 and DP4 was observed in mature grain extracts. Oligosaccharide profiles were obtained and evaluated from all 154 two-row spring barley lines and four checks (Table S1) and integrated peak areas extracted for 27 compounds (Table 1). The resulting data matrix was used for analyses of oligosaccharide distribution, abundance variation, metabolite correlations, and GWAS.

Large variations detected in oligosaccharide profiles

The abundance of most compounds followed a log-logistic distribution. Only the two most abundant compounds, sucrose and raffinose, followed a normal distribution (Figure S2). Oligosaccharide profiles were grouped separately for each of the biological replicates. Applying Neural Gas (NG) clustering to the data identified four statistically significant patterns of abundance (clusters). Each cluster can be interpreted as a prototypic abundance profile of the underlying metabolite values (Figure S3). They mainly differed in height of the normalised peak areas. Overlaps between cluster 1 and 4 and cluster 2 and 3 were detected. Samples in the latter clusters were characterised by significantly higher levels of sugar monomers and sucrose, lower fructosylraffinose and higher overall fructan values compared to clusters 1 and 4 samples (Figure S3). Accordingly, we rationalised the four clusters into two profile groups (Figure1); cluster 1 and 4 forming profile group 1 (‘low’) and cluster 2 and 3 forming profile group 2 (‘high’). The largest peak in each sample was sucrose, whilst among the oligosaccharides, the highest values were detected for raffinose, the co-eluting fructans 1-kestose/6-kestose, nystose and the co-eluting 1&6-kestotetraose (KT, bifurcose)/6G&1-KT (NS-DP4). Generally, a higher abundance of fructans with DP3 and DP4 was observed for all accessions, and differentiation between individuals was mainly attributed to the overall abundance of all fructan types.

We then assigned individual barley accessions to profile groups according to abundance profiles in each individual replicate (Table S1). Accessions with only two biological

replicates and mixed representation of their clusters in the two profile groups were assigned as 'mixed', as their oligosaccharide profile group was not distinct. In total 76, 77, and 5 accessions were assigned to the profile groups ‘low’, ‘high’, and ‘mixed’ respectively.

Significant correlations are observed between metabolites

Of the 349 pair-wise correlations, 184 (52.72%) were highly significantly correlated ($p < 0.001$), 204 (58.45%) moderately significantly ($p < 0.01$) and 228 (65.33%) were just significant ($p < 0.05$) (Table S4). Several regions with highly correlated metabolites were identified in the results matrix, reflecting in many cases, biochemical relationships (Figure 2). The most significant positive correlations were observed between the various branched neoseries-type fructans as well as between the linear inulin-type fructans. Significant positive correlations were also detected between the monosaccharides and their related disaccharides as well as between the maltose-type oligosaccharides. The most significant negative correlations were detected between fructosylraffinose and fructans being highest with nystose (Figure 2).

Differences in grain oligosaccharide profiles are genetically controlled in barley

GWAS for variation in mature grain oligosaccharides identified a single highly significant association for two compounds, neoseries-DP7 (LOD = 8.65, $p = 2.25 \times 10^{-9}$) and inulin-DP9 (LOD = 6.74, $p = 1.81 \times 10^{-7}$), with other less significant associations for both of these on chromosome 7H (Figure 3A). Regression of these two traits showed a high level of correlation ($R^2 = 0.86$, Figure 2, Table S4). Both QTL on 7H overlapped and the most significant marker from the analysis was the same, JHI-Hv50k-2016-438638 (Table 2A). This marker had adjusted p-values after FDR correction of $p = 0.00004$ for neoseries-DP7 and $p = 0.003$ for inulin-DP9. We anchored this QTL to the physical map (Mascher *et al.*, 2017), which based on local LD spans 3.88 MB from 174,327 (JHI-Hv50k-2016-435062) to 4,056,691 bases (JHI-Hv50k-2016-439312).

In total 194 gene models were detected within the QTL, of which 65 are unannotated (Table S5). The highest number of annotated gene models was involved in protein modification (32) and degradation (17) (Figure 3B) with others involved in transcription/translation (27), lipid/sterol/terpenoid metabolism (12), transcription factors (TFs) (11), or carbohydrate metabolism (11). Among the latter category, we identified five candidates that could influence fructan content (Table 2B). These included HORVU7Hr1G000250.3, HORVU7Hr1G000260.2, and HORVU7Hr1G001040.6, which were genetically similar or

identical to *1-FFT*, *1-SST* and *6-SFT* from *Hordeum vulgare*, respectively. Of the two others, HORVU7Hr1G000270.1 was similar to *6(G)-fructosyltransferase* (*6G-FFT*) from *Aegilops tauschii*, and HORVU7Hr1G001070.17 to *vacuolar invertase1* (*VI-1*) from *Triticum monococcum.*

To further explore relationships between compounds we used a hypothesis-free analysis of metabolite ratios in a GWAS. This analysis generates a 'p-gain' statistic which is calculated from the significance of increases in -log10(p) values of the metabolite ratios compared to an estimated threshold derived from the p-values obtained in GWAS of the individual compounds (Petersen *et al.*, 2012). Using the Log transformed ratios between neoseries-DP7 and inulin-DP9 with all other compounds, 17 pairs of compounds correlated with a QTL that passed the FDR threshold of -log 10(p) 6.02 in the same region of chromosome 7H as neoseries-DP7 and inulin-DP9 alone. The ratios of neoseries-DP7:inulin-DP10, inulin-DP9:neoseries-DP8 and inulin-DP9:inulin-DP10 passed the p-gain threshold of 5.2×10^2 ($p < 0.05$) for markers with a MAF of > 10% with 1.97×10^5, 4.08×10^9 and 3.71×10^3, as well as 1.93×10^5 and 6.43×10^5, respectively (Table 3, Figure S4), indicating metabolic links between these compounds. The QTL on 7H overlapped for all ratios. Significant markers identified were SCRI_RS_8079, JHI-Hv50k-2016-435510, and JHI-Hv50k-2016-438638, the latter being the same as identified with the metabolite concentrations for neoseries-DP7 and inulin-DP9 alone (Table 2A). GWAS for the ratio neoseries-DP7:inulin-DP9 did not identify any significant associations (Table 3).

Evaluation of exome capture data revealed several non-synonymous SNPs

Mascher *et al.* (2017) presented exome capture data for 25 of the genotypes included in our study, which we evaluated to identify putative casual SNPs for our five regional candidates involved in fructan biosynthesis. Eight non-synonymous SNPs in *1-FFT*, three in *VI-1*, two in *6G-FFT*, and one in *1-SST* were identified (Table S6). All identified SNPs are located within functional protein coding regions of the genes (Figure 4). Changes in just one out of 25 genotypes were observed for three markers among the eight SNPs detected in *1-FFT*. The other five SNPs in *1-FFT* represent changes from methionine to leucine (position 7H_ 262685), alanine to threonine (7H_263547 and 7H_263700), isoleucine to threonine (7H_264127), and leucine to isoleucine (7H_264198). They showed significant effects ($p < 0.05$) on 1-kestose and several neoseries-type fructans (Figure S5). The two SNPs in *6G-FFT* were in LD and represent changes from glycine to glutamic acid (7H_321608), and alanine to threonine (7H_319284). Notably, they have a significant effect on 1-kestose and several inulin-type fructans (Figure S5). However, the SNP in *1-SST*, representing a change from

threonine to isoleucine (7H_279526), as well as the three SNPs in *VI-1*, representing changes from glutamic acid to aspartic acid (7H_2423349), tryptophan to arginine (7H_2425560), and arginine to cysteine (7H_2425578), did not show a significant effect on either trait. For *6-SFT* no SNP was identified.

Fructan biosynthesis genes show developmental stage and tissue specific expression patterns

We compared expression patterns of the five candidate genes (Table 2) and three known fructan hydrolyase encoding genes in various tissues across barley plant development (Figure 5). In the vegetative phase, highest expression for *1-SST* (facilitating the biosynthesis of 1-kestose, the precursor for production of inulin- and graminan-type fructans) was observed during early germination in embryo and all seedling tissues (Figure 5A). During the reproductive phase, *1-SST* is expressed in all vegetative tissues with peak expression in the leaf epidermis, as well as in all reproductive tissues and stages with a pronounced peak of expression in the ovary wall, the embryo sac (ES), the egg apparatus and central cell (EC+CC), and the antipodal cells (ANT) during late pistil development (stages W8 to W10, Figure 5B). In the grain development phase, *1-SST* expression is highest in the early stages (7 to 9 DAP) in maternal grain tissues (pericarp, aleurone, sub-aleurone/outer starchy endosperm (SA)) while decreasing during the storage stage (from 11 DAP onwards) in all tissues (Figure 5C). *1-FFT* (mediating the biosynthesis of inulin-type fructans) showed tight co-expression with *1-SST* during early germination in embryo and all seedling tissues (Figure 5A) as well as all vegetative tissues (Figure 5B), while in reproductive (Figure 5B) and grain tissues (Figure 5C) much lower expression levels were observed. In contrast, *6-SFT* (mediating the biosynthesis of graminans-type fructans) showed very tight co-expression with *1-SST* during meiosis and pistil development (Figure 5B) as well as at early grain development (Figure 5C). During germination, *6-SFT* expression was extremely low while it was observed to be moderate in seedling (Figure 5A) and all vegetative tissues (Figure 5B). Notably, expression of *6G-FFT* (mediating the biosynthesis of neoseries-type fructans) was restricted to the outer grain tissues (see aleurone tissues in Figure 5A and aleurone, pericarp and endosperm tissues in Figure 5C) during late grain development (from 11 DAP onwards). *VI-1*, with yet unknown function, showed low expression levels in germinated grain tissues, all vegetative tissues, in pericarp at late grain development and senescing leaf, while higher levels were notable during late pistil development (Figure 5). Among the fructan hydrolyases, *1-FEH* (HORVU6Hr1g011260) and *6-FEH* (HORVU2Hr1G109120) seem to be involved in

balancing fructan biosynthesis, with *1-FEH* tightly co-expressed with *1-SST* in all tissues and stages and pronounced *6-FEH* expression during the reproductive phase in all tissues and in the pericarp at late grain development. In contrast, only marginal expression levels were observed for *6-FEH/CWI2* (HORVU2Hr1G118820) (Figure 5).

Fructan biosynthesis genes show differential co-expression patterns in developing barley grain

Besides the five fructan biosynthesis genes, the association of differential oligosaccharide profiles with other candidates in the identified genomic region may be possible. We hypothesised similar expression patterns for fructan biosynthesis genes and other candidates influencing the fructan levels in barley. Therefore, transcript expression levels were evaluated for all gene models within the QTL interval and co-expression of genes was assessed individually within the developmental phases and tissues (Table S7). We have focused on developing barley grain and significant correlations for the expression of fructan metabolism genes with each other and with TFs (Table 4).

Highly positive correlations among the fructan metabolism genes were observed between *1-FFT*, *1-SST* and *6-SFT*; between *6G-FFT* and *6-FEH*; and for *1-FEH* with *1-FFT*, and *6-FEH*. Notable negative correlations were observed for *6G-FFT* with *1-FFT*, and *1-SST*. Co-expression patterns with TFs were highly similar for *1-FFT*, *1-SST* and *6-SFT* in the developing grain. Notable positive correlations for *1-FFT*, *1-SST* and *6-SFT* expression were identified with the *WD_REPEATS_REGION domain-containing protein* (HORVU7Hr1G000820.1), the *ALWAYS EARLY 3* (HORVU7Hr1G001120.1), and the two *scarecrow-like protein* genes (HORVU7Hr1G001300.3, HORVU7Hr1G001310.1). In contrast, *6G-FFT* showed significant negative correlations with the *WD_REPEATS_REGION domain-containing protein* (HORVU7Hr1G000820.1), *scarecrow-like protein 22* (HORVU7Hr1G001310.1), and *HTH myb-type domain-containing protein* (HORVU7Hr1G001830.3). Significant positive correlations for *VI-1* were observed in developing grain with the *protein ALWAYS EARLY 3* (HORVU7Hr1G001120.1) and *scarecrow-like protein 22* (HORVU7Hr1G001310.1), as observed for *1-FFT*, *1-SST*, and *6-SFT*. *1-FEH* showed co-expression patterns partly like those observed for *1-FFT*. Besides the positive correlation with a *myb-type transcription factor* (HORVU7Hr1G001830.3) a strong negative correlation with the *AP2/ERF domain-containing protein* (HORVU7Hr1G001050.1) and a positive interaction (not significant) with a *NAC domain-containing protein gene* (HORVU7Hr1G000910.1) were identified. Highest positive correlations were noted for *6-*

FEH with *6G-FFT* and the *HTH myb-type domain-containing protein* (HORVU7Hr1G001830.3), which in contrast was negatively associated with *6G-FFT* (Table 4).

Additional and partly different patterns for the co-expression of fructan metabolism genes with other genes were observed across a range of developmental phases and tissues (Table S7, Supplementary results).

Discussion

Neoseries-type fructans are abundant in mature barley grain

Profiling of DP3 to DP10 oligosaccharides revealed the abundance of 6G-kestose and higher DP neoseries-type fructans in mature barley grain (Table 1, Figure S1, Table S2). While the presence of 6G-kestose has been reported in wheat and barley grain (Nilsson *et al.*, 1986; Henry and Saini, 1989) higher DP variants of this fructan-type have not been previously identified. Recent studies revealed the presence of neoseries-type fructans in oat, rye, spelt and wheat flour (Verspreet *et al.*, 2015b; Verspreet *et al.*, 2017) but claimed its absence in barley (Verspreet *et al.*, 2017). These contrasting observations may be explained both by the plant materials used (flour vs. whole grain) and the technical constraints of fructan profiling. Noticeable accumulation of unidentified higher DP fructans was reported for the outer pericarp of developing wheat grain (Schnyder *et al.*, 1993). Thus, utilising whole grain here may have facilitated the detection of neoseries-type fructans in barley, likely accumulating in outer grain parts as suggested by expression analysis for *6G-FFT* (Figure 5). Additionally, electronic properties of PAD, typically used for fructan profiling, require higher concentrations for the detection of higher molecular weight fructans (Rocklin and Pohl, 1983), resulting in a pronounced log-logistic distribution for those compounds that is also observed in our study (Figure S2) and which may have led to the underrepresentation of fructans with >DP4 in other studies. In the future, comprehensive grain fructan profiling could be improved by employing recently established LC-MS methodologies, as reviewed in Matros *et al.* (2019).

Barley accessions group according to their oligosaccharide accumulation patterns

Reported genotypic variation in grain fructan content ranges from 0.9-4.2% of dry matter (DM) among 20 barley breeding lines (Nemeth et al., 2014) and from 1.1-1.6% of DM

among seven barley cultivars (De Arcangelis *et al.,* 2019). These results correspond well with the variability for total fructan values (0.02-1.94% of grain dry weight) determined here. When discriminating between different chain lengths Henry and Saini (1989) measured varying amounts of FOS with 0.26% (DP3), 0.2% (DP4), 0.03% (DP5) and 0.23% of DM (>DP5) in mature barley grains, which was confirmed by results from Jenkins *et al.*, (2011). In our study, the lowest abundance range showed FOS with DP5 (traces to 0.16% of DM) while FOS with DP3 and DP4 ranged from 0.02-0.53% and traces to 0.44% of DM, respectively. Nemeth *et al.*, (2014) observed a positive correlation between fructan values and the content of long chain fructans (> DP9, $r = 0.54$, $p = 0.021$). However, such an association could not be found in our dataset. Clustering of the oligosaccharide profiles from the 154 lines revealed two major profile groups, one each of higher and lower sugar values (Figure 1). We detected significant positive correlations between biosynthetically closely related metabolites (e.g. within and between the different fructan-types) with negative associations for antagonistic compounds (e.g. fructosylraffinose with all fructans, sugar monomers and dimers, Figure 2). Co-occurrence of fructans and RFO has been reported for many plant species including wheat (Haska *et al.*, 2008) and barley (Henry, 1988), with the proposal that strong RFO and fructan accumulation do not occur together in a single plant species (Van den Ende, 2013). Notably, fructosylraffinose was only speculated to occur in barley (Cerning and Guilbot, 1973), while its presence was described in wheat decades ago (White and Secor, 1953; Saunders, 1971).

Differences in oligosaccharide profiles are genetically controlled in barley

A significant QTL on chromosome 7H affecting barley grain fructan levels was identified (Figure 3A) and five genes involved in fructan metabolism were detected in this region (Table 2 and Table 3, Table S5). We increased the power of GWAS by analysing metabolite ratios using the p-gain approach. As the p-gain passed an appropriate threshold (defined by the data), using ratios provided more information about the traits, and the genomic locus underlying them, than looking at the traits individually. Using ratios reduces background 'noise' in datasets, increasing statistical power to detect significant associations between traits and genomic loci (Petersen *et al.*, 2012). Previous studies have demonstrated that including ratios between pairs of traits can strengthen associations identified and uncover novel information about biochemical pathways (Gieger *et al.*, 2008; Illig *et al.*, 2010; Suhre *et al.*, 2011). Thus, ratio-GWAS represents an innovative approach for the discovery of new biologically meaningful associations in plants, as shown for the linked oligosaccharide

pathways described here. However, when we used the ratio between neoseries-DP7:inulin-DP9 in the GWAS we did not identify an association on 7H, indicating less information provided by the ratios than the individual values. This may relate to the high positive correlation of these two compounds across the barley lines ($R^2 = 0.86$, Figure 2, Table S4) and the close genomic location of the related fructan biosynthesis genes (Table 2). In contrast, for the ratios between inulin-DP7:inulin-DP10 ($R^2 = 0.011$, $p = 0.80$), inulin-DP9:neoseries-DP8 ($R^2 = 0.56$, $p < 0.05$) and inulin-DP9:inulin-DP10 ($R^2 = 0.028$, $p = 0.54$) a significant QTL was identified. This points towards a stronger association between the identified genomic locus and the molecular weight of the fructans than with the fructan structure.

In wheat, two loci for differential total fructan contents in grain were identified on chromosomes 7A and 6D, which did not show significant interactions (Huynh *et al.*, 2008b). Subsequent physical mapping provided indications for clustering of fructan biosynthesis genes in the genomes of both dicots as well as monocots (Huynh *et al.*, 2012). For wheat and barley the formation of a functional cluster was shown containing *1-SST* (provided are the IDs of the most probable barley gene product; J7GM45_HORVV), *1-FFT* (J7GHS0_HORVV), and *6-SFT* (Q96466_HORVU) (Huynh *et al.*, 2012), which were also identified here. Additionally, the authors found two *vacuolar invertase* (*VI*) genes (J7GIU6_HORVV, J7GR98_HORVV) in this cluster, of which we identified one, which is similar to *6G-FFT* (J7GIU6_HORVV). The identification of *6G-FFT* matches the detection of neoseries-type fructans in our study. Among the five candidates we identified was also a gene coding for an uncharacterised gene product (M0X3V0_HORVV) which is similar to a *VI-1* from *T. monococcum* (Q6PVN1_TRIMO) that has not been described or annotated in barley before.

The evaluation of exome capture data (Mascher *et al.*, 2017) led to the identification of several significant SNPs in the five fructan biosynthesis genes. SNPs in *1-FFT* were associated with grain neoseries-type fructan content while SNPS in *6G-FFT* were associated with inulin-type fructan content (Figure 4, Table S6). Ideally, the influence of these SNPs would be validated in the complete set of germplasm used to quantify fructan content. This analysis would likely reveal additional SNPs that have not been identified in this subset of lines.

Developmental and tissue specific nature of barley fructan biosynthesis

Throughout plant development, *1-SST*, *1-FFT* and *1-FEH* showed strong co-expression, starting in embryo tissue during germination, accompanied later by *6-SFT* expression in root, leaf and stem, likely leading to the biosynthesis of inulin- and graminan-type fructans in those tissues until senescence (Figure 5). These observations matched the consensus of inulin- and graminan-type fructans being the predominant polymers in barley tissues (Pollock and Cairns, 1991; Bonnett *et al.*, 1997; Huynh *et al.*, 2008a). Accordingly, *1-SST*, *1-FFT*, *6-SFT* and *1-FEH* are the best studied fructan biosynthesis genes (Duchateau *et al.*, 1995; Henson, 2000; Lüscher *et al.*, 2000; Huynh *et al.*, 2012). A key role was assigned to *1-SST* (Wagner *et al.*, 1983) and correlated transcription and activity was reported for *1-SST* and *6-SFT* in barley leaves (Nagaraj *et al.*, 2004). A role for fructans as a temporal carbohydrate reserve has been widely accepted in vegetative tissues and roots (Pollock *et al.*, 1996; Vijn and Smeekens, 1999; Housley, 2000) and can be assumed for the inulin and graminan-type fructans in barley.

We observed co-expression of *1-SST*, *6-SFT*, *6-FEH*, and *1-FEH* in reproductive tissues with a pronounced peak during late pistil development in ovary tissues (Figure 5B) probably leading to specific accumulation of graminan-type fructans. In *Campanula rapunculoides*, the largest inulin-type fructan concentrations were found in petals and ovaries (Vergauwen *et al.*, 2000). Based on the observation that petals in daylily (*Hemerocallis*) (Bieleski, 1993) and *C. rapunculoides* (Vergauwen *et al.*, 2000) and leaves of *Phippsia algida* (Solhaug and Aares, 1994) rapidly degrade fructans upon flower opening, a role for them in flower expansion was suggested. However, the function of the different fructan-types accumulating in *C. rapunculoides* ovary (inulin-type) and barley ovary tissues (graminan-type) remains unresolved at present. Also, the newly identified *VI-1* showed peak expression specific to ovary tissues at late pistil development, while its function in fructan biosynthesis remains unclear.

In accordance with the detection of neoseries-type fructans and the identification of *6G-FFT* in the significant QTL region we observed *6G-FFT* expression in barley grain (Figure 5A and C). Notably, its expression was restricted to developing grain from 11 DAP onwards and confined to the outer endosperm and maternal tissues. Reports on this fructan type in developing grains of other cereals do not exist to our knowledge. Accumulation of neoseries-type fructans in the aleurone of mature grain may be related to favourable structural characteristics when compared to inulins and graminans and to the function of this tissue during germination. Some reports showed that fructan branching architecture is critical to

physicochemical properties, such as water solubility or formation of aggregates at high concentration (Eigner *et al.*, 1988; Wolff *et al.*, 2000; Ponce *et al.*, 2008). The more compact shape of neoseries-type fructans would allow higher concentrations to be stored in the desiccated aleurone. Better water solubility and pH-stability of neoseries-type fructans would be advantageous during germination, when the aleurone hydrates and enzymes must be quickly activated and reach their substrates. Proving these hypotheses will require the comparative evaluation of physicochemical properties of neoseries- inulin- and graminan-type fructans in the future. Indeed, several studies have provided strong evidence for a positive relationship between enhanced fructan concentrations with better malting characteristics in barley varieties (Smith *et al.*, 1980; Cozzolino *et al.*, 2016 and references therein)

Potential regulators of barley grain fructan biosynthesis

Despite the increasing evidence of tissue specificity, there is limited knowledge of how fructan metabolism is orchestrated to adjust the storage and use of photosynthates during grain development. Within the significant QTL interval, we found several genes differentially co-expressed with the various fructan biosynthesis genes in developing grain. Among them were several TFs (Table 4, Table S7, Figure 6).

It is generally agreed that initiation of fructan biosynthesis is triggered by an organ-specific sucrose threshold (Lu *et al.*, 2002, Jin *et al.*, 2017). Also, several molecular components in sucrose-mediated induction of plant fructan biosynthesis, such as protein phosphatases and kinases (Noël *et al.*, 2001), second messenger Ca^{2+} (Martinez-Noël *et al.*, 2006), small GTPases and phosphatidylinositol 3-kinase (Ritsema *et al.*, 2009), as well as the plant hormones abscisic acid and auxin (Valluru, 2015) were shown to be required for activation of fructosyltransferase genes.

An opposing sugar-sensing system was recently identified in barley, whereby a single gene on chromosome 2H encodes two functionally distinct TF variants [SUSIBA (sugar signaling in barley) 1 and 2], which respond differently to sucrose concentrations (Jin *et al.*, 2017). However, no distinction was made between different tissues and fructan types and it remains unclear if this system coordinates fructan and starch biosynthesis in general.

In wheat, TaMYB13, a R2R3-MYB TF, was described as a transcriptional activator of fructan biosynthesis (Xue *et al.*, 2011; Kooiker *et al.*, 2013). In the promoter of the barley genes *1-FFT*, *1-SST*, *6-SFT* and *VI,* binding motifs for TaMYB13 were identified, suggesting that the co-expression of these genes may be driven by a TaMYB13 homolog (Huynh et al., 2012).

However, the *HTH myb-type domain-containing protein* (HORVU7Hr1G001830.3) identified here did not show similarity to TaMYB13 and we could not observe a clear homolog in barley. While three myb-type TFs were also described to activate promoters of genes involved in fructan biosynthesis and degradation in chicory (Wei *et al.*, 2017a and b), involvement of other TF family genes has not yet been reported.

Conclusions

A new genomic region and several causal SNPs involved in the regulation of barley grain fructan content were identified. The genomic region includes a physical cluster of functionally related fructan biosynthetic genes and several potential regulatory genes. While the clustering of fructan biosynthetic genes may hint at the co-evolution of these gene families, a conserved gene co-expression suggesting an equal contribution to grain fructan biosynthesis was not observed. Instead the spatiotemporal dynamics for fructan biosynthetic genes point towards versatile roles of the different fructan types. Phylogenetic relationships between fructosyltransferases and invertases within *Poaceae* suggest that *6-SFT* may have evolved from a *Poaceae* ancestor genome after the major clade of vacuolar invertases diverged, followed then by *1-FFT* and *1-SST* (Huynh *et al.*, 2012). The analysis also showed the presence of a unique barley clade of four vacuolar invertase genes, among them the newly annotated *6G-FFT*, between the *6-SFT* and the *1-FFT* and *1-SST* clades, suggesting that extra duplication might have occurred in barley. Accordingly, in developing grain we observed similar co-expression with a set of TFs for *1-FFT*, *1-SST*, and *6-SFT*, which was different from the associations found for *6G-FFT*. The proposed dynamics of fructan biosynthesis in barley grain and potential regulators are presented in Figure 6. Assuming a specific spatiotemporal control of grain fructan biosynthesis, breeding or genetic engineering for high fructan content related to grain specific traits (e.g. nutritional quality or germination) will require careful approaches targeting certain tissues and developmental stages as recently suggested for engineering mixed linkage (1,3;1,4)-β-glucan biosynthesis in the endosperm (Lim *et al.*, 2019).

689 **Supplementary data**

690 **Supplementary Table S1:** List of germplasms with their corresponding oligosaccharide
691 profile group.

692 **Supplementary Table S2:** Oligosaccharide annotation information.

693 **Supplementary Table S3:** Oligosaccharide peak area entry means values.

694 **Supplementary Figure S1:** Representative chromatogram of mature barley grain soluble
695 carbohydrates

696 **Supplementary Figure S2:** Metabolite distribution among the lines.

697 **Supplementary Figure S3:** Oligosaccharide profile prototypes as obtained by Neural Gas
698 clustering.

699 **Supplementary Table S4:** Correlation and significance values for associations among the 27
700 metabolites.

701 **Supplementary Figure S4:** Manhattan plots and box plots for all significant results from the
702 ratio GWAS.

703 **Supplementary Table S5:** List of all detected gene models in the significant QTL interval.

704 **Supplementary Table S6:** Results from exome capture data evaluation.

705 **Supplementary Figure S5:** Box plots for significant effects of SNPs identified from exome
706 capture data.

707 **Supplementary Table S7:** Heatmap of correlations for the expression of fructan genes with
708 other gene models from the QTL interval for all developmental phases and tissues.

709 **Supplementary Table S8:** Expression data and Pearson correlation values for associations
710 among the gene models from the GWA scan for all developmental phases and tissues.

711

712

713 **Conflict of Interest Statement**

714 All authors state no conflict of interest concerning this manuscript.

715

716

717

Acknowledgements

This work was supported by grants from the German Research Foundation (DFG, MA 4814/3-1 & 3-2) to Andrea Matros and the Australian Research Council (ARC) Centre of Excellence in Plant Energy Biology (http://www.plantenergy.uwa.edu.au/) to Rachel Burton. Kelly Houston and Robbie Waugh acknowledge the support of the Rural & Environment Science & Analytical Services Division of the Scottish Government We greatly acknowledge the excellent technical assistance of Shi F. Khor, Jelle Lahnstein. We wish to thank Natalie S. Betts and Helen M. Collins for their help with the geminated grain dataset. The authors acknowledge the use of the facilities, and scientific and technical assistance of the Australian Plant Phenomics Facility, which is supported by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS).

Author Contributions

AM and RAB designed and developed the concept of the study. RW provided the material and genomic data of the barley panel and was involved in the design of the GWAS. AM and BB conducted the growth experiments and harvested the mature grain material at TPA. AM performed the oligosaccharide profiling analysis and evaluated the HPAEC-PAD data. AM and KW conducted the experiments related to the identification of fructan structures (isolation and MS identification). US performed the clustering of the data and the distribution analysis. KH performed the GWAS and ratio-GWAS as well as the evaluation of the exome capture data. MRT, MKA and LGW conducted the transcriptomic analyses of pistil tissues and developing grain tissues. MS and RW conducted the transcriptomic analysis of developing anther tissues. AM evaluated the transcript expression datasets provided and conducted the co-expression analyses. AM, KH and RAB wrote and provided a draft manuscript which has been revised and accepted by all authors.

References

Andrade AC, Bautista CR, Cabrera MR, Guerra RS, Chávez EG, Ahumada CF, Lagunes AG. 2019. *Agave salmiana* fructans as gut health promoters: Prebiotic activity and inflammatory response in Wistar healthy rats. International Journal of Biological Macromolecules 136, 785-795.

Aubert MK, Coventry S, Shirley NJ, Betts NS, Würschum T, Burton RA, Tucker MR. 2018. Differences in hydrolytic enzyme activity accompany natural variation in mature aleurone morphology in barley (*Hordeum vulgare* L.). Scientific Reports 8, 1-14, doi: 10.1038/s41598-018-29068.

Bayer MM, Rapazote-Flores P, Ganal M, Hedley PE, Macaulay M, Plieske J, Ramsay L, Russell J, Shaw PD, Thomas W. 2017. Development and evaluation of a barley 50k iSelect SNP array. Frontiers in Plant Science 8, 1792, doi: 10.3389/fpls.2017.01792.

Barakate A, Orr J, Schreiber M, Colas I, Lewandowska D, McCallum N, Macaulay M, Morris J, Arrieta M, Hedley PE. 2020. Time-resolved transcriptome of barley anthers and meiocytes reveals robust and largely stable gene expression changes at meiosis entry. bioRxiv, doi: 10.1101/2020.04.20.051425.

Betts NS, Dockter C, Berkowitz O, Collins HM, Hooi M, Lu Q, Burton RA, Bulone V, Skadhauge B, Whelan J. 2019. Transcriptional and biochemical analyses of gibberellin expression and content in germinated barley grain. Journal of Experimental Botany, doi: 10.1093/jxb/erz546.

Bieleski RL. 1993. Fructan hydrolysis drives petal expansion in the ephemeral daylily flower. Plant Physiology 103 (1), 213-219.

Bonnett GD, Sims IM, Simpson RJ, Cairns AJ. 1997. Structural diversity of fructan in relation to the taxonomy of the Poaceae. New Phytologist 136 (1), 11-17.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23 (19), 2633-2635.

Burton RA, Fincher GB. 2009. (1, 3; 1, 4)-β-d-Glucans in cell walls of the Poaceae, lower plants, and fungi: a tale of two linkages. Molecular Plant 2 (5), 873-882.

Cairns AJ, Pollock CJ, Gallagher JA, Harrison J. 2000. Fructans: synthesis and regulation. In: Leegood, RC, Sharkey, TD, Von Caemmerer S, eds, Photosynthesis: physiology and metabolism. Kluwer Academic Publishers, pp 301-320.

Carpita NC, Housley TL, Hendrix JE. 1991. New features of plant-fructan structure revealed by methylation analysis and C-13 NMR-spectroscopy. Carbohydrate Research 217, 127-136.

Cerning J, Guilbot A. 1973. Changes in the carbohydrate composition during development and maturation of the wheat and barley kernel. Cereal Chem 220-232.

Cochrane MP. 2000. Seed carbohydrates. In: Black M, Brewley JD, eds. Seed technology and its biological basis. USA, Boca Rathon: CRC Press, pp 85-120.

Colmsee C, Beier S, Himmelbach A, Schmutzer T, Stein N, Scholz U, Mascher M. 2015. BARLEX–the barley draft genome explorer. Molecular Plant 8 (6), 964-966.

Comadran J, Kilian B, Russell J, Ramsay L, Stein N, Ganal M, Shaw P, Bayer M, Thomas W, Marshall D. 2012. Natural variation in a homolog of Antirrhinum centroradialis contributed to spring growth habit and environmental adaptation in cultivated barley. Nature Genetics 44 (12), 1388.

Cozzolino D, Degner S, Eglinton J. 2016. Relationships between fructans content and barley malt quality. Food Analytical Methods 9 (7), 2010-2015.

Darrier B, Russell J, Milner SG, Hedley PE, Shaw PD, Macaulay M, Ramsay LD, Halpin C, Mascher M, Fleury DL. 2019. A comparison of mainstream genotyping

platforms for the evaluation and use of barley genetic resources. Frontiers in Plant Science 10, 544, doi: 10.3389/fpls.2019.00544.

De Arcangelis E, Messia MC, Marconi E. 2019. Variation of polysaccharides profiles in developing kernels of different barley cultivars. Journal of Cereal Science 85, 273-278.

Duchateau N, Bortlik K, Simmen U, Wiemken A, Bancal P. 1995. Sucrose: fructan 6-fructosyltransferase, a key enzyme for diverting carbon from sucrose to fructan in barley leaves. Plant Physiology 107 (4), 1249-1255.

Dwivedi S, Sahrawat K, Puppala N, Ortiz R. 2014. Plant prebiotics and human health: biotechnology to breed prebiotic-rich nutritious food crops. Electronic Journal of Biotechnology 17 (5), 238-245.

Eigner W-D, Abuja P, Beck RH, Praznik W. 1988. Physicochemical characterization of inulin and sinistrin. Carbohydate Research 180 (1), 87-95.

Gieger C, Geistlinger L, Altmaier E, De Angelis MH, Kronenberg F, Meitinger T, Mewes H-W, Wichmann H-E, Weinberger KM, Adamski J. 2008. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS Genetics 4 (11), doi: 10.1371/journal.pgen.1000282.

Haskå L, Nyman M, Andersson R. 2008. Distribution and characterisation of fructan in wheat milling fractions. Journal of Cereal Science 48 (3), 768-774.

Hendry GAF, Wallace RK. 1993. The origin, distribution, and evolutionary significance of fructans. In: Suzuki M, Chatterton NJ, eds. Science and technology of fructans. USA, Boca Rathon: CRC Press, pp 119-139.

Henry RJ. 1988. The carbohydrates of barley grains—A review. Journal of the Institute of Brewing 94 (2), 71-78.

Henry R, Saini H. 1989. Characterization of cereal sugars and oligosaccharides. Cereal Chemistry 66 (5), 362-365.

Henson CA. 2000. Enzymology of fructan polymerization and depolymerization in grasses. In: Gupta AK, Kaur N, eds. Carbohydrate Reserves in Plants - Synthesis and Regulation. The Netherlands, Amsterdam: Elsevier Science, pp 269-281.

Housley T. 2000. Role of fructans redistributed from vegetative tissues in grain filling of wheat and barley. In: Gupta AK, Kaur N, eds. Developments in Crop Science, Elsevier, pp 207-221.

Huynh B-L, Palmer L, Mather DE, Wallwork H, Graham RD, Welch RM, Stangoulis JC. 2008a. Genotypic variation in wheat grain fructan content revealed by a simplified HPLC method. Journal of Cereal Science 48 (2), 369-378.

Huynh B-L, Wallwork H, Stangoulis JC, Graham RD, Willsmore KL, Olson S, Mather DE. 2008b. Quantitative trait loci for grain fructan concentration in wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 117 (5), 701-709.

Huynh B-L, Mather D, Schreiber A, Toubia J, Baumann U, Shoaei Z, Stein N, Ariyadasa R, Stangoulis JR, Edwards J, et al. 2012. Clusters of genes encoding fructan biosynthesizing enzymes in wheat and barley. Plant Molecular Biology 80 (3), 299-314.

Illig T, Gieger C, Zhai G, Römisch-Margl W, Wang-Sattler R, Prehn C, Altmaier E, Kastenmüller G, Kato BS, Mewes H-W. 2010. A genome-wide perspective of genetic variation in human metabolism. Nature genetics 42 (2), 137, doi: 10.1038/ng.507.

Jenkins CL, Lewis D, Bushell R, Belobrajdic DP, Bird AR. 2011. Chain length of cereal fructans isolated from wheat stem and barley grain modulates in vitro fermentation. Journal of Cereal Science 53 (2), 188-191.

Jin Y, Fei M, Rosenquist S, Jin L, Gohil S, Sandström C, Olsson H, Persson C, Höglund A-S, Fransson G. 2017. A dual-promoter gene orchestrates the sucrose-coordinated synthesis of starch and fructan in barley. Molecular plant 10 (12), 1556-1570.

Julkowska MM, Saade S, Agarwal G, Gao G, Pailles Y, Morton MJ, Awlia M, Tester M. 2019. MVApp-Multivariate analysis application for streamlined data analysis and curation. Plant Physiology 180 (3), 1261-1276.

Kaspar-Schoenefeld S, Merx K, Jozefowicz AM, Hartmann A, Seiffert U, Weschke W, Matros A, Mock HP. 2016. Label-free proteome profiling reveals developmental-dependent patterns in young barley grains. Journal of Proteomics 143, 106-121.

Kooiker M, Drenth J, Glassop D, McIntyre CL, Xue G-P. 2013. TaMYB13-1, a R2R3 MYB transcription factor, regulates the fructan synthetic pathway and contributes to enhanced fructan accumulation in bread wheat. Journal of Experimental Botany 64 (12), 3681-3696.

Langridge P, Waugh R. 2019. Harnessing the potential of germplasm collections. Nature Genetics 51, 200.

Lim WL, Collins HM, Byrt CS, Lahnstein J, Shirley NJ, Aubert MK, Tucker MR, Peukert M, Matros A, Burton RA. 2019. Overexpression of HvCslF6 in barley grain alters carbohydrate partitioning plus transfer tissue and endosperm development. Journal of Experimental Botany 71 (1), 138-153.

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z. 2012. GAPIT: genome association and prediction integrated tool. Bioinformatics 28 (18), 2397-2399.

Liu FT, Li P, Chen MX, Luo YM, Prabhakar M, Zheng HM, He Y, Qi Q, Long HY, Zhang Y, et al. 2017. Fructooligosaccharide (FOS) and galactooligosaccharide (GOS) increase bifidobacterium but reduce butyrate producing bacteria with adverse glycemic metabolism in healthy young population. Scientific Reports 7 (1), 11789.

Lu C, Koroleva OA, Farrar JF, Gallagher J, Pollock CJ, Tomos AD. 2002. Rubisco small subunit, chlorophylla/b-binding protein and sucrose: fructan-6-fructosyl transferase gene expression and sugar status in single barley leaf cells *in situ*. Cell type specificity and induction by light. Plant Physiology 130 (3), 1335-1348.

Lüscher M, Hochstrasser U, Boller T, Wiemken A. 2000. Isolation of sucrose: sucrose 1-fructosyltransferase (1-SST) from barley (*Hordeum vulgare*). The New Phytologist 145 (2), 225-232.

Martínez-Noël G, Tognetti J, Nagaraj V, Wiemken A, Pontis H. 2006. Calcium is essential for fructan synthesis induction mediated by sucrose in wheat. Planta 225 (1), 183-191.

Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V, Dockter C, Hedley PE, Russell J. 2017. A chromosome conformation capture ordered sequence of the barley genome. Nature 544, 427-433.

Matros A, Peukert M, Lahnstein J, Seiffert U, Burton R. 2019. Determination of fructans in plants: Current analytical means for extraction, detection, and quantification. Annual Plant Reviews Online, 1-39.

Meints B, Cuesta-Marcos A, Fisk S, Ross A, Hayes P. 2016. Food barley quality improvement and germplasm utilization. In Exploration, identification and utilization of barley germplasm. Elsevier, pp 41-73.

Nagaraj VJ, Altenbach D, Galati V, Lüscher M, Meyer AD, Boller T, Wiemken A. 2004. Distinct regulation of sucrose: sucrose-1-fructosyltransferase (*1-SST*) and sucrose: fructan-6-fructosyltransferase (*6-SFT*), the key enzymes of fructan synthesis in barley leaves: *1-SST* as the pacemaker. New Phytologist 161 (3), 735-748.

Nemeth C, Andersson AA, Andersson R, Mangelsen E, Sun C, Åman P. 2014. Relationship of grain fructan content to degree of polymerisation in different barleys. Food and Nutrition Sciences 5 (6), 581-589.

Noël G, Tognetti JA, Pontis HG. 2001. Protein kinase and phosphatase activities are involved in fructan synthesis initiation mediated by sugars. Planta 213 (4), 640-646.

Nilsson U, Dahlqvist A, Nilsson B. 1986. Cereal fructosans: part 2—Characterization and structure of wheat fructosans. Food chemistry 22 (2), 95-106.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nature methods 14 (4), 417.

Petersen A-K, Krumsiek J, Wägele B, Theis FJ, Wichmann H-E, Gieger C, Suhre K. 2012. On the hypothesis-free testing of metabolite ratios in genome-wide and metabolome-wide association studies. BMC bioinformatics 13 (1), 120, doi: 10.1186/1471-2105-13-120.

Peukert M, Lim WL, Seiffert U, Matros A. 2016. Mass spectrometry imaging of metabolites in barley grain tissues. Current Protocols in Plant Biology 1 (4), 574-591.

Peukert M, Thiel J, Peshev D, Weschke W, Van den Ende W, Mock HP, Matros A. 2014. Spatio-temporal dynamics of fructan metabolism in developing barley grains. The Plant Cell 26 (9), 3728-3744.

Pollock CJ, Cairns AJ. 1991. Fructan metabolism in grasses and cereals. Annual Review of Plant Physiology and Plant Molecular Biology 42 (1), 77-101.

Pollock CJ, Cairns AJ, Sims IM, Housley TL. 1996. Fructans as reserve carbohydrates in crop plants In: Zamski E, Shaffer AA, eds. Photoassimilate distribution in plants and crops: Source-sink relationships. USA, New York: Marcel Dekker. pp 97-114.

Ponce JA, Macías ER, Soltero JA, Fernández VV, Zúñiga V, Escalona HB. 2008. Physical-Chemical and non-linear rheological properties of aqueous solutions of agave fructans. e-Gnosis 6, 1-23.

Ritsema T, Brodmann D, Diks SH, Bos CL, Nagaraj V, Pieterse CM, Boller T, Wiemken A, Peppelenbosch MP. 2009. Are small GTPases signal hubs in sugar-mediated induction of fructan biosynthesis? PLoS One 4. e6605

Rocklin RD, Pohl CA. 1983. Determination of carbohydrates by anion exchange chromatography with pulsed amperometric detection. Journal of Liquid Chromatography 6 (9), 1577-1590.

Saunders R. 1971. Fructosylraffinose, a tetrasaccharide in wheat bran. Phytochemistry 10 (2), 491-493.

Schnyder H, Gillenberg C, Hinz J. 1993. Fructan contents and dry matter deposition in different tissues of the wheat grain during development. Plant, Cell & Environment 16 (2), 179-187.

Smith D, Morgan A, Aastrup S. 1980. Variation in the biochemical composition of acid extracts from barleys of contrasting malting quality. Journal of the Institute of Brewing 86 (6), 277-283.

Solhaug KA, Aares E. 1994. Remobilization of fructans in *Phippsia algida* during rapid inflorescence development. Physiologia Plantarum 91 (2), 219-225.

Suhre K, Shin S-Y, Petersen A-K, Mohney RP, Meredith D, Wägele B, Altmaier E, Deloukas P, Erdmann J, Grundberg E. 2011. Human metabolic individuality in biomedical and pharmaceutical research. Nature 477, 54-60, doi: 10.1038/nature10354.

Valluru R. 2015. Fructan and hormone connections. Frontiers in Plant Science 6, 180, doi: 10.3389/fpls.2015.00180.

Van den Ende W. 2013. Multifunctional fructans and raffinose family oligosaccharides. Frontiers in Plant Science 4, 247, doi: 10.3389/fpls.2013.00247.

Veenstra LD, Jannink J-L, Sorrells ME. 2017. Wheat fructans: A potential breeding target for nutritionally improved, climate-resilient varieties. Crop Science 57 (3), 1624-1640.

Veenstra LD, Santantonio N, Jannink J-L, Sorrells ME. 2019. Influence of genotype and environment on wheat grain fructan content. Crop Science 59 (1), 190-198.

Vergauwen R, Van den Ende W, Van Laere A. 2000. The role of fructan in flowering of *Campanula rapunculoides*. Journal of Experimental Botany 51 (348), 1261-1266.

Verspreet J, Dornez E, Delcour JA, Harrison SJ, Courtin CM. 2015a. Purification of wheat grain fructans from wheat bran. Journal of Cereal Science 65, 57-59.

Verspreet J, Dornez E, Van den Ende W, Delcour JA, Courtin CM. 2015b. Cereal grain fructans: structure, variability and potential health effects. Trends Food Science & Technology 43 (1), 32-42.

Verspreet J, Hansen AH, Dornez E, Delcour JA, Van den Ende W, Harrison SJ, Courtin CM. 2015c. LC-MS analysis reveals the presence of graminan- and neo-type fructans in wheat grains. Journal of Cereal Science 61, 133-138.

Verspreet J, Hansen AH, Harrison SJ, Vergauwen R, Van den Ende W, Courtin CM. 2017. Building a fructan LC–MS2 library and its application to reveal the fine structure of cereal grain fructans. Carbohydrate Polymers 174, 343-351.

Verspreet J, Pollet A, Cuyvers S, Vergauwen R, Van den Ende W, Delcour JA, Courtin CM. 2012. A simple and accurate method for determining wheat grain fructan content and average degree of polymerization. Journal of Agricultural and Food Chemistry 60 (9), 2102-2107.

Vijn I, Smeekens S. 1999. Fructan: more than a reserve carbohydrate?. Plant Physiology 120 (2), 351-360.

Wagner W, Keller F, Wiemken A. 1983. Fructan metabolism in cereals: induction in leaves and compartmentation in protoplasts and vacuoles. Zeitschrift für Pflanzenphysiologie 112 (4), 359-372.

Wei H, Bausewein A, Greiner S, Dauchot N, Harms K, Rausch T. 2017a. *Ci MYB 17*, a stress-induced chicory R2R3-MYB transcription factor, activates promoters of genes involved in fructan synthesis and degradation. New Phytologist 215 (1), 281-298.

Wei H, Zhao H, Su T, Bausewein A, Greiner S, Harms K, Rausch T. 2017b. Chicory R2R3-MYB transcription factors *CiMYB5* and *CiMYB3* regulate fructan 1-exohydrolase expression in response to abiotic stress and hormonal cues. Journal of Experimental Botany 68 (15), 4323-4338.

White LM, Secor GE. 1953. Chromatographic evidence for the occurrence of a fructosyl raffinose in wheat flour and wheat. Archives of Biochemistry and Biophysics 44 (1), 244-245.

Wilkinson LG, Yang X, Burton R, Würschum T, Tucker MR. 2019. Natural variation in ovule morphology is influenced by multiple tissues and impacts downstream grain development in barley (*Hordeum vulgare* L). Frontiers in Plant Science 10, 1374, doi: 10.3389/fpls.2019.01374.

Wolff D, Czapla S, Heyer A, Radosta S, Mischnick P, Springer J. 2000. Globular shape of high molar mass inulin revealed by static light scattering and viscometry. Polymer 41 (22), 8009-8016.

Xue GP, Kooiker M, Drenth J, McIntyre CL. 2011. TaMYB13 is a transcriptional activator of fructosyltransferase genes involved in β-2, 6-linked fructan synthesis in wheat. The Plant Journal 68 (5), 857-870.

Yamamori A, Okada H, Kawazoe N, Ueno K, Onodera S, Shiomi N. 2015. Structure of fructan prepared from onion bulbs (*Allium cepa* L.). Journal of Applied Glycoscience 62 (3), 95-99.

Zhang Z, Ersoz E, Lai C-Q, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM. 2010. Mixed linear model approach adapted for genome-wide association studies. Nature Genetics 42 (4), 355-360.

Tables

Table 1: List of the 27 annotated metabolites. Peaks were annotated by comparison with analytical standards and isolated fractions from barley grain and onion bulb samples as well as based on fructanase digestion and mild acid hydrolysis. Further details of compound annotation are provided in Table S2. Abbreviations: DP, degree of polymerisation; KP, kestopentaose; KT, kestotetraose; NS, neoseries-type fructan; RT, retention time

Peak #	Compound	RT [min]	DP	Molecular family	Fructan-type
1	Glucose	*3.43*	1	Monosaccharide	
2	Fructose	*3.80*	1	Monosaccharide	
3	Melibiose	*4.36*	2	Disaccharide	
4	Sucrose	*5.95*	2	Disaccharide	
5	Raffinose	*9.20*	3	Raffinose family oligosaccharides	
6	1-Kestose / 6-Kestose	*9.74*	3	Fructan	Inulin/Levan
7	Maltose	*10.47*	2	Maltose-type oligosaccharides	
8	Fructosylraffinose	*10.73*	4	Raffinose family oligosaccharides	
9	B-Type Procyanidin 1	*11.53*	2	Flavonoids	
10	6G-Kestose	*12.13*	3	Fructan	NS-Inulin
11	Nystose	*12.76*	4	Fructan	Inulin
12	B-Type Procyanidin 2	*13.08*	2	Flavonoids	
13	1&6-KT (Bifurcose) / 6G&1-KT (NS-DP4)	*13.52*	4	Fructan	Graminan, NS-Inulin
14	Maltotriose	*14.11*	3	Maltose-type oligosaccharides	
15	UK-Fructan1	*14.68*	-	Fructan	unknown
16	1,1,1-Kestopentaose	*15.68*	5	Fructan	Inulin
17	6G,1-KP (NS-DP5)	*16.38*	5	Fructan	NS-Inulin
18	6G&1-KP (NS-DP5)	*16.61*	5	Fructan	NS-Inulin
19	Maltotetraose	*17.51*	4	Maltose-type oligosaccharides	
20	Inulin-DP6	*18.47*	6	Fructan	Inulin
21	Neoseries-DP6	*19.28*	6	Fructan	NS-Inulin
22	Inulin-DP7	*21.10*	7	Fructan	Inulin
23	Neoseries-DP7	*22.73*	7	Fructan	NS-Inulin
24	Inulin-DP8	*23.48*	8	Fructan	Inulin
25	Neoseries-DP8	*25.15*	8	Fructan	NS-Inulin
26	Inulin-DP9	*25.94*	9	Fructan	Inulin
27	Inulin-DP10	*28.10*	10	Fructan	Inulin

Table 2: Significant GWA results for the two metabolites neoseries-DP7 and inulin-DP9. Abbreviations: bp, base pair; DP, degree of polymerisation; LOD, logarithm of odds; MAF, minimum allele frequency; QTL, quantitative trait loci

A: Information on the detected significant QTL

Trait	Chromosome	Peak marker	Peak marker bp	MAF	LOD	QTL start and end bp
Neoseries-DP7	7H	JHI-Hv50k-2016-438638	3407292	0.14	8.65	174327..4056691
Inulin-DP9					6.74	

B: Candidate gene models related to fructan biosynthesis

Gene model	HORVU7Hr1G000250.3	HORVU7Hr1G000260.2	HORVU7Hr1G000270.1	HORVU7Hr1G001040.6	HORVU7Hr1G001070.17
Gene model location	261065..262018	276441..280334	318543..322514	2257431..2260153	2423328..2427280
Description	Acid beta-fructofuranosidase, GH family 32 protein	Acid beta-fructofuranosidase, GH family 32 protein	Acid beta-fructofuranosidase, GH family 32 protein	Acid beta-fructofuranosidase, GH family 32 protein	Acid beta-fructofuranosidase, GH family 32 protein
PFAM	PF00251, PF08244, PF11837	PF00251, PF08244, PF11837	PF00251, PF08244	PF00251, PF08244	PF00251, PF08244
UniprotKB	A0A287VBC4	M0XA31	J7GIU6	M0YBF9	A0A287VC78
Similar proteins	J7GHS0	B6ECP1	M8C9V9	Q9AUH3	Q6PVN1
Similarity	0.9	1	0.9	1	0.9
Description	Fructan:fructan 1-fructosyltransferase	Sucrose:sucrose 1-fructosyltransferase	6(G)-fructosyltransferase	Sucrose:fructan 6-fructosyltransferase	Vacuolar invertase1
Species	*Hordeum vulgare*	*Hordeum vulgare*	*Aegilops tauschii*	*Hordeum vulgare*	*Triticum monococcum*
Abbreviation	1-FFT	1-SST	6G-FFT	6-SFT	VI-1

Table 3: Significant GWA results for the ratios of neoseries-DP7 and inulin-DP9 with all other metabolites. Abbreviations: bp, base pair; DP, degree of polymerisation; FDR, false discovery rate; LOD, logarithm of odds; MAF, minimum allele frequency; *P*, probability value; *P*-gain, ratio of the lowest p-value of the two individual metabolites and the p-value of the metabolite ratio; P-values and FDR adjusted p-values from the initial GWAS for traits that did not identify significant associations but when included as a ratio do identify significant associations are included in columns '*P* for non sig trait' and 'FDR adjusted p-value for non sig trait'; * indicates significant results passing the p-gain threshold of 5.2×10^2, which are also highlighted in light grey. Corresponding Manhattan and box plots are shown in Supplementary Figure S4.

Ratio combination	Peak marker	Peak marker bp	MAF	LOD	*P*	*P* for non sig trait	FDR adjusted p-value	FDR adjusted p-value for non sig trait	*P*-gain	FDR adjusted p-gain
Neoseries-DP7:Fructosylraffinose	JHI-Hv50k-2016-438638 JHI-Hv50k-2016-438980	3407292 3689408	0.14	7.0	1.08E-07	6.26E-04	2.24E-03	6.76E-01	2.08E-02	2.08E-02
Neoseries-DP7:B-Type Procyanidin1				6.9	1.39E-07	4.73E-05	2.87E-03	1.00E+00	1.62E-02	1.62E-02
Neoseries-DP7:1,1,1-Kestopentaose				9.1	8.62E-10	8.22E-05	1.79E-05	2.84E-01	2.61E+00	2.61E+00
Neoseries-DP7:6G,1-KP				9.0	8.96E-10	9.06E-01	1.86E-05	1.00E+00	2.51E+00	2.51E+00
Neoseries-DP7:Inulin-DP6				8.4	3.88E-09	9.78E-04	8.04E-05	1.00E+00	5.80E-01	5.80E-01
Neoseries-DP7:Neoseries-DP6				7.9	1.28E-08	1.51E-01	2.66E-04	1.00E+00	1.75E-01	1.75E-01
Neoseries-DP7:Inulin-DP7				9.1	8.45E-10	3.17E-03	1.75E-05	1.00E+00	2.66E+00	2.66E+00
Neoseries-DP7:Neoseries-DP8				10.8	1.64E-11	7.89E-02	2.14E-07	1.00E+00	1.37E+02	2.17E+02
Neoseries-DP7:Inulin-DP10	SCRI_RS_8079	2325562	0.11	7.4	4.24E-08	3.46E-01	1.72E-03	1.00E+00	1.97E+05*	5.80E+02
Inulin-DP9:1-Kestose	JHI-Hv50k-2016-438638 JHI-Hv50k-2016-438980	3407292 3689408	0.14	8.0	1.03E-08	1.72E-02	2.14E-04	1.00E+00	1.75E+01	1.75E+01
Inulin-DP9:1,1,1-Kestopentaose				9.4	4.38E-10	8.22E-05	9.08E-06	2.84E-01	4.14E+02	4.14E+02
Inulin-DP9:6G,1-KP				7.3	6.29E-06	9.06E-01	9.40E-04	1.00E+00	2.88E-02	4.00E+00
Inulin-DP9:Inulin-DP6				8.4	4.03E-09	9.78E-04	8.35E-05	1.00E+00	4.50E+01	4.50E+01
Inulin-DP9:Neoseries-DP6				8.0	1.04E-08	1.51E-01	2.15E-04	1.00E+00	1.75E+01	1.75E+01
Inulin-DP9:Inulin-DP7				9.0	9.04E-10	3.17E-03	1.87E-05	1.00E+00	2.01E+02	2.01E+02
Inulin-DP9:Neoseries-DP8	JHI-Hv50k-2016-435510 JHI-Hv50k-2016-437376	280043 2350955	0.07	10.7	1.93E-11	3.71E-04	4.01E-07	1.00E+00	4.08E+09*	2.02E+06
	JHI-Hv50k-2016-438638 JHI-Hv50k-2016-438980	3407292 3689408	0.14	10.3	4.89E-11	7.89E-02	5.06E-07	1.00E+00	3.71E+03*	7.42E+03
Inulin-DP9:Inulin-DP10	JHI-Hv50k-2016-435510 JHI-Hv50k-2016-437376	280043 2350955	0.07	6.9	1.36E-07	3.21E-01	1.93E-03	1.00E+00	1.93E+04*	4.18E+02
	SCRI_RS_8079	2325562	0.11	6.9	1.40E-07	3.46E-01	1.93E-03	1.00E+00	6.43E+05*	5.17E+02
Neoseries-DP7:Inulin-DP9	JHI-Hv50k-2016-438638 JHI-Hv50k-2016-438980	3407292, 3689408	0.14	1.6	2.31E-02	1.80E-07	1.00E+00	3.74E-03	9.75E-08	4.66E-05

Table 4: Significant correlations for the expression of fructan metabolism genes with each other and with potential regulatory gene models from the detected QTL interval for developing barley grain. Positive correlations are shown in blue and negative ones in orange, whereas the color code is indicative for the strength of the correlation (the darker, the stronger). Significance threshold was $p > 0.05$. Numbered boxes without formatting indicate values just above significance ($p < 0.055$). Bold framed boxes indicate correlations which were detected in various developing grain datasets. Datasets correspond to the ones presented in Figure 5 and methods are detailed in section 'material and methods'. The raw data are presented in Table S8.

Gene Name	Annotation	Molecular Function	1-FFT	1-SST	6G-FFT	6-SFT	VI1	1-FEH	6-FEH
HORVU2Hr1G109120.2	**Fructan 6-exohydrolase (6-FEH)**	**Carbohydrate metabolism**			0.81			0.62	
HORVU6Hr1g011260.19	**Fructan 1-exohydrolase (1-FEH)**	**Carbohydrate metabolism**							0.62
HORVU7Hr1G000250.3	**Fructan-fructan 1-fructosyltransferase (1-FFT)**	**Carbohydrate metabolism**		0.92	-0.97	0.88			
HORVU7Hr1G000260.2	**Sucrose:sucrose 1-fructosyltransferase (1-SST)**	**Carbohydrate metabolism**	0.92		-0.84	0.99			
HORVU7Hr1G000270.1	**6(G)-fructosyltransferase (6G-FFT)**	**Carbohydrate metabolism**	-0.97	-0.84		-0.80			0.81
HORVU7Hr1G000820.1	WD_REPEATS_REGION domain-containing protein	Transcription factor	0.95	0.99	-0.87	0.98			
HORVU7Hr1G000910.1	NAC domain-containing protein	Transcription factor						0.81	
HORVU7Hr1G001040.6	**Sucrose:fructan 6-fructosyltransferase (6-SFT)**	**Carbohydrate metabolism**	0.73	0.98					
HORVU7Hr1G001050.1	AP2/ERF domain-containing protein	Transcription factor						-0.83	
HORVU7Hr1G001070.17	**Vacuolar invertase1 (VI-1)**	**Carbohydrate metabolism**							
HORVU7Hr1G001120.1	Protein ALWAYS EARLY 3	Transcription factor	0.80	0.92		0.95	0.78		
HORVU7Hr1G001300.3	GRAS domain-containing protein (Scarecrow-like protein 6)	Transcription factor	0.68	0.84		0.88			
HORVU7Hr1G001310.1	Scarecrow-like protein 22	Transcription factor	0.99	0.96	-0.95	0.94	0.80		
HORVU7Hr1G001320.1	GRAS domain-containing protein	Transcription factor				0.80			
HORVU7Hr1G001830.3	HTH myb-type domain-containing protein	Transcription factor	0.88		-0.91			0.68	0.88

Figure legends

Figure 1: Oligosaccharide profile groups as obtained by Neural Gas clustering. Shown are the mean profiles of the two major profile groups. Value (y-axis) represents the peak area [nC*min] of the individual compounds (x-axis) listed in Table 1. Profile group 2 was characterised by higher levels of sugar monomers and sucrose, lower fructosylraffinose and higher overall fructan values. Abbreviations: G-glucose, F-fructose, Me-melibiose, S-sucrose, R-raffinose, 1-K/6-K-1-kestose/6-kestose, M-maltose, FR-fructosylraffinose, P1-procyanidin B1, 6G-K-6G-kestose, N-Nystose, P2-procyanidin B2, B-1&6-KT (Bifurcose)/6G&1-KT (NS-DP4), Mtr-maltotriose, UK-unknown fructan, I-DP5-inulin DP5, NS-DP5-neoseries-DP5 (probably 6G,1-KP and 6G&1-KP), Mte-maltotetraose, I-DP6-inulin DP6, NS-DP6-neoseries-DP6, I-DP7-inulin DP7, NS-DP7-neoseries-DP7, I-DP8-inulin DP8, NS-DP8-neoseries-DP8, I-DP9-inulin DP9, I-DP10-inulin DP10.

Figure 2: Correlation pattern among metabolites. Pair-wise Pearson correlations are shown in a heat map representation, whereas metabolites are sorted according to correlation-based hierarchical cluster analysis. High positive correlations are represented by dark blue and negative ones by red circles, whereas the circle diameter is indicative for the strength of the correlation. X-not significantly correlated ($p>0.05$). The raw data are presented in Table S4.

Figure 3: Results of the GWA scan. (A) Manhattan plots are shown for the two fructans neoseries-DP7 and inulin-DP9. The $-\log_{10}$ (p-value) is shown on the y-axis and the 7 barley chromosomes are shown on the x-axis. Marker-trait association analysis was based on mean integrated peak areas. Integrated box plots show information for the top SNP (JHI-Hv50k-2016-438638). The variable alleles found (A and G) are shown on the x-axis. The y-axis shows the median of the metabolite amount for all lines with the respective allele variant; the width of the box is indicative for the number of lines with this particular allele. The false discovery rate significance threshold = -log 10(p) 6.02. (B) Assembling of the annotated gene models from the significant QTL region according to functional categories as obtained from the UniProt database (https://www.uniprot.org/uniprot/, June 2019). Numbers represent the count of gene models with the respective functional annotation.

Figure 4: Location of non-synonymous SNPs in fructan biosynthesis candidate genes. All significant causal SNPs (position assigned in bold red) are located within functional protein coding regions of the genes (black regions of the transcripts). All identified SNPs are

presented in Table S6 and box plots for significant effects are shown in Figure S5. For *6-SFT* no SNP was identified.

Figure 5: Transcript expression of the fructan biosynthesis genes in barley across plant development and in various tissues. Datasets included in the analysis are detailed in the materials and methods section. Genes included comprise the five fructan biosynthesis genes from the detected significant QTL region (*1-FFT*, HORVU7Hr1G000250; *1-SST*, HORVU7Hr1G000260; *6G-FFT*, HORVU7Hr1G000270; *6-SFT*, HORVU7Hr1G001040; and *VI-1*, HORVU7Hr1G001070) as well as three known fructan hydrolyase encoding genes (*6-FEH*, HORVU2Hr1G109120; *6-FEH/CWI2*, HORVU2Hr1G118820; and *1-FEH*, HORVU6Hr1g011260). Expression levels are colour coded, whereas different scales were used for TPM and FPKM values as indicated in the legend. (A) shows expression levels in the early vegetative phase for whole germinated grain tissues and for isolated germinated grain tissues from 0 to 96 hours after imbibition (HAI). Also, expression levels in seedling are shown for germinated embryo (GE), root and shoot. (B) shows data from the reproductive phase. Vegetative tissues included are: EPI, epidermal strips (4 weeks after planting, W4); ROO2, roots (W4); RAC, inflorescences, rachis (W5); LEM, inflorescences, lemma (W6); LOD, inflorescences, lodicule (W6); PAL, dissected inflorescences, palea (W6); INF2, inflorescence (10 mm); and NOD, internode. Meiosis stages included are: A.Pre, premeiosis anthers; A.LepZyg, leptotene/zygotene anthers; M.Lep/Zyg leptotene/zygotene meiocytes; A.MetTet, metaphaseI-tetrad anthers; A.PacDip, pachytene/diplotene anthers; M.PacDip, pachytene/diplotene meiocytes. Waddington (W) stages for pistil development are: W8; W8.5; W9; W9.5; and W10. Isolated pistil tissues are: nucellus (including nucellus and embryo sac for W8.5); integument; ovary wall; ES, embryo sac; EC+CC, egg apparatus and central cell; ANT, antipodal cells; and chalaza. (C) shows data from the grain development phase for whole developing grain (from 7 to 20 days after pollination, DAP) and for isolated developing grain tissues from 7 to 25 DAP. Abbreviations are SA, sub-aleurone/outer starchy endosperm; and SE, starchy endosperm/inner starchy endosperm. Also, data from senescing leaf (SN) are presented. Other abbreviations are: At, anthesis; CI, collar initiation; Em, emergence; Hv, harvest; and Sw, sowing.

Figure 6: Fructan-types, suggested biosynthesis routes and potential regulators in developing barley grain. Specific spatiotemporal biosynthesis of oligofructans was observed for barley

grains. Continuous bold arrows illustrate the major route of biosynthesis during the pre-storage phase (until 14 DAP) and the dashed bold arrows indicate the major route during the storage phase (until 20 DAP). During the pre-storage phase high transcript levels for *1-SST* and *6-SFT* were observed for the endosperm leading to an accumulation of 6-kestose and bifurcose. With transition to the storage phase a transcriptional switch was observed resulting in high transcript levels of *1-SST* in the nucellar projection (NP). *1-FFT* was found to be exclusively expressed in the NP during the storage phase. Induction of the inulin-type fructan biosynthesis pathway led to high amounts of 1-kestose and nystose accumulating in the endosperm cavity (Peukert et al., 2014). The dotted bold arrow illustrates the major route of biosynthesis during the late storage phase with *6G-FFT* transcripts detected in the outer endosperm (from 30 DAP onwards, Figure 5), which matched the detection of neoseries-type oligofructans in mature barley grains (Figure 1). Transcription factors (TF) showing significant correlation of transcript expression pattern in developing grain with *1-FFT, 1-SST* and *6-SFT* (positive), *6G-FFT* (negative), and *VI-1* (positive) are listed in the inserted text box. Inulin-neoseries represents linear fructans with β(2,1) & β(2,6) linked fructosyl units at the glucose (1F, 6G-di-β-D-fructofuranosylsucrose is shown; m=1, n= 1), graminan-type represents branched fructans with β(2,1) & β(2,6) linked fructosyl units (bifurcose is shown; m= 1, n= 1), inulin-type illustrates linear fructans with β(2,1) linked fructosyl units (1-kestose is shown; n= 1); , and levan-type shows linear fructans with β(2,6) linked fructosyl units (6-kestose is shown; n= 1). The arrows indicate direction of further polymerisation. Abbreviations are: *1-FFT*, fructan:fructan 1-fructosyltransferase; *1-SST*, sucrose:sucrose 1-fructosyl-transferase; *6-SFT*, sucrose:fructan 6-fructosyltransferase; *6G-FFT*, fructan:fructan 6G-fructosyltransferase; *VI-1*, vacuolar invertase 1 (unknown role in developing grain).

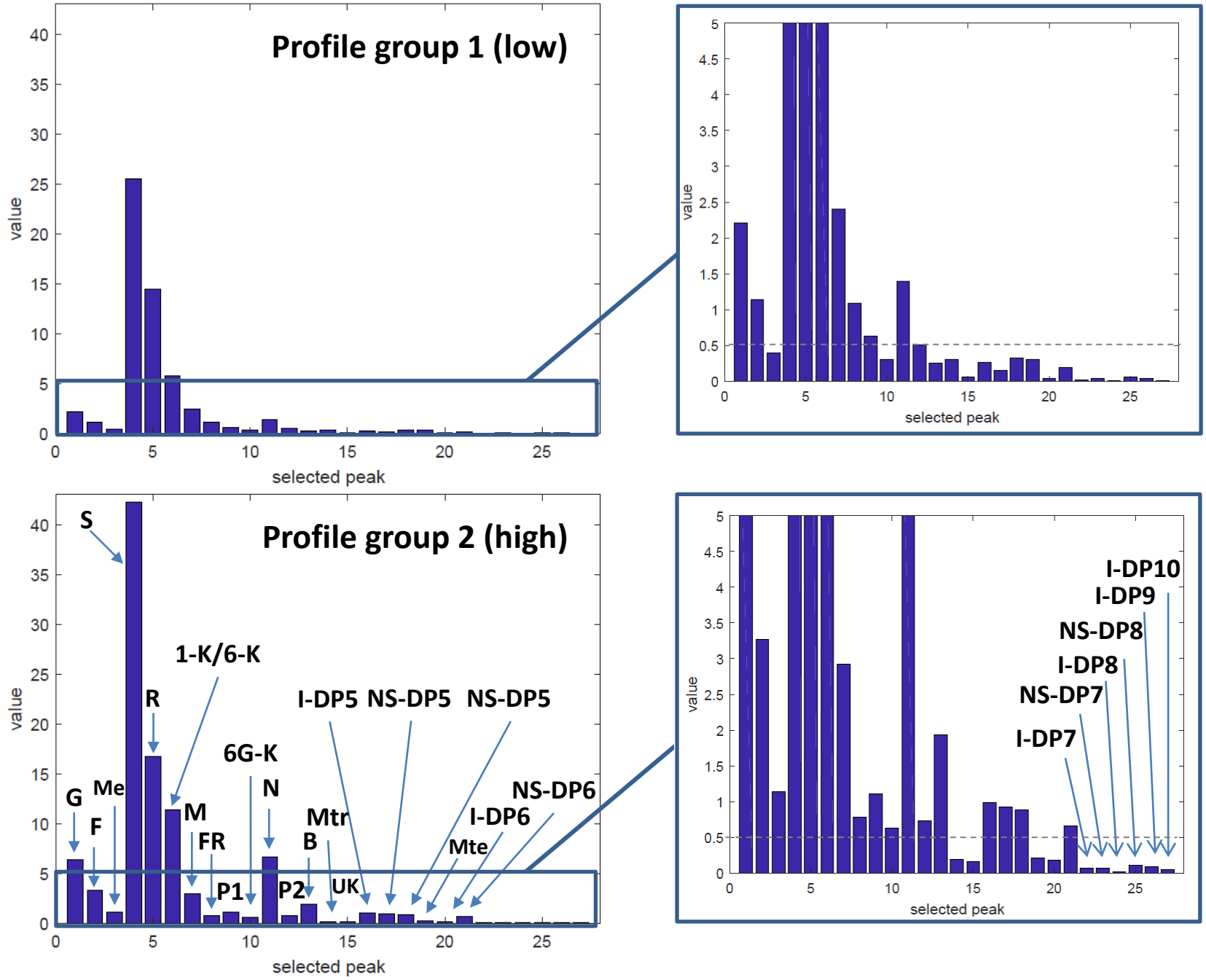

Profile group 1 (low)
value
selected peak
Profile group 2 (high)
S
G
Me
F
R
1-K/6-K
M
FR
P1
6G-K
N
P2
B
Mtr
UK
I-DP5
NS-DP5
NS-DP5
I-DP6
Mte
NS-DP6
I-DP7
NS-DP7
I-DP8
NS-DP8
I-DP9
I-DP10

Figure 1: Oligosaccharide profile groups as obtained by Neural Gas clustering. Shown are the mean profiles of the two major profile groups. Value (y-axis) represents the peak area [nC*min] of the individual compounds (x-axis) listed in Table 1. Profile group 2 was characterised by higher levels of sugar monomers and sucrose, lower fructosylraffinose and higher overall fructan values. Abbreviations: G-glucose, F-fructose, Me-melibiose, S-sucrose, R-raffinose, 1-K/6-K-1-kestose/6-kestose, M-maltose, FR-fructosylraffinose, P1-procyanidin B1, 6G-K-6G-kestose, N-Nystose, P2-procyanidin B2, B-1&6-KT (Bifurcose)/6G&1-KT (NS-DP4), Mtr-maltotriose, UK-unknown fructan, I-DP5-inulin DP5, NS-DP5-neoseries-DP5 (probably 6G,1-KP and 6G&1-KP), Mte-maltotetraose, I-DP6-inulin DP6, NS-DP6-neoseries-DP6, I-DP7-inulin DP7, NS-DP7-neoseries-DP7, I-DP8-inulin DP8, NS-DP8-neoseries-DP8, I-DP9-inulin DP9, I-DP10-inulin DP10.

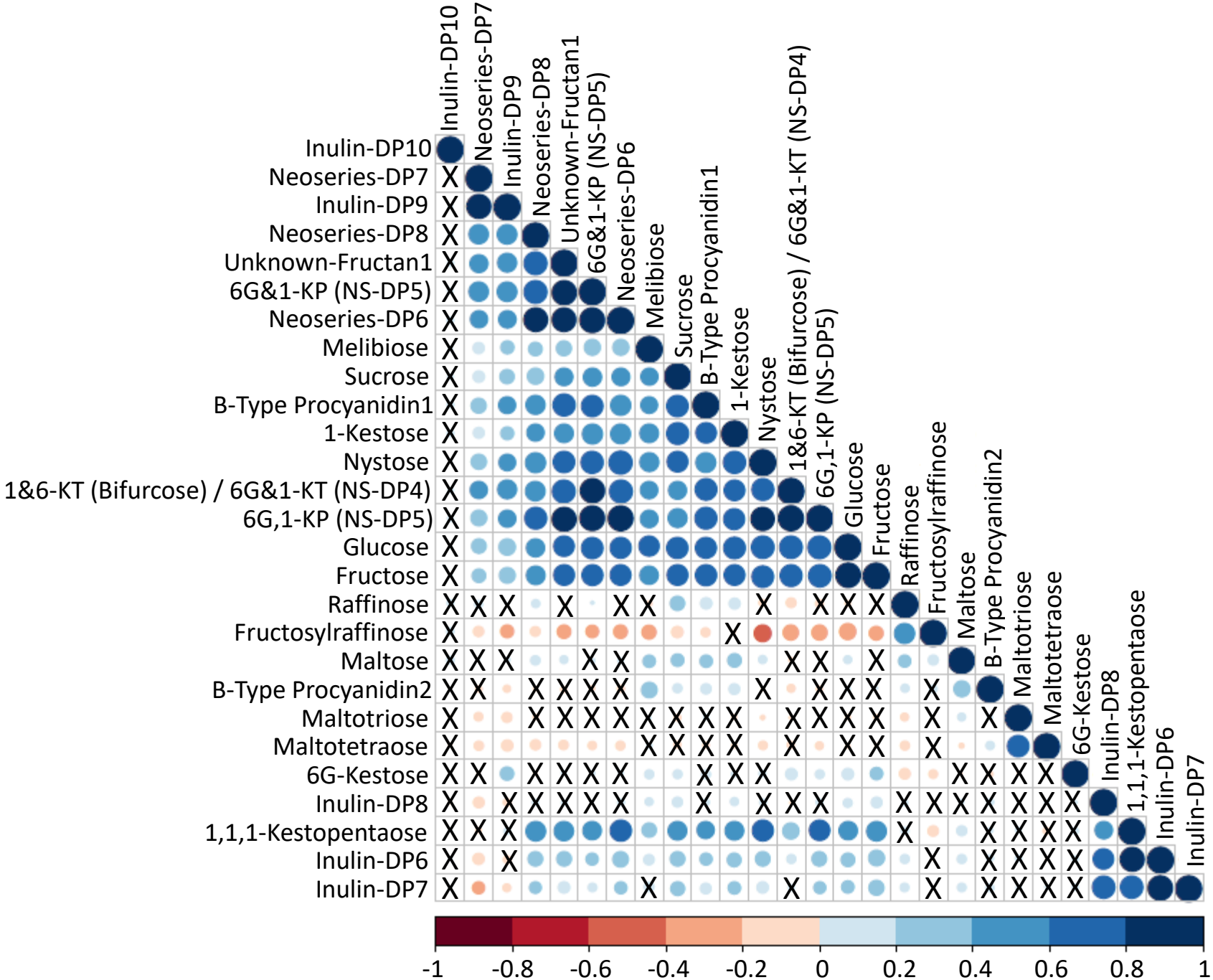
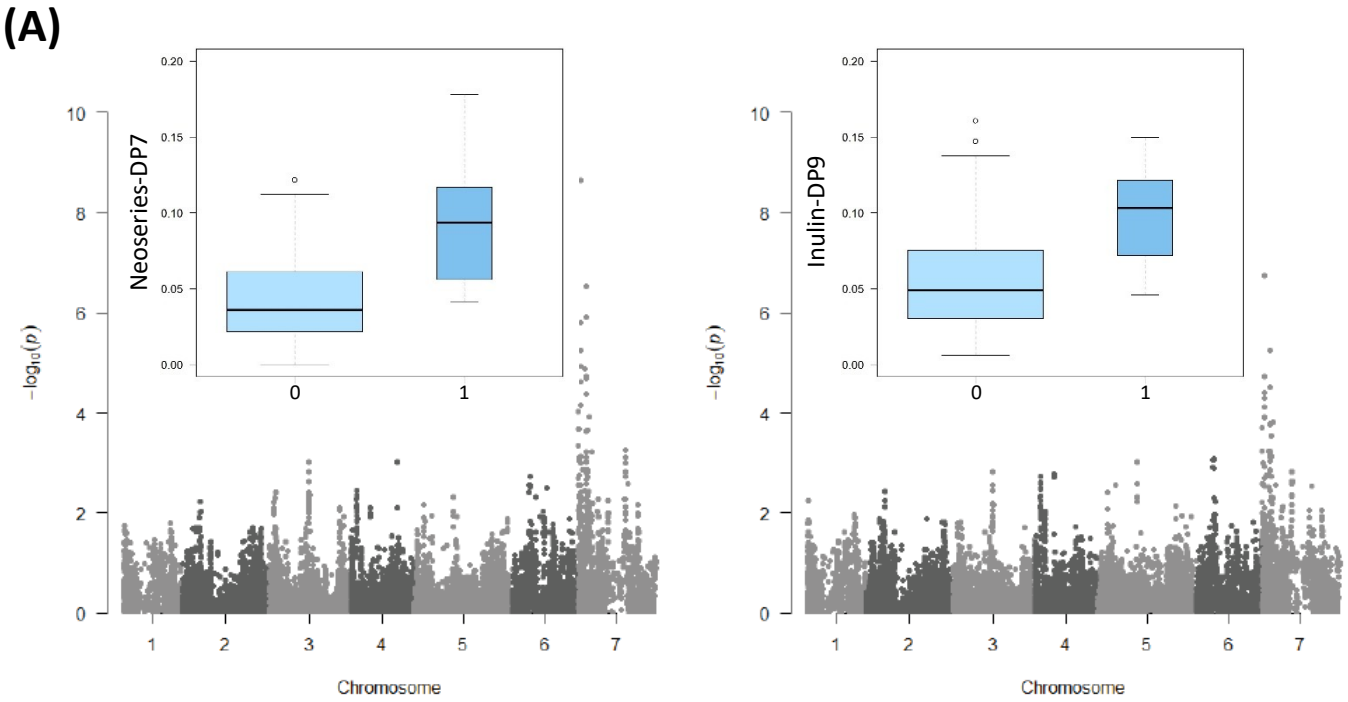

Inulin-DP10
Neoseries-DP7
Inulin-DP9
Neoseries-DP8
Unknown-Fructan1
6G&1-KP (NS-DP5)
Neoseries-DP6
Melibiose
Sucrose
B-Type Procyanidin1
1-Kestose
Nystose
1&6-KT (Bifurcose) / 6G&1-KT (NS-DP4)
6G,1-KP (NS-DP5)
Glucose
Fructose
Raffinose
Fructosylraffinose
Maltose
B-Type Procyanidin2
Maltotriose
Maltotetraose
6G-Kestose
Inulin-DP8
1,1,1-Kestopentaose
Inulin-DP6
Inulin-DP7
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

Figure 2: Correlation pattern among metabolites. Pair-wise Pearson correlations are shown in a heat map representation, whereas metabolites are sorted according to correlation-based hierarchical cluster analysis. High positive correlations are represented by dark blue and negative ones by red circles, whereas the circle diameter is indicative for the strength of the correlation. X-not significantly correlated ($p>0.05$). The raw data are presented in Table S4.



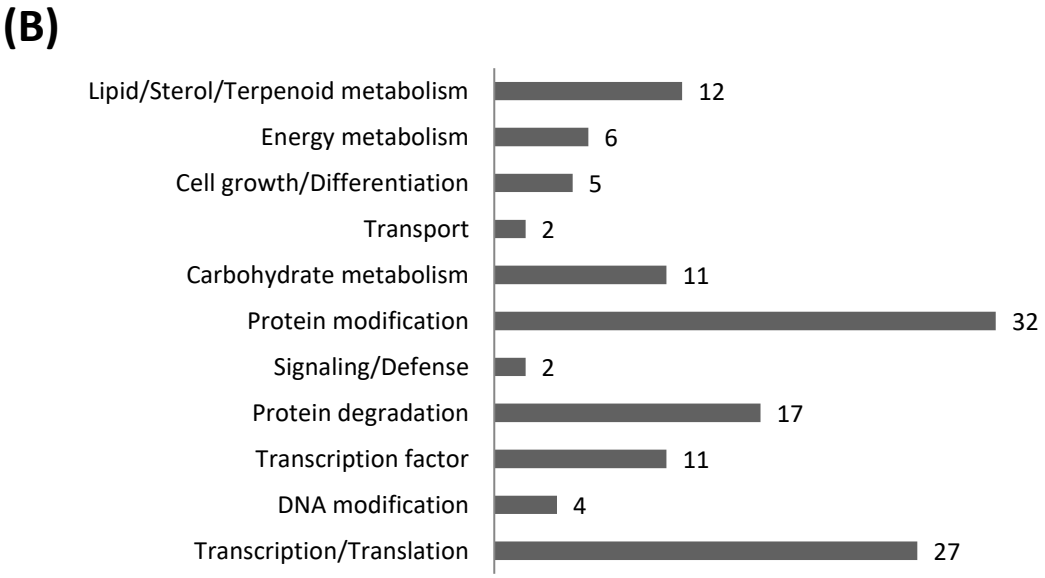


Figure 3: Results of the GWA scan. (A) Manhattan plots are shown for the two fructans neoseries-DP7 and inulin-DP9. The –log10 (p-value) is shown on the y-axis and the 7 barley chromosomes are shown on the x-axis. Marker-trait association analysis was based on mean integrated peak areas. Integrated box plots show information for the top SNP (JHI-Hv50k-2016-438638). The variable alleles found (A and G) are shown on the x-axis. The y-axis shows the median of the metabolite amount for all lines with the respective allele variant; the width of the box is indicative for the number of lines with this particular allele. The false discovery rate significance threshold = -log 10(p) 6.02. (B) Assembling of the annotated gene models from the significant QTL region according to functional categories as obtained from the UniProt database (https://www.uniprot.org/uniprot/, June 2019). Numbers represent the count of gene models with the respective functional annotation.

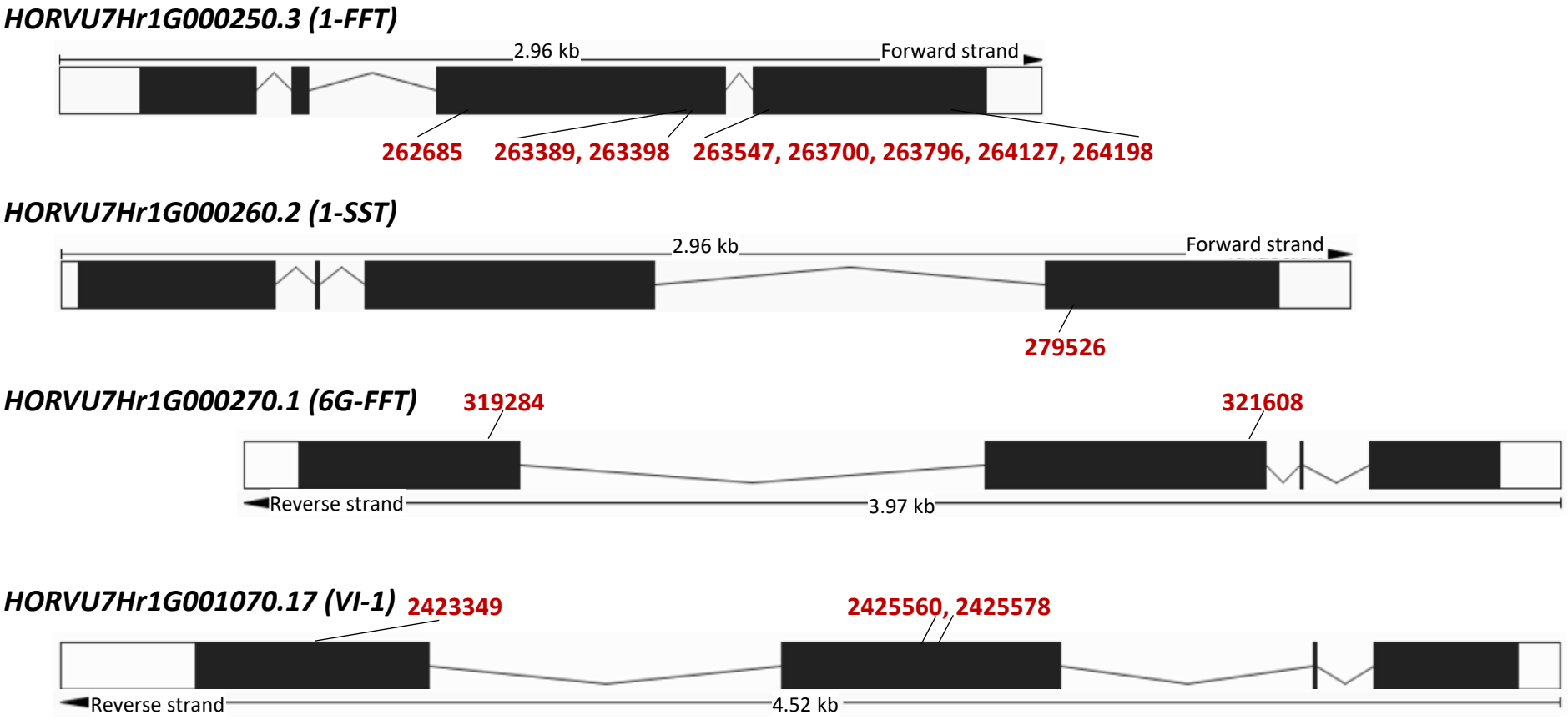
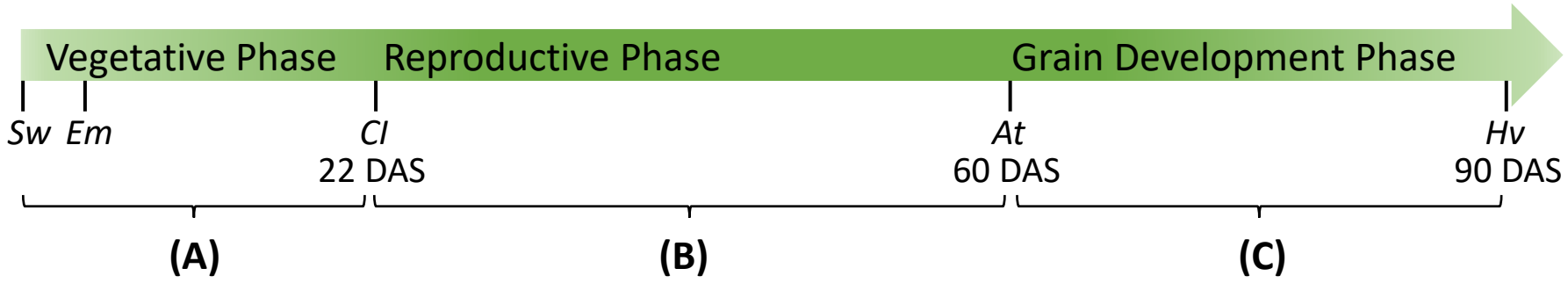


Figure 4: Location of non-synonymous SNPs in fructan biosynthesis candidate genes. All significant causal SNPs (position assigned in bold red) are located within functional protein coding regions of the genes (black regions of the transcripts). All identified SNPs are presented in Table S6 and box plots for significant effects are shown in Figure S5. For *6-SFT* no SNP was identified.



(A)

Germinating Grain

	0	24	48	72	96 HAI
1-FFT	2	67	26	20	32
1-SST	18	232	67	68	113
6G-FFT	4	2	0	0	0
6-SFT	0	0	0	0	0
VI-1	0	0	2	1	0
6-FEH	0	1	2	2	1
6-FEH/CWI2	0	0	0	0	0
1-FEH	6	25	41	25	16

Seedling

	GE	Root	Shoot
1-FFT	49	34	59
1-SST	239	212	268
6G-FFT	0	0	0
6-SFT	38	19	80
VI-1	134	79	42
6-FEH	20	67	18
6-FEH/CWI2	0	0	0
1-FEH	21	24	23

Germinating Grain Tissues

	Proximal Aleurone					Middle Aleurone					Distal Aleurone					Embryo					Scutellum				
	0	24	48	72	96	0	24	48	72	96	0	24	48	72	96	0	24	48	72	96	0	24	48	72	96 HAI
1-FFT	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	5	138	35	35	80	0	5	1	1	21
1-SST	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	38	475	104	142	243	5	52	3	3	10
6G-FFT	19	3	1	1	1	18	12	3	1	0	18	14	4	2	1	0	0	0	0	0	0	0	0	0	0
6-SFT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
VI-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	2	1	0
6-FEH	0	1	2	3	0	0	0	0	0	0	0	0	0	0	0	1	1	2	2	0	0	1	1	1	0
6-FEH/CWI2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0
1-FEH	7	163	180	107	18	4	8	122	140	65	4	3	53	75	26	7	11	5	3	3	12	33	21	17	8

(B)

Vegetative Tissues

	EPI	ROO2	RAC	LEM	LOD	PAL	INF2	NOD
1-FFT	503	253	341	169	328	236	446	373
1-SST	566	127	108	337	356	172	312	422
6G-FFT	0	0	0	0	0	0	76	0
6-SFT	234	436	449	153	292	374	416	262
VI-1	551	154	631	735	223	387	252	385
6-FEH	509	134	389	675	234	852	351	90
6-FEH/CWI2	346	0	204	349	161	709	0	63
1-FEH	219	169	89	215	565	301	318	86

Meiosis Stages

	A.Pre	A.LepZyg	M.LepZyg	A.MetTet	A.PacDip	M.PacDip
1-FFT	3	1	1	0	2	0
1-SST	13	33	34	32	15	15
6G-FFT	0	0	0	0	0	0
6-SFT	5	61	41	62	15	9
VI-1	0	0	0	0	0	0
6-FEH	40	19	66	69	41	54
6-FEH/CWI2	0	0	0	0	0	0
1-FEH	17	16	27	25	20	27

Pistil Development

	W8	W8.5	W9	W9.5	W10
1-FFT	3	4	4	6	10
1-SST	94	153	172	228	238
6G-FFT	0	0	0	0	0
6-SFT	15	39	93	379	460
VI-1	0	0	1	4	16
6-FEH	10	11	11	15	8
6-FEH/CWI2	0	0	0	0	0
1-FEH	10	11	12	12	14

Pistil Tissues

	Nucellus			Integument			Ovary wall			ES	EC+CC		ANT		Chalaza
	W8.5*	W9.5	W10	W8.5	W9.5	W10	W8.5	W9.5	W10	W9.5	W9.5	W10	W9.5	W10	W10
1-FFT	0	0	0	0	0	0	0	1	2	0	0	0	0	0	2
1-SST	3	5	1	3	4	4	11	109	101	78	190	199	92	1	4
6G-FFT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6-SFT	0	0	0	0	0	0	0	7	99	0	44	44	0	0	0
VI-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6-FEH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6-FEH/CWI2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1-FEH	1	2	0	1	2	1	1	2	3	0	5	1	2	1	2

* Nucellus and embryo sac

(C) Grain Development Stages

	7	9	11	13	15	20 DAP
1-FFT	9	8	6	5	2	1
1-SST	156	118	44	21	18	4
6G-FFT	0	2	9	14	32	32
6-SFT	791	525	126	35	11	7
VI-1	0	0	0	0	1	0
6-FEH	1	0	0	0	2	1
6-FEH/CWI2	1	0	0	0	1	0
1-FEH	14	15	19	16	15	7

Grain Development Tissues

	Pericarp				Aleurone				SA				SE			
	7	9	13	25	7*	9	13	25	7	9	13	25	7	9	13	25 DAP
1-FFT	10	6	8	2	0	1	1	0	1	2	1	0	0	0	0	0
1-SST	68	61	4	2	47	53	3	0	21	25	0	0	7	6	0	0
6G-FFT	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0
6-SFT	212	267	0	1	203	223	0	0	473	456	0	5	320	85	1	3
VI-1	0	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0
6-FEH	0	0	8	533	0	0	5	7	0	0	0	0	0	0	1	0
6-FEH/CWI2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
1-FEH	15	14	5	51	40	41	9	1	1	1	0	0	1	4	4	0

* 7 DAP sample not aleurone yet

Senescing Leaf

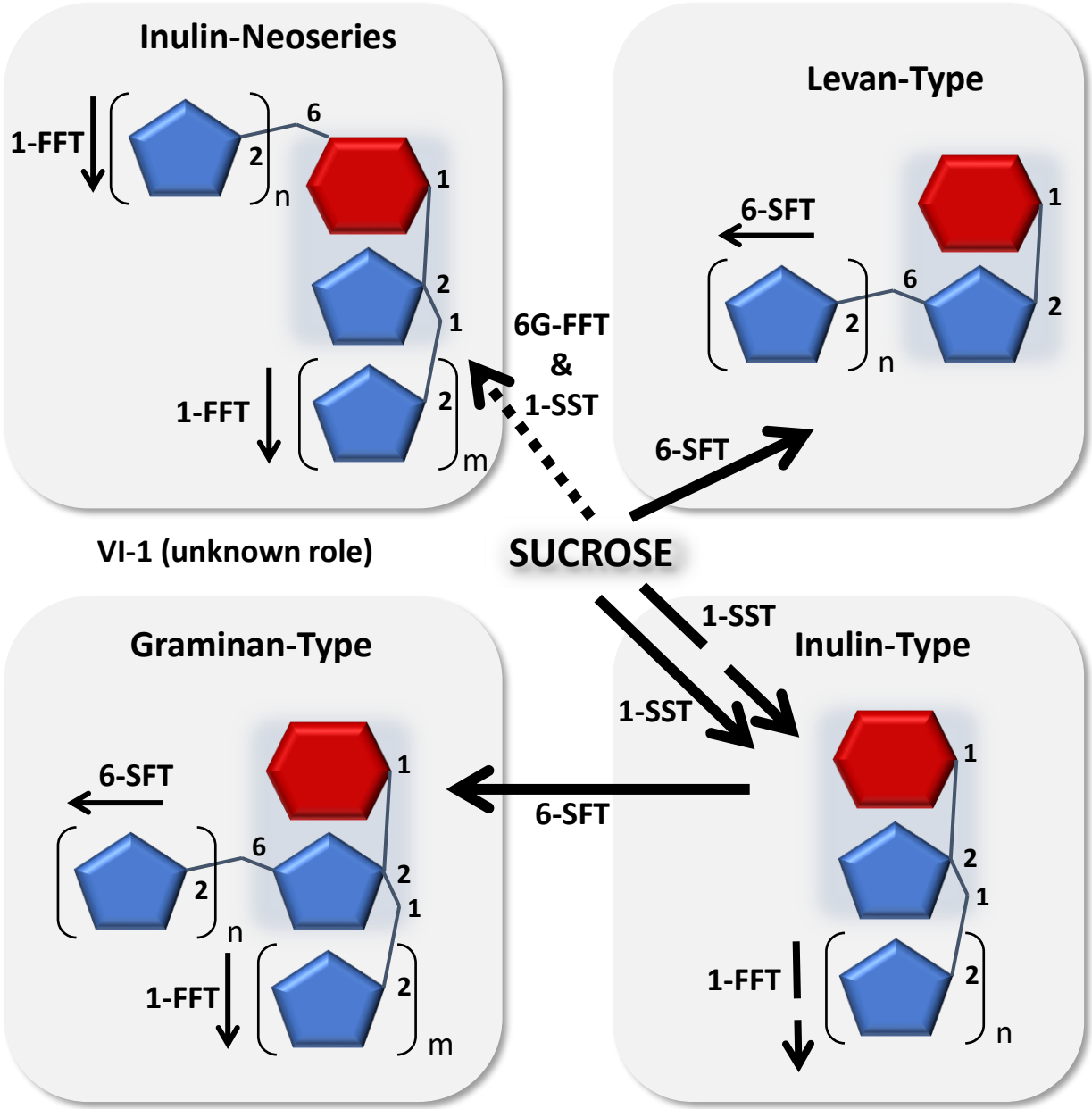
	SN
1-FFT	434
1-SST	138
6G-FFT	0
6-SFT	477
VI-1	194
6-FEH	551
6-FEH/CWI2	137
1-FEH	169

TPM Scale: Low 0-10 | Medium 10-250 | High >250

FPKM Scale: Low 0-100 | Medium 100-500 | High >500

0 in white box, no expression detected; 0 in light blue or light grey box, expression close to zero detected

Figure 5: Transcript expression of the fructan biosynthesis genes in barley across plant development and in various tissues. Data sets included in the analysis are detailed in the materials and methods section. Genes included comprise the five fructan biosynthesis genes from the detected significant QTL region (*1-FFT*, HORVU7Hr1G000250; *1-SST*, HORVU7Hr1G000260; *6G-FFT*, HORVU7Hr1G000270; *6-SFT*, HORVU7Hr1G001040; and *VI-1*, HORVU7Hr1G001070) as well as three known fructan hydrolyase encoding genes (*6-FEH*, HORVU2Hr1G109120; *6-FEH/CWI2*, HORVU2Hr1G118820; and *1-FEH*, HORVU6Hr1g011260). Expression levels are colour coded, whereas different scales were used for TPM and FPKM values as indicated in the legend. (A) shows expression levels in the early vegetative phase for whole germinated grain tissues and for isolated germinated grain tissues from 0 to 96 hours after imbibition (HAI). Also, expression levels in seedling are shown for germinated embryo (GE), root and shoot. (B) shows data from the reproductive phase. Vegetative tissues included are: EPI, epidermal strips (4 weeks after planting, W4); ROO2, roots (W4); RAC, inflorescences, rachis (W5); LEM, inflorescences, lemma (W6); LOD, inflorescences, lodicule (W6); PAL, dissected inflorescences, palea (W6); INF2, inflorescence (10 mm); and NOD, internode. Meiosis stages included are: A.Pre, premeiosis anthers; A.LepZyg, leptotene/zygotene anthers; M.Lep/Zyg leptotene/zygotene meiocytes; A.MetTet, metaphaseI-tetrad anthers; A.PacDip, pachytene/diplotene anthers; M.PacDip, pachytene/diplotene meiocytes. Waddington (W) stages for pistil development are: W8; W8.5; W9; W9.5; and W10. Isolated pistil tissues are: nucellus (including nucellus and embryo sac for W8.5); integument; ovary wall; ES, embryo sac; EC+CC, egg apparatus and central cell; ANT, antipodal cells; and chalaza. (C) shows data from the grain development phase for whole developing grain (from 7 to 20 days after pollination, DAP) and for isolated developing grain tissues from 7 to 25 DAP. Abbreviations are SA, sub-aleurone/outer starchy endosperm; and SE, starchy endosperm/inner starchy endosperm. Also, data from senescing leaf (SN) are presented. Other abbreviations are: At, anthesis; CI, collar initiation; Em, emergence; Hv, harvest; and Sw, sowing.



TFs associated with 1-FFT, 1-SST, 6-SFT:

- *WD_REPEATS_REGION domain-containing protein* (HORVU7Hr1G000820.1),
- *Protein ALWAYS EARLY 3* (HORVU7Hr1G001120.1)
- *Two scarecrow-like protein genes* (HORVU7Hr1G001300.3, HORVU7Hr1G001310.1)

TFs associated with 6G-FFT:

- *WD_REPEATS_REGION domain-containing protein* (HORVU7Hr1G000820.1)
- *Scarecrow-like protein 22* (HORVU7Hr1G001310.1)
- *HTH myb-type domain-containing protein* (HORVU7Hr1G001830.3)

TFs associated with VI-1:

- *Protein ALWAYS EARLY 3* (HORVU7Hr1G001120.1)
- *Scarecrow-like protein 22* (HORVU7Hr1G001310.1)

Pre-storage phase

Storage phase

Late storage phase

Figure 6: Fructan-types, suggested biosynthesis routes and potential regulators in developing barley grain. Specific spatiotemporal biosynthesis of oligofructans was observed for barley grains. Continuous bold arrows illustrate the major route of biosynthesis during the pre-storage phase (until 14 DAP) and the dashed bold arrows indicate the major route during the storage phase (until 20 DAP). During the pre-storage phase high transcript levels for *1-SST* and *6-SFT* were observed for the endosperm leading to an accumulation of 6-kestose and bifurcose. With transition to the storage phase a transcriptional switch was observed resulting in high transcript levels of *1-SST* in the nucellar projection (NP). *1-FFT* was found to be exclusively expressed in the NP during the storage phase. Induction of the inulin-type fructan biosynthesis pathway led to high amounts of 1-kestose and nystose accumulating in the endosperm cavity (Peukert *et al.*, 2014). The dotted bold arrow illustrates the major route of biosynthesis during the late storage phase with *6G-FFT* transcripts detected in the outer endosperm (from 30 DAP onwards, Figure 6), which matched the detection of neoseries-type oligofructans in mature barley grains (Figure 1). Transcription factors (TF) showing significant correlation of transcript expression pattern in developing grain with *1-FFT, 1-SST* and *6-SFT* (positive), *6G-FFT* (negative), and *VI-1* (positive) are listed in the inserted text box. Inulin-neoseries represents linear fructans with β(2,1) & β(2,6) linked fructosyl units at the glucose (1F, 6G-di-β-D-fructofuranosylsucrose is shown; m=1, n= 1), graminan-type represents branched fructans with β(2,1) & β(2,6) linked fructosyl units (bifurcose is shown; m= 1, n= 1), inulin-type illustrates linear fructans with β(2,1) linked fructosyl units (1-kestose is shown; n= 1); , and levan-type shows linear fructans with β(2,6) linked fructosyl units (6-kestose is shown; n= 1). The arrows indicate direction of further polymerisation. Abbreviations are: *1-FFT*, fructan:fructan 1-fructosyltransferase; *1-SST*, sucrose:sucrose 1-fructosyl-transferase; *6-SFT*, sucrose:fructan 6-fructosyltransferase; *6G-FFT*, fructan:fructan 6G-fructosyltransferase; *VI-1*, vacuolar invertase 1 (unknown role in developing grain).