Check for updates

# Phenotypic characterization and candidate gene analysis of a short kernel and brassinosteroid insensitive mutant from hexaploid oat (*Avena sativa*)

Nikos Tsardakas Renhuldt[1*†], Johan Bentzer[1], Dag Ahrén[2], Sofia Marmon[1] and Nick Sirijovski[1,3†]

[1]ScanOats Industrial Research Centre, Department of Chemistry, Division of Pure and Applied Biochemistry, Lund University, Lund, Sweden, [2]National Bioinformatics Infrastructure Sweden (NBIS), SciLifeLab, Department of Biology, Lund University, Lund, Sweden, [3]CropTailor AB, Department of Chemistry, Division of Pure and Applied Biochemistry, Lund University, Lund, Sweden

In an ethyl methanesulfonate oat (*Avena sativa*) mutant population we have found a mutant with striking differences to the wild-type (WT) cv. Belinda. We phenotyped the mutant and compared it to the WT. The mutant was crossed to the WT and mapping-by-sequencing was performed on a pool of F2 individuals sharing the mutant phenotype, and variants were called. The impacts of the variants on genes present in the reference genome annotation were estimated. The mutant allele frequency distribution was combined with expression data to identify which among the affected genes was likely to cause the observed phenotype. A brassinosteroid sensitivity assay was performed to validate one of the identified candidates. A literature search was performed to identify homologs of genes known to be involved in seed shape from other species. The mutant had short kernels, compact spikelets, altered plant architecture, and was found to be insensitive to brassinosteroids when compared to the WT. The segregation of WT and mutant phenotypes in the F2 population was indicative of a recessive mutation of a single locus. The causal mutation was found to be one of 123 single-nucleotide polymorphisms (SNPs) spanning the entire chromosome 3A, with further filtering narrowing this down to six candidate genes. In-depth analysis of these candidate genes and the brassinosteroid sensitivity assay suggest that a Pro303Leu substitution in AVESA.00010b.r2.3AG0419820.1 could be the causal mutation of the short kernel mutant phenotype. We identified 298 oat proteins belonging to orthogroups of previously published seed shape genes, with AVESA.00010b.r2.3AG0419820.1 being the only of these affected by a SNP in the mutant. The AVESA.00010b.r2.3AG0419820.1 candidate is functionally annotated as a GSK3/SHAGGY-like kinase with homologs in Arabidopsis, wheat, barley, rice, and maize, with several of these proteins having known mutants giving rise to brassinosteroid insensitivity and shorter seeds. The substitution in AVESA.00010b.r2.3AG0419820.1 affects a residue with a known gain-of function substitution in Arabidopsis BRASSINOSTEROID-INSENSITIVE2. We propose a gain-

of-function mutation in *AVESA.00010b.r2.3AG0419820.1* as the most likely cause of the observed phenotype, and name the gene *AsGSK2.1*. The findings presented here provide potential targets for oat breeders, and a step on the way towards understanding brassinosteroid signaling, seed shape and nutrition in oats.

# 1 Introduction

Hexaploid oat (*Avena sativa*) is a staple cereal known for its high content of beta-glucan dietary fiber, protein, lipids, and other nutrients such as avenanthramides [EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA) 2010; Zhang et al., 2021]. Nutrients are not uniformly distributed within oat kernels, and the proportion of nutrients differs across kernel tissues (Liu and Wise, 2021). It has previously been shown that seed shape and size are correlated to specialized metabolite abundance including beta-glucan content (Zimmer et al., 2021; Brzozowski et al., 2022). Disentangling the genetic determinants of seed size and shape could potentially help us understand connections to both oat nutrition and agricultural traits.

Hormones are involved in innumerable processes in plants, affecting plant growth and development in a multitude of ways. Mutations affecting plant hormone biosynthesis or signaling often result in phenotypes that alter many aspects of the plant, including plant size, architecture, placement of stomata, and a number of other properties, sometimes all at once (Smith et al., 2017). Grain size and shape have been extensively studied (Li et al., 2019) with several pathways shown to be involved in determining these phenotypes, many of them involving hormones. Seed shape is affected by both maternal and zygote genotypes. Phytohormone signaling plays a role in both maternal and zygote tissues. Brassinosteroids (BRs) are one class of plant hormones known to positively regulate seed size in several plant species (Tong et al., 2012; Jiang et al., 2013; Cheng et al., 2020; Kloc et al., 2020; Sun et al., 2021), promoting both cell expansion and cell proliferation in maternal tissues (Li et al., 2019). Auxins also affect seed shape through maternal tissues, with various auxin response factors having different effects on seed shape and size (Schruff et al., 2006; Hu et al., 2018; Jia et al., 2021). BRs affect seed size through zygotic tissues as well, regulating the HAIKU pathway together with abscisic acid, with cytokinins affecting seed size downstream of this pathway (Jiang et al., 2013; Li et al., 2013; Cheng et al., 2014). Auxins also affect seed size through zygotic tissues, with some controversy regarding which exact genes are involved in this (Bernardi et al., 2012; Ishimaru et al., 2013; Bernardi et al., 2019; Nonhebel et al., 2020; Kabir and Nonhebel, 2021). Beyond these, seed shape is influenced by a number of different mechanisms acting in maternal tissues, including G-protein signaling, transcriptional regulators, mitogen-activated protein kinase signaling, and through the ubiquitin-proteasome pathway (Li et al., 2019).

Extensive work has been performed in identifying genes affecting seed shape and size, with more than 80 genes and 400 quantitative trait loci (QTLs) identified only in rice (Li et al., 2022). Several of these are actively used in rice breeding programs (Zuo and Li, 2014). Mutants are commonly used to identify gene function. Mapping-by-sequencing is one way to map genes and does so by sequencing pools of recombinant plants sharing a phenotype of interest, similar to bulked segregant analysis (Schneeberger et al., 2009). For sufficiently large pools and with good enough phenotyping, recombination ensures that the mutant allele frequency will be close to one only in genomic regions that are close to the causative mutation, as the remaining mutations will have been lost randomly for the individuals within the pool. Mapping-by-sequencing has previously been employed to map the *AsCer-q* gene in oat (Kamal et al., 2022).

As a part of an oat TILLING population (Chawade et al., 2010), we have identified an oat (*Avena sativa*) mutant with short kernels, compact spikelets, and altered plant architecture when compared to the wild-type (WT) cultivar Belinda. In this paper, we characterize the mutant and provide evidence of BR insensitivity. We use a mapping-by-sequencing approach and the recently published oat cv. Sang reference genome (Kamal et al., 2022) to identify a likely causative gain-of-function mutation in a gene coding for a glycogen synthase kinase 3/SHAGGY-like kinase (GSK3/SHAGGY-like kinase, GSK), which we designate AsGSK2.1.

# 2 Materials and methods

## 2.1 Plant material and F2 segregating pools

The short kernel mutant was identified among the Belinda TILLING population (Chawade et al., 2010) during a field amplification experiment in 2017, and a forward genetics investigation of the genetic cause was initiated. Compared to the parental cv. Belinda, cv. Sang, and other members of the TILLING population, the mutant exhibited distinctly compact spikelets and short kernels and, to a lesser extent, altered plant architecture. The

mutant phenotype was stable under greenhouse conditions with an 18 h photoperiod. A cross was made using Belinda as mother and the mutant as pollen donor (Belinda × mutant). The resulting $F_1$ seeds were germinated in soil in a greenhouse and allowed to self-fertilize. Upon ocular inspection, the Belinda × mutant $F_1$ plants resembled Belinda, indicative of the mutation being recessive. An $F_2$ population was raised as in Kamal et al. (2022). Leaf tissue from 15 short kernel $F_2$ plants was pooled for DNA extraction and sequencing. To study the mutant phenotype in more detail, eight non-backcrossed mutant plants and eight Belinda plants were grown in parallel in a greenhouse under ambient light conditions in Lund, Sweden, spring 2022.

## 2.2 Phenotyping

### 2.2.1 Shoot and seed measurements

Photos of the growing plants were taken, and the height and the number of tillers at maturity were recorded. The primary tiller and two secondary tillers were individually collected for each plant, and the rest of the seeds were collected and threshed together using a Halderup LT-15 laboratory thresher. For a minimum of five plants, the primary panicle was characterized, and its total length, length of main branch of the first node (i.e., first whorl), and bottom panicle node distance were measured using a ruler. The length of the lower glume for the five top spikelets of the primary tiller was also measured with a ruler. The number of spikelets, the number of primary, secondary and tertiary seeds divided for each whorl, and the number of whorls were manually counted. Manually harvested seeds, divided into groups of primary, secondary, and tertiary seeds, were photographed and photos analyzed using the program SmartGrain (Tanabata et al., 2012) to determine seed length, seed width, area, perimeter length, and circularity.

### 2.2.2 Macromolecular composition

The overall macromolecular composition was analyzed on threshed seeds, i.e., whole grains including hulls (n = 8) using near infrared spectroscopy with a handheld GrainSense calibrated for oat, giving an estimate of total lipids, protein, carbohydrates, and water content. Macromolecular composition was then measured using biochemical methods as described below for a selection of the samples. Around 40 of the seeds from each plant were manually dehulled and milled in a Fritsch Pulverisette 23 at 50 oscillations/s for 2 min. Milled oat was stored at −20°C until analysis.

### 2.2.3 Protein content

Protein (n = 4) was measured using the Dumas method (Thermo Electron Corp., Flash EA, 1112 Series, Waltham, MA, USA, AOAC International, 2019, Method 990.03). For each sample, 25–27 mg was analyzed on a 33 mm tin disc. A tin disk without added sample was used for blank, aspartic acid was used as reference, and 5.4 (Mariotti et al., 2008) was used as nitrogen-to-protein conversion factor.

### 2.2.4 Lipid content

Total lipids (n = 4) were extracted using chloroform:methanol at a 1:1 ratio according to a protocol adapted from Bligh and Dyer

(1959). Shortly, the extraction used 3.75 mL of methanol/chloroform (2:1 v/v, with 0.05% butylated hydroxytoluene (BHT) w/v), 1 mL of 0.15 M acetic acid, 1.25 mL of chloroform, and 1.25 mL of water, with vortexing between each addition. An amount of 100 mg of milled seed material was used for extraction together with an internal standard of 0.40 mg of tripentadecanoin (Larodan). Extracted lipids were methylated according to Cavonius et al. (2014) and run on Thermo Scientific TRACE GC 1300-1310 Gas Chromatograph with a NUKOL™ fused silica capillary column (Supelco™). The GC program started at 80°C with 10°C increase per min to 150°C, 7°C increase per min until 215°C, and held at 215°C until a total runtime of 40 min. GLC426 (NuCheck) was used as an external reference standard.

### 2.2.5 Beta-glucan content

Beta-glucan (n = 4) was analyzed using the McKleary method with the Megazyme kit (Megazyme, AOAC International, 2019, Method 995.16). The method was scaled down to work in 2 mL of sample tubes, using 2 × 15 mg (technical duplicates) of oat as starting material.

### 2.2.6 Water content

Water content (n = 4) was measured gravimetrically on an analytical balance before and after freeze drying until stable weight.

### 2.2.7 Brassinosteroid sensitivity

The sensitivity to BR was investigated with the leaf unrolling test (Wada et al., 1985; as described by Chono et al. (2003)). Samples were planted in vermiculite grown in darkness at 24°C for 8–9 days for mutant and WT plants. Leaf segments of 1.5 cm were incubated for 72 h in $10^{-5}$ M Epibrassinolide (Sigma-Aldrich) with water as control. The level of unrolling was measured using digital calipers. At least seven samples were used for the mutant and WT, respectively, and the experiment was repeated three times.

### 2.2.8 Statistics

Phenotype statistics were calculated in R (R Core Team, 2023, v4.2.3), with Wilcoxon rank sum tests used to compute p-values, and values of p < 0.05 were considered significant. The tidyverse packages (Wickham et al., 2019) and gtsummary (Sjoberg et al., 2021) were used in data processing and table creation.

## 2.3 DNA extraction, genome resequencing, and variant calling and analysis

Whole-genome short-read resequencing of the pooled tissue was done as that in the work of Kamal et al. (2022). The mutant tissue pool was ground to a fine powder using liquid nitrogen in a mortar and pestle. Genomic DNA was isolated using the Qiagen DNeasy Plant Mini Kit according to the manufacturer's protocol. A library of 350-bp insert size was prepared using the Illumina TruSeq PCR-free library preparation kit (Illumina, San Diego, CA). Samples were sequenced on NovaSeq6000 (NovaSeq Control Software 1.6.0/RTA v3.4.4) with a 2 × 151 setup using "NovaSeqXp" workflow in "S1" or "SP" mode flowcell to generate reads corresponding to 30× coverage. The Bcl to FastQ conversion

was performed using bcl2fastq_v2.20.0.422 from the CASAVA software suite. The quality scale used is Sanger/phred33/Illumina 1.8+. The reads were trimmed using fastp (Chen et al., 2018; v0.20.1, options - -correction - -adapter_sequence=AGATCGGAAGAG CACACGTCTGAACTCCAGTCA- -adapter_sequence_r2=A GATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT), mapped to the *Avena sativa* cv. Sang reference genome (Kamal et al., 2022) using BWA-MEM2 (Vasimuddin et al., 2019; v2.2.1), and variants were called using DeepVariant (Poplin et al., 2018; v1.1.0), GNU Parallel (Tange, 2021; v20210422), and GLNexus (Lin et al., 2018; Yun et al., 2021; v1.3.1). The full variant calling pipeline is available at GitHub at nikostr/dna-seq-deepvariant-glnexus-variant-calling (Tsardakas Renhuldt, 2021; v0.3.1) and was run using Snakemake (Mölder et al., 2021; v6.5.1). Variants in independent lines from the TILLING population as well as in Sang or Belinda (Supplementary Table S1) were called as a part of Kamal et al. (2022) and filtered out using bcftools (Danecek et al., 2021, v1.12). SnpEff (Cingolani et al., 2012, v4.3.1t) and v1.1 of the Sang annotation of protein coding genes (Kamal et al., 2022) were used to annotate effects of single-nucleotide polymorphisms for the variants.

The mutant allele frequency of variants with a total allelic depth ≥ 15 was plotted and used to identify conserved regions among the pooled samples. Variants were filtered on the basis of read support (read depth of alternate allele ≥ 15, no reads supporting the reference allele), being located in a region found to be conserved in the pool, being classified by SnpEff as having moderate or high impact, and being located in a gene with expression (>0.5 transcripts per million) in all seed, glume, and spikelet tissue samples. Gene expression values in transcripts per million (three developing seed samples from cv. Sang, seven glume samples, and eight spikelet samples from cv. Belinda, and 21 additional developing seed samples) as provided by Kamal et al. (2022) were used (Supplementary Table S2). The filtered genes were investigated individually. Orthogroups from Kamal et al. (2022) were used to identify genes homologous to the genes affected in the mutant. Multiple sequence alignments were produced using MUSCLE (Edgar, 2004, v3.8.1551), with fasttree (Price et al., 2010; v2.1.10) used to build phylogenetic trees, which, in turn, were visualized using iTOL (Letunic and Bork, 2021; v6.7).

Oat genes potentially involved in the regulation of seed shape were also identified using the orthogroups and a recent review of such genes (Li et al., 2019). Only genes from species included in the orthogroups (*Arabidopsis*, rice, wheat, and maize) were considered. The Sang positions of the markers used by Zimmer et al. (2021) were taken from Tinker et al. (2022). We used KofamScan (Aramaki et al., 2020), v1.3.0, to search the Sang proteins against the eukaryote-specific profiles in the Kofam database (https://www.genome.jp/ftp/db/kofam/archives/2021-12-01/). Only KEGG Orthology matches with scores above the pre-computed KofamScan thresholds were considered.
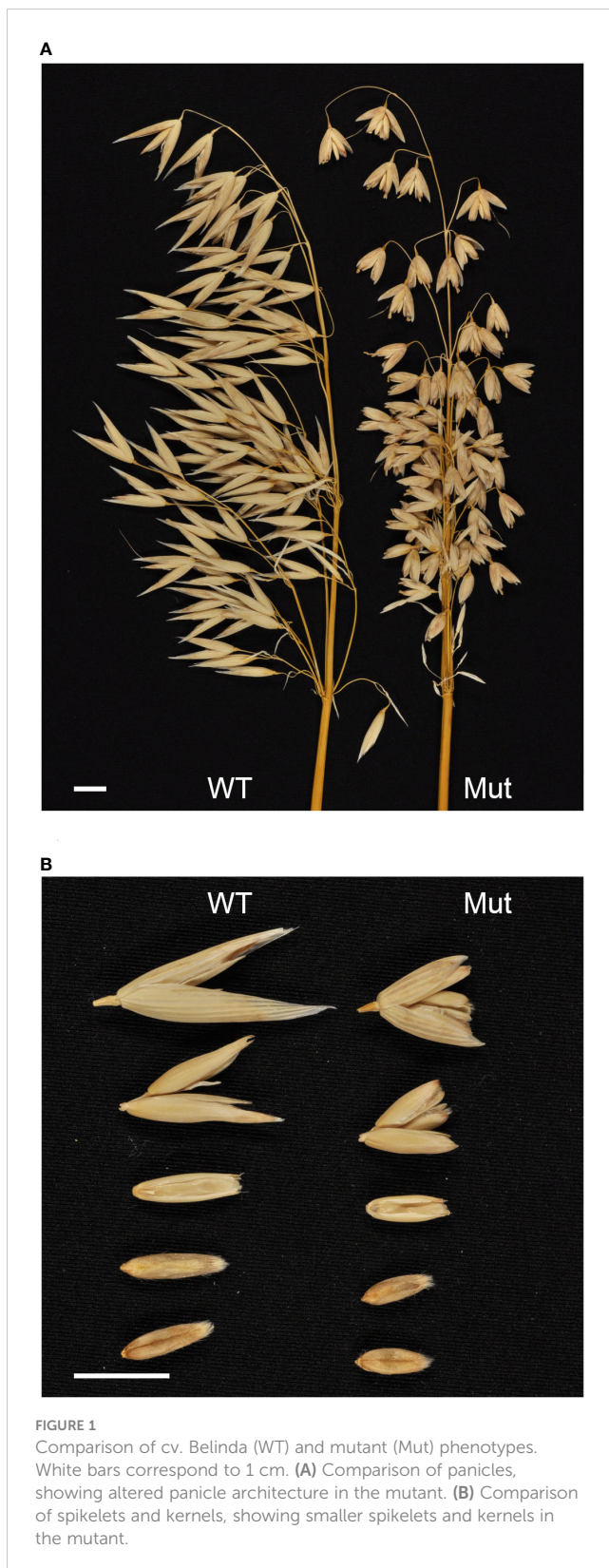
# 3 Results

The oat mutant was initially identified in a TILLING population (Chawade et al., 2010) due to its strikingly different panicles. Out of the 69 $F_2$ plants, 15 shared the short kernel phenotype. This ratio is in line with what would be expected of a recessive trait caused by a single locus (null hypothesis is that homozygous mutant allele: heterozygous + homozygous WT allele have distribution 1:3, $x^2(1, N = 69) = 0.39$, $P = 0.53$). Figure 1 shows a comparison of the mutant to cv. Belinda. The mutant kernels and glumes were significantly (Wilcoxon rank sum test, $p \leq 0.03$) shorter than the corresponding tissues in the mother variety Belinda; see Table 1. Only the length of the seeds and associated tissues were affected, no change was seen in width, nor were the seeds misshaped, as shown in Figure 1B. The shorter seeds also had a significantly lower seed area (mutant, 18.9 ± 1.3 mm$^2$; cv. Belinda, 28 ± 0.5 mm$^2$) and thousand kernel weight (mutant, 27 ± 2.2 g; cv. Belinda, 39 ± 2.6 g). Studying the plant phenotype, it was evident that many different parts of the plant were affected and that the effects were not limited to the seeds and associated tissues being short. Total plant height, both at intermediate growth (observed, not quantified) and at maturity, as well as distance between bottom panicle nodes and length of the main branch of the first node were significantly shorter than in the control plants. Total panicle length in the mutant was not shorter than the control, despite the shorter node distances. This seems to be a result of altered panicle architecture, with, in general, one additional whorl. The higher whorls (three and up) also contained a significantly higher number of seeds in the short kernel panicles. Further, the mutant spikelets also more often contained tertiary grains. The nutritional quality of the short kernel oat was also somewhat altered, with the protein content being higher than in the mother cultivar (mutant, 19.3 ± 1.3%; cv. Belinda, 16.7 ± 1.0%, according to the Dumas method described above), and a corresponding lower carbohydrate content (Supplementary Table S3). The beta-glucan content and lipid content were similar in both seed types.

The mutant allele frequency was higher across chromosome 3A than throughout the remainder of the genome, as shown in Figure 2A. The number of genes remaining when performing filtering as outlined above is shown in Figure 2B, with a total of six genes remaining following all filtering steps: characteristics and orthology information regarding these candidate genes are shown in Tables 2 and 3, respectively. Supplementary Table S4 connects previously published gene names to the gene identifiers present in the orthogroups. All six candidate genes had other expressed oat genes present in the same orthogroup (Supplementary Figure S1). Out of the six genes, two had orthologs known to have some effect on seed shape.

The mutant had a C to T substitution at chr3A:64189534 (32 reads supporting the alternate allele, no reads supporting the reference allele, QUAL value of 61.7) affecting AVESA.00010b.r2.3AG0419820.1 (Supplementary Figure S2). This leads to a Pro303Leu substitution in the protein, which is annotated as a SHAGGY-like kinase. The protein belongs to the same orthogroup (OG0000448) as *Arabidopsis* BIN2 (also known as AtSK21 and UCU1), as well as the barley HvGSK proteins, and OsSK/OsGSK from rice. It is also homologous to TaSG-D1 in wheat. Figure 2C; Supplementary Figure S3 show multiple sequence alignments of homologous proteins previously known from the literature, including positions of known gain-of-function substitutions from *Arabidopsis* and wheat. Supplementary Figure S4 shows a phylogenetic tree of OG0000448

**FIGURE 1**

Comparison of cv. Belinda (WT) and mutant (Mut) phenotypes. White bars correspond to 1 cm. **(A)** Comparison of panicles, showing altered panicle architecture in the mutant. **(B)** Comparison of spikelets and kernels, showing smaller spikelets and kernels in the mutant.

and TaSG-D1 with subgroups indicated along with proteins previously known from other species (Cheng et al., 2020; Kloc et al., 2020; Li et al., 2021; Hou et al., 2022; Li et al., 2022; Zhang et al., 2022). The Pro303Leu substitution affects the same residue as the weak gain-of-function mutation *ucu1-3* in *Arabidopsis*. These

homologous proteins are known negative regulators of BR signaling, with mutants known to affect kernel size and plant architecture (Pérez-Pérez et al., 2002; Tong et al., 2012; Cheng et al., 2020; Kloc et al., 2020). The second candidate protein with orthologs known to affect seed shape is encoded by *AVESA.00010b.r2.3AG0437710.1*, orthologous to the transcriptional corepressor LEUNIG and OsLUGL. When *OsLUGL* is mutated in rice, it results in seeds that are deformed and reduced seed set, neither of which were observed in the short kernel mutant. Mutation in *OsLUGL* also does not affect plant height and overall architecture (Yang et al., 2019), whereas the contrary was observed in the mutant studied here. Further, we had identified a mutation in *AVESA.00010b.r2.3AG0437710.1* in another oat line lacking the phenotype (data not shown). Thus, the Gly558Asp substitution in AVESA.00010b.r2.3AG0437710.1 is unlikely to be the cause of the mutant phenotype. A functional orthology analysis of the remaining four candidate genes (*AVESA.00010b.r2.3AG0417810.1*, *AVESA.00010b.r2.3AG0424310.1*, *AVESA.00010b.r2.3AG0426560.1*, and *AVESA.00010b.r2.3AG0430410.1*) revealed potential biological functions that include a chlorophyll reductase, exonuclease, hydroxyproline O-arabinosyltransferase, and a DCD domain (Development and Cell Death)–containing protein, respectively. A detailed investigation of the literature on orthologs of these proteins did not provide support for these candidates playing a role in seed shape or plant architecture, and we therefore considered the identified mutations in these genes to be very unlikely to be responsible for the mutant phenotype.

To further study if the identified amino acid substitution in the SHAGGY-like kinase candidate AVESA.00010b.r2.3AG0419820.1 affects BR signaling in the oat mutant, a leaf unrolling test was performed (Wada et al., 1985; as described by Chono et al. (2003)). The Belinda control plants were significantly more sensitive to BR as shown by the test, with the leaves being $90(\pm14)\%$ unrolled compared to the mutant leaves being only $58(\pm11)\%$ unrolled (Wilcoxon rank sum test, $p = 0.004$) with no difference (Wilcoxon rank sum test, $p = 0.13$) between the samples subjected to water, $36(\pm 6)\%$ compared to $39(\pm 2)\%$ unrolling for the Belinda control and the mutant, respectively.

To help exclude alternate genes causing the phenotype, gene candidates were identified on the basis of homology to previously published seed shape genes from other species. Out of the 89 *Arabidopsis*, wheat, rice, and maize proteins listed in the review of regulators of seed size control written by Li et al. (2019), an orthogroup was identified for 86 of these, corresponding to a total of 63 orthogroups containing a total of 298 oat proteins, five of which are encoded by genes located on chromosome 3A. Among all oat genes identified through these orthogroups, only the SHAGGY-like kinase *AVESA.00010b.r2.3AG0419820.1* contains a mutation classified as having at least moderate impact by SnpEff. The full list of candidate seed shape proteins can be found in Supplementary Table S5, and the locations of their corresponding genes are shown together with the seed shape markers from Zimmer et al. (2021) in Supplementary Figure S5.

Taking the above into account, we propose the Pro303Leu substitution in AVESA.00010b.r2.3AG0419820.1 is a gain-of-function mutation that is most likely among the six candidates to cause the oat short kernel phenotype reported herein.

TABLE 1  Characteristics of the plants and the seeds of cv. Belinda and the mutant.

| Characteristic | Belinda | | Mutant | | |
|---|---|---|---|---|---|
| | N[1] | Mean (SD[2]) | N[1] | Mean (SD[2]) | p-value[4] |
| Plant height (cm) | 8 | 118.9 (6.9) | 8 | 103.1 (2.9) | <0.001 |
| Main panicle length (mm) | 6 | 212.5 (10.3) | 6 | 197.7 (14.2) | 0.054 |
| Main branch of first node (mm) | 6 | 118.8 (13.2) | 6 | 93.3 (1.6) | 0.005 |
| Bottom panicle node distance (mm) | 6 | 59.0 (3.3) | 6 | 48.0 (3.2) | 0.005 |
| Lower glume length (mm) | 6 | 21.6 (0.7) | 6 | 10.8 (0.3) | 0.005 |
| Kernel length including hull (mm)[3] | 4 | 13.6 (0.3) | 4 | 8.6 (0.4) | 0.030 |
| Kernel width including hull (mm)[3] | 4 | 2.8 (0.1) | 4 | 2.9 (0.1) | 0.4 |
| Kernel area (sq. mm)[3] | 4 | 28.0 (0.5) | 4 | 18.9 (1.3) | 0.030 |
| Number of kernels | 6 | 140.2 (22.2) | 6 | 179.0 (29.7) | 0.045 |
| Number of tertiary kernels | 6 | 1.5 (2.1) | 6 | 16.5 (7.9) | 0.005 |
| TKW for primary and secondary kernels (g) | 6 | 39.0 (2.6) | 6 | 26.7 (2.2) | 0.005 |

[1]Number of individuals used to calculate mean and standard deviation (SD).
[2]Standard deviation.
[3]P-values are based on Wilcoxon rank sum test comparisons of Belinda and the mutant.
[4]Based on averages of >98 seeds per replicate.

Following accepted naming convention, we designate AVESA.00010b.r2.3AG0419820.1 as AsGSK2.1.

# 4 Discussion

Above, we have identified a substitution in the protein encoded by *AVESA.00010b.r2.3AG0419820.1* as the likely cause of the mutant phenotype. Beyond the experimental evidence presented here, it is the only homolog of seed shape proteins identified in Li et al. (2019) that is found to be affected in the mutant. The candidate is homologous to GSK3 proteins, including *Arabidopsis* BIN2, as well as TaSG-D1 in wheat, the HvGSK proteins in barley, and OsSK/OsGSKs in rice. In *Arabidopsis*, the GSK3/SHAGGY-like kinase BIN2 has been extensively characterized as a transcriptional corepressor of BR signaling, with several known gain-of-function mutations that give a BR-deficient phenotype, including dwarfing (Pérez-Pérez et al., 2002). It has also been shown to interact with and phosphorylate auxin response factors in *Arabidopsis*, with its homolog OsSK41/OsGSK5 similarly phosphorylating OsARF4 in rice (Cho et al., 2014; Hu et al., 2018). In wheat, barley and rice, there is evidence that these affect kernel shape and size, with the gain-of-function mutation affecting TaSG-D1 being known to produce small, round seeds as well as altered plant architecture in *Triticum sphaerococcum* (Cheng et al., 2020). The Pro303Leu substitution in the protein encoded by the candidate gene causing a gain-of-function, similar to the *Arabidopsis ucu1-3* mutation (Pérez-Pérez et al., 2002), would help explain how this single mutation causes the mutant phenotype in spite of there being a total of 25 oat proteins with evidence of RNA expression present in the same orthogroup. That the phenotype of the mutant studied here diverges from the WT in many different parts of the plant is in line with what is seen for other mutants of GSK3 proteins (see, e.g., Song et al. (2023) and Li et al. (2021)). That the

causative mutation affects hormone signaling is also expected given the systemic effect seen in the mutant.

Youn and Kim (2015) proposed to unify the naming of the GSK3/SHAGGY-like kinases in *Arabidopsis* and rice, with the first digit of the gene name given by the subgroup (I–IV). This naming scheme has not been universally adopted. In this paper, we opt for a naming scheme similar to that used in barley, where the digit prior to the dot indicates the subgroup. Subgroups are indicated in Supplementary Figure S4.

For growers of oat, Yan et al. (2013) identify several ideal traits for a milling oat cultivar: high and stable grain yield, resistance to lodging, proper maturity, resistance to relevant biotic and abiotic stresses, high test weight, large kernels, and high straw yield. Among the ideal traits relevant for consumers, they include a high beta-glucan content. The identification of *AsGSK2.1* and its homologs provides a potential target for oat breeders. Downregulating the *HvGSK* genes has previously been shown to give barley with higher thousand kernel weight under normal growing conditions, as well as increased biomass in both normal growing conditions and under salt stress (Kloc et al., 2020). In rice, upregulation of *OsSK22/OsGSK2* gives rise to short, round grains, and suppression of the same gene gives longer grains (Tong et al., 2012). Loss-of-function mutations affecting *OsSK41/OsGSK5* or *OsSK23/OsGSK3* cause larger and heavier grain (Hu et al., 2018; Gao et al., 2019). Downregulation or knock-outs of *AsGSK2.1* could potentially provide oats with larger kernels and increased oat yields. Achieving such downregulation might seem daunting given the large number of oat genes in the orthogroup, but it has been shown that expression is positively correlated among the paralogs in barley (Kloc et al., 2020), meaning that it may not be necessary to target all 25 oat genes of the orthogroup to see effects on grain shape and weight.

Beyond targeting the identified candidate gene and its homologs directly, these findings may function as a stepping stone in helping to understand BR signaling in oats and monocots more widely. BR
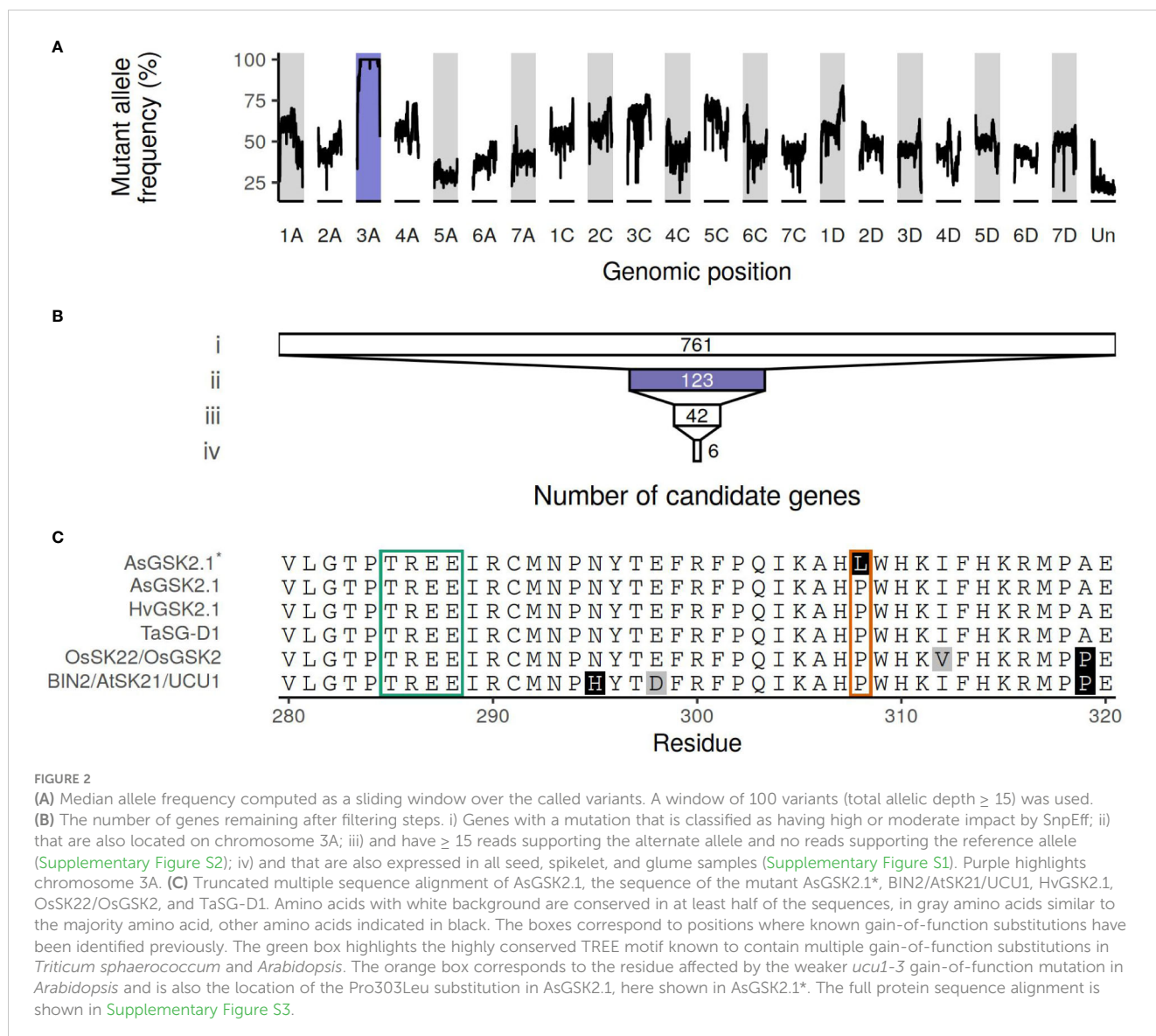
FIGURE 2
**(A)** Median allele frequency computed as a sliding window over the called variants. A window of 100 variants (total allelic depth ≥ 15) was used.
**(B)** The number of genes remaining after filtering steps. i) Genes with a mutation that is classified as having high or moderate impact by SnpEff; ii) that are also located on chromosome 3A; iii) and have ≥ 15 reads supporting the alternate allele and no reads supporting the reference allele (Supplementary Figure S2); iv) and that are also expressed in all seed, spikelet, and glume samples (Supplementary Figure S1). Purple highlights chromosome 3A. **(C)** Truncated multiple sequence alignment of AsGSK2.1, the sequence of the mutant AsGSK2.1*, BIN2/AtSK21/UCU1, HvGSK2.1, OsSK22/OsGSK2, and TaSG-D1. Amino acids with white background are conserved in at least half of the sequences, in gray amino acids similar to the majority amino acid, other amino acids indicated in black. The boxes correspond to positions where known gain-of-function substitutions have been identified previously. The green box highlights the highly conserved TREE motif known to contain multiple gain-of-function substitutions in *Triticum sphaerococcum* and *Arabidopsis*. The orange box corresponds to the residue affected by the weaker *ucu1-3* gain-of-function mutation in *Arabidopsis* and is also the location of the Pro303Leu substitution in AsGSK2.1, here shown in AsGSK2.1*. The full protein sequence alignment is shown in Supplementary Figure S3.

TABLE 2  Functional annotation, KEGG orthology, and mutation of the proteins corresponding to the six candidate genes.

| Candidate | Functional annotation | KEGG orthology | Mutation |
|---|---|---|---|
| AVESA.00010b.r2.3AG0417810.1 | NAD(P)-binding Rossmann-fold superfamily protein | K13606 (chlorophyll(ide) b reductase [EC:1.1.1.294]) | Gly122Glu |
| AVESA.00010b.r2.3AG0419820.1 | SHAGGY-like kinase | K14502 (protein brassinosteroid insensitive 2 [EC:2.7.11.1]) | Pro303Leu |
| AVESA.00010b.r2.3AG0424310.1 | RNA exonuclease 4 | K18327 (R exonuclease 4 [EC:3.1.-.-]) | Gly273Asp |
| AVESA.00010b.r2.3AG0426560.1 | Hyp O-arabinosyltransferase-like protein | K20782 (hydroxyproline O- arabinosyltransferase [EC:2.4.2.58]) | Ala76Val |
| AVESA.00010b.r2.3AG0430410.1 | DCD (Development and Cell Death) domain protein | | Pro290Ser |
| AVESA.00010b.r2.3AG0437710.1 | Transcriptional corepressor LEUNIG | | Gly558Asp |

TABLE 3  Notable published homologs of the candidate genes.

| Candidate | Description | Homolog | Reference |
|---|---|---|---|
| *AVESA.00010b.r2.3AG0417810.1* | | | |
| | NYC1 encodes a chlorophyll b reductase. Deleterious mutations in this gene leads to hindered chlorophyll degradation/yellowing during sencescense. | AtNYC1 OsNYC1 | Horie et al. (2009) Kusaba et al. (2007) |
| *AVESA.00010b.r2.3AG0419820.1* | | | |
| | BIN2 is known to be involved in BR signaling. Mutations affecting the TREE motif are known to cause gain-of-function, repressing BR signaling. In barley and rice knockouts of homologous genes lead to increased grain size. Gain-of-function mutation in the wheat ortholog and overexpression of rice ortholog OsSK22 leads to smaller, rounder kernels and shorter stature. | BIN2/AtSK21 | Pérez-Pérez et al. (2002); Peng and Li (2003) |
| | | HvGSKs | Kloc et al. (2020) |
| | | OsSK22, OsSK23, OsSK41 | Tong et al. (2012); Hu et al. (2018); Gao et al. (2019) |
| | | TaSG-D1 | Cheng et al. (2020) |
| *AVESA.00010b.r2.3AG0424310.1* | | | |
| | The homologue in Arabidopsis contains exonuclease domain, but has no documented effect on miRNA. | AT3G15080.1 | Ramachandran and Chen (2008) |
| *AVESA.00010b.r2.3AG0426560.1* | | | |
| | MtRDN1 is involved in root nodule formation in Medicago truncatula. OsRDN1 is capable of rescuing MtRDN1 mutants. HPAT3 is involved in pollen tube formation in Arabidopsis. | MtRDN1 | Schnabel et al. (2011); Kassaw et al. (2017) |
| | | OsRDN1 | Schnabel et al. (2011); Kassaw et al. (2017) |
| | | HPAT3 | Beuder et al. (2020) |
| *AVESA.00010b.r2.3AG0430410.1* | | | |
| | Involved in *Arabidopsis* stress response and inhibits cell death. | NRP | Yang et al. (2021) |
| *AVESA.00010b.r2.3AG0437710.1* | | | |
| | Transcriptional corepressor involved in floral organ identity. Mutants are known to have reduced fertility, and may produce abnormal seeds. | LEUNIG OsLUGL | Liu and Meyerowitz (1995) Yang et al. (2019) |

signaling in cereals and monocots, in general, is not as well understood as that in *Arabidopsis*, in spite of BR playing an important role affecting yield and yield-related agronomic traits. Of particular note are semi-dwarf cultivars with lower risk of lodging and more erect cultivars allowing for increased planting density, which, in turn, contributes to a higher yield. BR affects both of these phenotypes (Gruszka, 2020). Even though the GSKs are a conserved part of BR signaling across multiple species including cereals, further studies of this mutant and the candidate gene could provide insight into other parts of oat BR signaling, which may be of relevance in other cereals as well.

In the populations studied by Zimmer et al. (2021), they identified negative correlations between kernel width/thickness and beta-glucan content, but no real correlation between kernel length and beta-glucan content. Our measurements show no significant difference neither in the kernel width nor in the beta-glucan content between the mutant and cv. Belinda. If there is no correlation between beta-glucan content and kernel length and if suppressing *AsGSK* genes does indeed produce longer seeds similar to the effect of suppressing *OsGSK2/OsSK22* (Tong et al., 2012),

then suppressing the *AsGSK* genes may potentially produce larger seeds without negatively affecting the beta-glucan content.

QTLs for kernel length, width, and thickness were also identified by Zimmer et al. (2021), with length QTLs being located in Mrg06, Mrg21, and Mrg24, and QTLs for width and thickness located in Mrg13. Using the recently published reference genomes and their corresponding nomenclature (International Oat Nomenclature Committee, 2021), these merge groups correspond to chromosomes 5D, 4D, 5A, and 2C, respectively. Out of the genes corresponding to the proteins present in OG0000448, one is located on 2C and two are located on 4D, but, on both chromosomes, there are other potential candidates located closer to the significant markers. None of the significant markers from Mrg24 are present on chromosome 5A in the Sang assembly. On chromosome 2C, we note that a gene coding for a protein in the same orthogroup as DA1, DAR1, and ZmDA1 (OG0008367, AVESA.00010b.r2.2CG0324970) is expressed in the relevant tissues and is located in the middle of the region containing significant markers. A member of the same orthogroup as the known grain width related protein GW8 (OG0002672, AVESA.00010b.r2.2CG0325500) is located very close

to these markers as well, but it is not present in the expression data. We also note that genes with homologs known to affect grain length are located close to the grain length QTLs on chromosomes 4D and 5D (AVESA.00010b.r2.4DG0789350 in orthogroup OG0000353 with TGW6; AVESA.00010b.r2.5DG0996410 in OG0017199 with GS3) but that these are not reliably expressed in the expression data that we consider here. The above is one demonstration of how the list of potential grain shape genes presented in Supplementary Table S5 may be used, and we hope that it will prove to be a useful resource for breeders as well as for future studies of kernel shape in oats.

We have identified a likely gain-of-function mutation in *AsGSK2.1* (*AVESA.00010b.r2.3AG0419820.1*), as the most likely candidate mutation causing the mutant to display a short stature, compact panicles, short kernels, and an altered plant architecture. We have also identified oat proteins homologous to proteins previously known to affect kernel shape in other plant species. These findings provide potential targets for oat breeders, and a step on the way toward understanding BR signaling, seed shape, and nutrition in oats.

## Data availability statement

The raw whole genome resequencing data have been deposited in the European Nucleotide Archive at EMBL-EBI under accession number PRJEB57056 https://www.ebi.ac.uk/ena/browser/view/PRJEB57056. The KEGG Orthology is available at doi: 10.5281/zenodo.7886703.

## Author contributions

NT: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. JB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. DA: Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing. SM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. NS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

## Funding

## Acknowledgments

## Conflict of interest

The authors declare that this study received funding from Lantmännen and Oatly AB as the industry support component of the ScanOats Research Center. The funders were not involved in the study design, data collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Author NS was employed by CropTailor AB and is currently employed by Oatly AB.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1358490/full#supplementary-material.

# References

AOAC International. (2019). *Official methods of analysis of AOAC international. 1st ed.*Ed. G. W. Latimer (Rockville, Maryland: AOAC International).

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2020). KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36, 2251–2525. doi: 10.1093/bioinformatics/btz859

Bernardi, J., Battaglia, R., Bagnaresi, P., Lucini, L., and Marocco, A. (2019). Transcriptomic and metabolomic analysis of zmYUC1 mutant reveals the role of auxin during early endosperm formation in maize. *Plant Sci.* 281, 133–145. doi: 10.1016/j.plantsci.2019.01.027

Bernardi, J., Lanubile, A., Li, Q.-B., Kumar, D., Kladnik, A., Cook, S. D., et al. (2012). Impaired auxin biosynthesis in the defective endosperm18 mutant is due to mutational loss of expression in the zmYuc1 gene encoding endosperm-specific YUCCA1 protein in maize. *Plant Physiol.* 160, 1318–1285. doi: 10.1104/pp.112.204743

Beuder, S., Dorchak, A., Bhide, A., Moeller, S. R., Petersen, B. L., and MacAlister, C. A. (2020). Exocyst mutants suppress pollen tube growth and cell wall structural defects of hydroxyproline O-arabinosyltransferase mutants. *Plant J.* 103, 1399–4195. doi: 10.1111/tpj.14808

Bligh, E.G., and Dyer, W.J. (1959). A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37, 911–175.

Brzozowski, L. J., Hu, H., Campbell, M. T., Broeckling, C. D., Caffe, M., Gutiérrez, L., et al. (2022). Selection for seed size has uneven effects on specialized metabolite abundance in oat (Avena sativa L.). *G3 Genes|Genomes|Genetics* 12, jkab419. doi: 10.1093/g3journal/jkab419

Cavonius, L. R., Carlsson, N.-G., and Undeland, I. (2014). Quantification of total fatty acids in microalgae: comparison of extraction and transesterification methods. *Analytical Bioanalytical Chem.* 406, 7313–7322.

Chawade, A., Sikora, P., Bräutigam, M., Larsson, M., Vivekanand, V., Nakash, M. A., et al. (2010). Development and characterization of an oat TILLING-population and identification of mutations in lignin and β-glucan biosynthesis genes. *BMC Plant Biol.* 10, 865. doi: 10.1186/1471-2229-10-86

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

Cheng, X., Xin, M., Xu, R., Chen, Z., Cai, W., Chai, L., et al. (2020). A single amino acid substitution in STKc_GSK3 kinase conferring semispherical grains and its implications for the origin of triticum sphaerococcum. *Plant Cell* 32, 923–934. doi: 10.1105/tpc.19.00580

Cheng, Z. J., Zhao, X. Y., Shao, X. X., Wang, F., Zhou, C., Liu, Y. G., et al. (2014). Abscisic acid regulates early seed development in arabidopsis by ABI5-mediated transcription of SHORT HYPOCOTYL UNDER BLUE1. *Plant Cell* 26, 1053–1685. doi: 10.1105/tpc.113.121566

Cho, H., Ryu, H., Rho, S., Hill, K., Smith, S., Audenaert, D., et al. (2014). A secreted peptide acts on BIN2-mediated phosphorylation of ARFs to potentiate auxin response during lateral root development. *Nat. Cell Biol.* 16, 66–76. doi: 10.1038/ncb2893

Chono, M., Honda, I., Zeniya, H., Yoneyama, K., Saisho, D., Takeda, K., et al. (2003). A semidwarf phenotype of barley uzu results from a nucleotide substitution in the gene encoding a putative brassinosteroid receptor. *Plant Physiol.* 133, 1209–1195. doi: 10.1104/pp.103.026195

Cingolani, P., Platts, A., Wang, Le L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, snpEff: SNPs in the genome of drosophila melanogaster strain W1118; iso-2; iso-3. *Fly* 6, 80–925. doi: 10.4161/fly.19695

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* 10, giab008. doi: 10.1093/gigascience/giab008

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA). (2010). Scientific opinion on the substantiation of a health claim related to oat beta glucan and lowering blood cholesterol and reduced risk of (Coronary) heart disease pursuant to article 14 of regulation (EC) no 1924/2006. *EFSA J.* 8(12), 1885. doi: 10.2903/j.efsa.2010.1885

Gao, X., Zhang, J.-Q., Zhang, X., Zhou, J., Jiang, Z., Huang, P., et al. (2019). Rice qGL3/osPPKL1 functions with the GSK3/SHAGGY-like kinase osGSK3 to modulate brassinosteroid signaling. *Plant Cell* 31, 1077–1093. doi: 10.1105/tpc.18.00836

Gruszka, D. (2020). Exploring the brassinosteroid signaling in monocots reveals novel components of the pathway and implications for plant breeding. *Int. J. Mol. Sci.* 21, 354. doi: 10.3390/ijms21010354

Horie, Y., Ito, H., Kusaba, M., Tanaka, R., and Tanaka, A. (2009). Participation of Chlorophyll b Reductase in the Initial Step of the Degradation of Light-harvesting Chlorophyll a/b-Protein Complexes in Arabidopsis. *J. Biol. Chem.* 284, 17449–17565. doi: 10.1074/jbc.M109.008912

Hou, L., Li, Z., Shaheen, A., Zhang, K., Wang, J., Gao, X., et al. (2022). Zea mays GSK2 gene is involved in brassinosteroid signaling. *Plant Growth Regul* 97, 117–125. doi: 10.1007/s10725-022-00806-z

Hu, Z., Lu, S.-J., Wang, M.-J., He, H., Sun, L., Wang, H., et al. (2018). A novel QTL qTGW3 encodes the GSK3/SHAGGY-like kinase osGSK5/osSK41 that interacts with osARF4 to negatively regulate grain size and weight in rice. *Mol. Plant* 11, 736–749. doi: 10.1016/j.molp.2018.03.005

International Oat Nomenclature Committee. (2021). Meeting Notes 30 April 2021.

Ishimaru, K., Hirotsu, N., Madoka, Y., Murakami, N., Hara, N., Onodera, H., et al. (2013). Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. *Nat. Genet.* 45, 707–711. doi: 10.1038/ng.2612

Jia, M., Li, Y., Wang, Z., Tao, S., Sun, G., Kong, X., et al. (2021). TaIAA21 represses taARF25-mediated expression of taERFs required for grain size and weight development in wheat. *Plant J.* 108, 1754–1767. doi: 10.1111/tpj.15541

Jiang, W.-B., Huang, H.-Y., Hu, Y.-W., Zhu, S.-W., Wang, Z.-Y., and Lin, W.-H. (2013). Brassinosteroid regulates seed size and shape in arabidopsis. *Plant Physiol.* 162, 1965–1775. doi: 10.1104/pp.113.217703

Kabir, M. R., and Nonhebel, H. M. (2021). Reinvestigation of THOUSAND-GRAIN WEIGHT 6 grain weight genes in wheat and rice indicates a role in pollen development rather than regulation of auxin content in grains. *Theor. Appl. Genet.* 134, 2051–2625. doi: 10.1007/s00122-021-03804-3

Kamal, N., Tsardakas Renhuldt, N., Bentzer, J., Gundlach, H., Haberer, G., Juhász, A., et al. (2022). The mosaic oat genome gives insights into a uniquely healthy cereal crop. *Nature* 606, 113–119. doi: 10.1038/s41586-022-04732-y

Kassaw, T., Nowak, S., Schnabel, E., and Frugoli, J. (2017). ROOT DETERMINED NODULATION1 is required for M. Truncatula CLE12, but not CLE13, peptide signaling through the SUNN receptor kinase. *Plant Physiol.* 174, 2445–2565. doi: 10.1104/pp.17.00278

Kloc, Y., Dmochowska-Boguta, M., Zielezinski, A., Nadolska-Orczyk, A., Karlowski, W. M., and Orczyk, W. (2020). Silencing of hvGSK1.1/SHAGGY-like kinase (Hordeum vulgare L.) growth in normal and in salt stress conditions. *Int. J. Mol. Sci.* 21. doi: 10.3390/ijms21186616

Kusaba, M., Ito, H., Morita, R., Iida, S., Sato, Y., Fujimoto, M., et al. (2007). Rice NON-YELLOW COLORING1 is involved in light-harvesting complex II and grana degradation during leaf senescence. *Plant Cell* 19, 1362–1375. doi: 10.1105/tpc.106.042911

Letunic, I., and Bork, P. (2021). Interactive tree of life (iTOL) V5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301

Li, C., Zhang, B., and Yu, H. (2021). GSK3s: nodes of multilayer regulation of plant development and stress responses. *Trends Plant Sci.* 26, 1286–13005. doi: 10.1016/j.tplants.2021.07.017

Li, H., Luo, L., Wang, Y., Zhang, J., and Huang, Y. (2022). Genome-wide characterization and phylogenetic analysis of GSK genes in maize and elucidation of their general role in interaction with BZR1. *Int. J. Mol. Sci.* 23, 80565. doi: 10.3390/ijms23158056

Li, J., Nie, X., Tan, J. L. H., and Berger, F. (2013). Integration of epigenetic and genetic controls of seed size by cytokinin in arabidopsis. *Proc. Natl. Acad. Sci.* 110, 15479–15845. doi: 10.1073/pnas.1305175110

Li, N., Xu, R., and Li, Y. (2019). Molecular networks of seed size control in plants. *Annu. Rev. Plant Biol.* 70, 435–635. doi: 10.1146/annurev-arplant-050718-095851

Li, P., Chen, Y.-H., Lu, J., Zhang, C.-Q., Liu, Q.-Q., and Li, Q.-F. (2022). Genes and their molecular functions determining seed structure, components, and quality of rice. *Rice* 15, 185. doi: 10.1186/s12284-022-00562-8

Lin, M. F., Rodeh, O., Penn, J., Bai, X., Reid, J. G., Krasheninina, O., et al. (2018). GLnexus: joint variant calling for large cohort sequencing. *bioRxiv*, 343970. doi: 10.1101/343970

Liu, K., and Wise, M. L. (2021). Distributions of nutrients and avenanthramides within oat grain and effects on pearled kernel composition. *Food Chem.* 336, 127668. doi: 10.1016/j.foodchem.2020.127668

Liu, Z., and Meyerowitz, E. M. (1995). LEUNIG regulates AGAMOUS expression in arabidopsis flowers. *Development* 121, 975–991. doi: 10.1242/dev.121.4.975

Mariotti, F., Tomé, D., and Mirand, P. P. (2008). Converting nitrogen into protein 6.25 and jones' Factors. *Crit. Rev. Food Sci. Nutr.* 48, 177–845. doi: 10.1080/10408390701279749

Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., Tomkins-Tinch, C. H., Sochat, V., et al. (2021). Sustainable data analysis with snakemake. *F1000Research* 10, 33. doi: 10.12688/f1000research.29032.2

Nonhebel, H. M., Griffin, K., Nonhebel, H. M., and Griffin, K. (2020). Production and roles of IAA and ABA during development of superior and inferior rice grains. *Funct. Plant Biol.* 47, 716–265. doi: 10.1071/FP19291

Peng, P., and Li, J. (2003). Brassinosteroid signal transduction: A mix of conservation and novelty. *J. Plant Growth Regul.* 22, 298–3125. doi: 10.1007/s00344-003-0059-y

Pérez-Pérez, J. M., Ponce, M. R., and Micol, J. L. (2002). The UCU1 arabidopsis gene encodes a SHAGGY/GSK3-like kinase required for cell expansion along the proximodistal axis. *Dev. Biol.* 242, 161–735. doi: 10.1006/dbio.2001.0543

Poplin, R., Chang, P.-C., Alexander, D., Schwartz, S., Colthurst, T., Ku, A., et al. (2018). A universal SNP and small-indel variant caller using deep neural networks. *Nat. Biotechnol.* 36, 983–987. doi: 10.1038/nbt.4235

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS One* 5, e94905. doi: 10.1371/journal.pone.0009490

Ramachandran, V., and Chen, X. (2008). Degradation of microRNAs by a family of exoribonucleases in arabidopsis. *Science* 321, 1490–1492. doi: 10.1126/science.1163728

R Core Team (2023). *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing). Manual.

Schnabel, E. L., Kassaw, T. K., Smith, L. S., Marsh, J. F., Oldroyd, G. E., Long, S. R., et al. (2011). The ROOT DETERMINED NODULATION1 gene regulates nodule number in roots of medicago truncatula and defines a highly conserved, uncharacterized plant gene family. *Plant Physiol.* 157, 328–405. doi: 10.1104/pp.111.178756

Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A. H., Nielsen, K. L., et al. (2009). SHOREmap: simultaneous mapping and mutation identification by deep sequencing. *Nat. Methods* 6, 550–515. doi: 10.1038/nmeth0809-550

Schruff, M. C., Spielman, M., Tiwari, S., Adams, S., Fenby, N., and Scott, R. J. (2006). The AUXIN RESPONSE FACTOR 2 gene of arabidopsis links auxin signalling, cell division, and the size of seeds and other organs. *Development* 133, 251–615. doi: 10.1242/dev.02194

Sjoberg, D. D., Whiting, K., Curry, M., Lavery, J. A., and Larmarange, J. (2021). Reproducible summary tables with the gtsummary package. *R J.* 13, 570–805. doi: 10.32614/RJ-2021-053

Smith, S. M., Li, C., and Li, J. (2017). "1 - hormone function in plants," in *Hormone metabolism and signaling in plants*. Eds. J. Li, C. Li and S. M. Smith (Cambridge, Massachusetts: Academic Press), 1–38. doi: 10.1016/B978-0-12-811562-6.00001-3

Song, Y., Wang, Y., Yu, Q., Sun, Y., Zhang, J., Zhan, J., et al. (2023). Regulatory network of GSK3-like kinases and their role in plant stress response. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1123436

Sun, F., Ding, L., Feng, W., Cao, Y., Lu, F., Yang, Q., et al. (2021). Maize transcription factor zmBES1/BZR1-5 positively regulates kernel size. *J. Exp. Bot.* 72, 1714–1726. doi: 10.1093/jxb/eraa544

Tanabata, T., Shibaya, T., Hori, K., Ebana, K., and Yano, M. (2012). SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis. *Plant Physiol.* 160, 1871–1880. doi: 10.1104/pp.112.205120

Tange, O. (2021). *GNU parallel 20210422 ('Ever given')*. doi: 10.5281/ZENODO.4710607

Tinker, N. A., Wight, C. P., Bekele, W. A., Yan, W., Jellen, E. N., Tsardakas Renhuldt, N., et al. (2022). Genome analysis in avena sativa reveals hidden breeding barriers and opportunities for oat improvement. *Commun. Biol.* 5, 1–115. doi: 10.1038/s42003-022-03256-5

Tong, H., Liu, L., Jin, Y., Du, L., Yin, Y., Qian, Q., et al. (2012). DWARF AND LOW-TILLERING acts as a direct downstream target of a GSK3/SHAGGY-like kinase to mediate brassinosteroid responses in rice. *Plant Cell* 24, 2562–2577. doi: 10.1105/tpc.112.097394

Tsardakas Renhuldt, N. (2021). *Nikostr/dna-seq-deepvariant-glnexus-variant-calling: V0.3.1*. doi: 10.5281/ZENODO.5045703

Vasimuddin, Md., Misra, S., Li, H., and Aluru, S. (2019). "Efficient architecture-aware acceleration of BWA-MEM for multicore systems," in *2019 IEEE international parallel and distributed processing symposium (IPDPS)*, 314–324. doi: 10.1109/IPDPS.2019.00041

Wada, K., Kondo, H., and Marumo, S. (1985). A simple bioassay for brassinosteroids: A wheat leaf-unrolling test. *Agric. Biol. Chem.* 49, 2249–2251. doi: 10.1271/bbb1961.49.2249

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., et al. (2019). Welcome to the tidyverse. *J. Open Source Software* 4, 1686. doi: 10.21105/joss.01686

Yan, W., Frégeau-Reid, J., and Mitchell Fetch, J. (2013). "Breeding for ideal milling oat: challenges and strategies," in *Oats nutrition and technology* (New York: John Wiley & Sons, Ltd), 7–32. doi: 10.1002/9781118354100.ch2

Yang, C., Liu, X., Li, D., Zhu, X., Wei, Z., Feng, Z., et al. (2019). OsLUGL is involved in the regulating auxin level and osARFs expression in rice (Oryza sativa L.). *Plant Sci.* 288, 110239. doi: 10.1016/j.plantsci.2019.110239

Yang, Y., Liu, X., Zhang, W., Qian, Q., Zhou, L., Liu, S., et al. (2021). Stress response proteins NRP1 and NRP2 are pro-survival factors that inhibit cell death during ER stress. *Plant Physiol.* 187, 1414–1275. doi: 10.1093/plphys/kiab335

Youn, J.-H., and Kim, T.-W. (2015). Functional insights of plant GSK3-like kinases: multi-taskers in diverse cellular signal transduction pathways. *Mol. Plant* 8, 552–565. doi: 10.1016/j.molp.2014.12.006. Cell Signaling.

Yun, T., Li, H., Chang, P.-C., Lin, M. F., Carroll, A., and McLean, C. Y. (2021). Accurate, scalable cohort variant calls using deepVariant and GLnexus." Edited by peter robinson. *Bioinformatics* 36(24), 5582–5895. doi: 10.1093/bioinformatics/btaa1081

Zhang, K., Dong, R., Hu, X., Ren, C., and Li, Y. (2021). Oat-based foods: chemical constituents, glycemic index, and the effect of processing. *Foods* 10, 13045. doi: 10.3390/foods10061304

Zhang, P., Zhang, L., Chen, T., Jing, F., Liu, Y., Ma, J., et al. (2022). Genome-wide identification and expression analysis of the GSK gene family in wheat (Triticum aestivum L.). *Mol. Biol. Rep.* 49, 2899–29135. doi: 10.1007/s11033-021-07105-2

Zimmer, C. M., McNish, I. G., Klos, K. E., Eickholt, D. P., Arruda, K. M.A., Pacheco, M. T., et al. (2021). Genome-wide association mapping for kernel shape and its association with β-glucan content in oats. *Crop Sci.* 61, 3986–3995. doi: 10.1002/csc2.20605

Zuo, J., and Li, J. (2014). Molecular genetic dissection of quantitative trait loci regulating rice grain size. *Annu. Rev. Genet.* 48, 99–118. doi: 10.1146/annurev-genet-120213-092138