



Genetic Improvement of Sorghum: Crop Genome Designing for Nutraceuticals

M. T. Labuschagne and L. Elkonin

Contents

1	Introduction	196
2	Sorghum Grain Chemical Composition	198
2.1	Phenolic Compounds	198
2.2	Carotenoids	202
2.3	Lipids and Vitamin E	203
2.4	Amines	203
2.5	Carbohydrates	203
2.6	Fibers, Vitamins, and Minerals	204
2.7	Sorghum Protein	204
2.8	Other Proteins in Sorghum	205
3	Genetic Improvement of Sorghum Nutraceutical Content	206
3.1	Genetic Resources for Genetic Improvement of Nutraceuticals and Nutritional Value	206
3.2	Subset Collections as Sources for Marker-Assisted Breeding	206
3.3	Sorghum Linkage and Association Mapping Resources	207
3.4	The Sorghum Mutant Library	208
3.5	Genetic Engineering Approaches for Improving Nutritional Composition of Sorghum Grain	209
4	Conclusion	219
	References	220

Abstract

Sorghum is an ancient cereal crop grown widely in the dry regions of Africa and Asia, mainly by subsistence farmers. It is inherently adapted to drought and heat stress and has significant potential as a global food security crop in changing

M. T. Labuschagne (✉)

Department of Plant Sciences, University of the Free State, Bloemfontein, South Africa

e-mail: labuscm@ufs.ac.za

L. Elkonin

Department of Biotechnology, Federal Centre of Agriculture Research of the South-East Region, Saratov, Russia

climatic conditions. The nutritional profile of sorghum is comparable with that of other cereals, but is unique in that it has various bioactive compounds such as phenolic acids, procyanidins, flavonoids, and anthocyanins. The protein (kafirin) of sorghum is, however, poorly digestible, and genetic improvement of protein digestibility has remained a challenge. Large genetic variability has been shown for almost all the bioactive compounds in tested sorghum accessions, but there is a large gap in the knowledge of the genetics underlying their expression. There are a number of sorghum germplasm collections, reference genome sequences, association panels, and mutant populations available in the world, which could be used to screen for genetic variation and determine genetic architecture, in order to improve sorghum nutraceutical content. Genetic engineering techniques have been applied to improve kafirin digestibility and carotenoid content, and have significant potential for improving nutraceutical content in the future. Information on the genes linked to nutraceuticals and the gene action involved in their expression will allow improvement of nutraceutical content through conventional breeding and by use of molecular markers, genomics, and genetic engineering. All available information on grain nutraceuticals and other nutritional components in sorghum should be integrated as a resource to be used by the sorghum community for sorghum improvement and genetic studies of nutraceuticals.

Keywords

Sorghum · Phenolic compounds · Kafirin · Genetic resources · Genomics · Genetic engineering

1 Introduction

Sorghum [*Sorghum bicolor* (L.) Moench] is a stress-tolerant C4 plant, which belongs to the *Poaceae* family, which is well adapted to arid growing conditions. It is a versatile ancient cereal grown mainly by subsistence farmers. This annual crop is mostly photoperiod insensitive and completes its life cycle in 4 months. It is seen as a potential food security crop highly suited to climate change conditions. Maize, rice, and wheat are the most important cereals in the world, with sorghum ranking fifth in terms of importance. It is a staple food to more than 500 million people in Africa and Asia (Xin et al. 2021). The annual sorghum production is over 57 million tons (FAOSTAT 2019). In developed countries, it is mainly used for animal feed, biofuel production, and bioproducts, but it is a staple food in many developing countries (Xin et al. 2021).

Nutraceuticals are, by definition, bioactive compounds, which are biologically active, and are found in food. Various phytochemical substances are regarded as bioactive compounds, including the polyphenols. Dietary fiber, polyunsaturated fatty acids, micro- and macronutrients, vitamins, oligosaccharides, lactic acid bacteria, choline, and lecithin are all bioactive compounds (Chaudhari et al. 2017). Biopeptides, which are biologically active peptides, are also nutraceuticals.

Nutraceuticals have possible antioxidant properties and various other effects on the human body. Protein is a natural source of biopeptides, including antioxidant biopeptides (Hu et al. 2021). Certain amino acids, peptides, and proteins are also possible bioactive compounds (Szerszunowicz and Kłobukowski 2020). Reactive oxygen species (ROS) is important in metabolism and aging of humans, and some nutraceuticals contribute to decreasing the levels of ROS formed in the body through antioxidant activity (Li et al. 2021a). Human consumption of foods containing nutraceuticals have many therapeutic advantages. These compounds contribute to longevity, delaying aging, and preventing the development chronic diseases such as diabetes and hypertension, among others (Szerszunowicz and Kłobukowski 2020). Nutraceuticals can contribute to the prevention and management of viral infections (Li et al. 2021a). Different kinds of biological properties and activities have been ascribed to bioactive compounds, including antioxidant, anti-inflammatory, and antimicrobial properties, which can contribute to protection against human disease (Omrani et al. 2020).

Sorghum grains consist of carbohydrates, kafirin (protein), fiber, polyunsaturated fatty acids, and resistant starch (Khalid et al. 2022). Sorghum has a nutritional profile comparable with other cereals, but is unique to other cereals in that it contains different bioactive compounds such as anthocyanins, flavonoids, phenolic acids, and pro-cyanidins (Ofosu et al. 2021). The extensive biological activities of sorghum grains have been shown in a number of *in vitro* and *in vivo* studies. Sorghum bioactive compounds have been shown to inhibit oxidative stress, cardiovascular disease, high lipid levels, and hypertension. It has anticancer and antidiabetic properties and could lower cholesterol index and obesity through antioxidant and anti-inflammatory mechanisms (Li et al. 2021a). The phenolic compounds having bioactive properties include phenolic acid, flavonoids, stilbenes, and tannins. Sorghum grain also contains B-complex vitamins, fat-soluble vitamins A, D, E, and K, as well as minerals, including magnesium, potassium, phosphorus, and zinc. Some sorghum polyphenols, including tannin (proanthocyanidin) and 3-deoxyanthocyanidin, have the potential to protect individuals against inflammation, diabetes, and oxidative stress. Flavonoids are also known as antioxidants, which could positively influence inflammatory and neurodegenerative diseases, as well as cancer and diabetes. Sorghum grain also have a high fiber content, which can reduce blood cholesterol and glucose levels (Khalid et al. 2022). As sorghum is a good source of B vitamins, minerals, carbohydrates, and is also gluten free, it has significant potential as a source of food and beverages and gluten-containing replacement diets for people with celiac disease (Bouargalme et al. 2022).

Although sorghum is known for various phytochemicals, which contribute to human health, it is often not valued for these attributes due to the poor digestibility of its proteins (Duressa et al. 2018). Extensive research has been done by various research groups on sorghum protein digestibility. There have been sustained (although limited) global research efforts focusing on unraveling the biochemical and genetic basis of low protein digestibility and the improvement of kafirin bioavailability. The sequencing of the sorghum genome has supported these efforts, leading to a better scientific understanding of kafirin chemical properties, the

genetics underlying kafirins, and its protein body structure. Despite this, there have been several challenges in the genetic improvement of protein digestibility, as increased protein digestibility is linked to agronomically undesirable traits such as the opaque (soft) endosperm phenotype. It is technically possible to develop high-protein digestible sorghum with hard endosperm, as was the case with quality protein maize, especially with the help of marker-assisted breeding (Duressa et al. 2018). Li et al. (2021a) stated that kafirins are some of the best nutraceutical sources.

Significant genetic variability has been reported for almost all the bioactive compounds in numerous sorghum accessions, indicating the possibility for genetic improvement through selection. Despite this, very limited research has been done on the genetic basis of bioactive compounds in sorghum and its possible improvement through conventional breeding, the use of DNA markers, genomics, or genetic engineering. To improve sorghum, it is critical to use the available genetic variation. These genetic resources should be screened for valuable genes underlying the synthesis of nutraceutical compounds (Bouargalne et al. 2022).

In the world, there are a number of sorghum germplasm collections, reference genome sequences of good quality, and association panels, which can be used for genome-wide association studies for food and quality related traits. There are also mutant populations, which are useful to discover genes, which can be applied for sorghum improvement, as well as information on gene expression. Genetic engineering is becoming increasingly important in sorghum nutritional content research, and the technology has developed significantly in recent years.

The aim of this chapter was to determine the current status of sorghum as a source of nutraceuticals, to assess current genetic resources in sorghum for possible genetic manipulation and improvement of nutraceutical content, and the use of new technologies such as genomics and genetic engineering to improve sorghum as a source of nutraceuticals in the human diet.

2 Sorghum Grain Chemical Composition

Sorghum grain composition is similar to that of maize and millet. It consists of starch, lipids, protein, as well as non-starch polysaccharides. It also contains B vitamins and vitamins D, E, and K (fat-soluble), as well as minerals (Przybylska et al. 2019). Sorghum seed was reported to consist of protein (4.4–21.1%), fat (2.1–7.6%), crude fiber (1.0–3.4%), total carbohydrate (57.0–80.6%), starch (55.6–75.2%), and total minerals as ash (1.3–3.5%) (Cabrera et al. 2020). In contrast to other cereals, sorghum is also a rich source of phenolic compounds.

2.1 Phenolic Compounds

Sorghum has the highest amount of phenolic compounds of all cereals. A variety of phenolic compounds are present in sorghum grains, including phenolic acids, flavonols, 3-deoxyanthocyanidins, flavanones, flavones, and condensed tannins

(Shen et al. 2018). The outer layer of the grain (bran) has a high concentration of most of these compounds, which is unfortunately often removed during milling by decortication, which leads to significant reduction of sorghum health benefits. Therefore, the consumption of whole grain sorghum has many key health benefits, such as free radical scavenging activity (Kumari et al. 2021).

Two main categories of phenolic compounds have been identified, namely soluble and insoluble. The soluble compounds like flavonoids, quinones, and phenylpropanoids are found in the vacuole of plant cells, while the insoluble compounds like lignins, condensed tannins, and hydroxycinnamic acid are attached to the cell wall (Gharaati 2019). The phenolic compounds are important secondary metabolites and are major contributors to the antioxidant properties of sorghum, having significant physiological benefits for humans. For this reason, the consumption of sorghum-based foods with high levels of polyphenolic substances can contribute to the prevention and reduction of the risk of chronic diseases, including some cancers and diabetes. The sorghum variety or genotype, the grain pericarp color, and the testa pigmentation all influence the phenolic compound profiles in sorghum. There are four classes of pericarp color in sorghum, which are white, yellow, red, and black. Genetically, black sorghums are actually red, as sunlight during the maturation process causes the red color to turn black. The white sorghums, also known as food-type sorghum, have very low levels of total extractable phenol and usually have no tannins or anthocyanins. Red sorghum is caused by a red pericarp and contains significant levels of extractable phenols, although they have no tannins. Black pericarps lead to black sorghums, which have very high levels of anthocyanins, while the brown sorghums have varying levels of phenolic compounds and pigmented pericarps (Khalid et al. 2022).

The presence of total phenolic compounds in most sorghum whole grain is 0.46 ~ 20 mg gallic acid equivalent (GAE) per gram. The highest total phenolic compound content of almost 48 mg GAE/g was reported in whole grain red sorghum. Total phenolic content in sorghum bran varies even more than in seed. The total phenolic content of red sorghum bran was 20-fold that of white sorghum bran and 8.9-fold that of yellow sorghum bran, while that in black sorghum bran extracts was 7.5-fold that of white sorghum bran and 3.3-fold that of yellow sorghum bran extract (Burdette et al. 2010). It must be kept in mind that the variations in the quantities of total phenolic compounds reported in different studies could have been influenced by different extraction procedures and the solvents used, the specific genotypes tested, and environmental conditions in which plants were grown (Li et al. 2021a).

Free and bound forms of phenolic compounds are found in sorghum, but 70–95% of phenolic acids are in a bound form. Despite this, far more research has been done on the identification and biological activity of free phenolic compounds in sorghum than on the bound compounds (Li et al. 2021b). Bound phenolics bind to structural compounds of the cell wall, thereby reducing bioavailability. It is therefore important to find novel ways to promote the release of bound phenolic compounds in order to increase their bioavailability. The molecular bonding can be severed by creating an acidic or basic environment and by elevating the temperature, which can cause release (Li et al. 2021a; Espitia-Hernández et al. 2022).

Phenolic compounds can act as natural antioxidants by decreasing the oxidative damage of biomolecules, which reduces the effects of reactive oxidants. Phenolic extracts from sorghum were shown to decrease and inhibit the growth of cancer cells in organs such as the colon, liver, and esophagus (Kadri et al. 2017). Phenolic compounds vary significantly between sorghum seeds of different colors (black, brown, red, and white). The amount of phenolic compounds is highly correlated with antioxidant properties. Due to their phenolic compound content, sorghum grains are important as an ingestible form of antioxidants in the diet. The development of sorghum varieties with high levels of antioxidants can increase its nutritional value and health benefits. Sorghum tannins also have medicinal properties (Choi et al. 2019).

2.1.1 Phenolic Acids and Flavonoids

Various studies have reported different types and numbers of phenolic acids, but the most frequently reported phenolic acids isolated from sorghum were caffeic acid and ferulic acid, 3-deoxyanthocyanidins (luteolinidin and apigeninidin), flavanones (naringenin), flavones (luteolin and apigenin), and dihydroflavonol (taxifolin) (Wu et al. 2017). Caffeic acid, p-coumaric acid, sinapic acid, gallic acid, protocatechuic acid, and p-hydroxybenzoic acid have also been reported quite frequently. Ferulic acid is the predominant phenolic acid, especially in red and brown sorghum, and is uncountable for about 90% of the combined phenolic acids. In red sorghum, ferulic acid was the predominant phenolic acid, and p-coumaric, caffeic, and 3,4-dihydroxybenzoic acids were also identified, but there was a large genotype influence on phenolic acids (Li et al. 2021b). Red and brown sorghum grains had the most luteolinidin and apigeninidin, followed by black grains, while white pericarp varieties had very low amounts of these compounds. Many phenolic acids have been assayed, and of these, ferulic, p-coumaric, and protocatechuic acids were seen in the highest concentrations in both red and white sorghum grain (Przybylska-Balcerek et al. 2018).

The outer layers of the grain contain most of the flavonoids, which contribute to the coloring of the pericarp. Sorghum grains have been found to contain many flavonoids. As in the case of phenolic acids, a large genotype effect was seen (Li et al. 2021b). Both the color and thickness of the pericarp influence flavonoid concentrations and profiles. Flavonoids have antioxidant properties, and the daily consumption of foods with significant amounts of flavonoids can help to reduce the risk of cancers of the breast, colon, and pancreas. The main flavonoids in sorghum are anthocyanins, which are a group of anthocyanidin glycosides (Wu et al. 2017). Anthocyanin content in sorghum bran was three to four times higher than that of the flour. The highest anthocyanin content was reported in black sorghum bran (Kumari et al. 2021).

The most frequently reported flavones in sorghum grains are luteolin, apigenin, and naringenin. The widest known flavonols in sorghum grain are kaempferol and quercetin, and the most researched flavanol is iscatechin, and for dihydroflavonol, it is taxifolin (Luo et al. 2020). Ofosu et al. (2021) were the first to report the presence of formononetin, glycitein, and ononin in sorghum. High concentrations of flavanones were identified in yellow-pigmented sorghum genotypes, and those with colored testas had a higher content of condensed tannins (Szczeszunowicz and

Kłobukowski 2020). A high concentration of flavones was reported in red-brown sorghum bran (576.47 µg/g) and lower amounts in yellow sorghum flour (15.3 µg/g). The flavanone concentrations were higher in yellow bran (1773.47 µg/g) and lower in brown sorghum flour (4.29 µg/g). Flavanone concentrations were consistently higher in bran than in flour (Kumari et al. 2021). Some of the sorghum flavones had estrogenic effects, as well as anticancer effects *in vitro* (Cox et al. 2019).

Of all the cereals, only sorghum is a dietary source of 3-deoxyanthocyanidins (3-DXAs). These occur mainly as luteolinidin and apigeninidin, which are mostly water-soluble pigments (Luo et al. 2020). They produce yellow (apigeninidin) and orange (luteolinidin) colors in acidic solvents. They are effective natural colorants and have good antioxidant properties, which are beneficial for human health. Li et al. (2021a) reported that luteolinidin was the predominant 3-DXA, with its total content accounting for 40.55–78.36% of the total 3-DX. The concentration of 3-DXA was three to four times higher in black testa seeds than red or brown testa seeds (Shen et al. 2018). Kumari et al. (2021) reported that red-brown sorghum bran had the highest 3-DXA (4479.16 µg/g), while yellow sorghum flour had low 3-DXA levels (14.14 µg/g). Results from *in vitro* tests in sorghum showed that 3-DXA had both anticancer and antioxidant properties (Cox et al. 2019). The main antidiabetic substances in sorghum flavonoids have likewise been ascribed to condensed tannins and 3-DXAs (Li et al. 2021a).

2.1.2 Stilbenoids

Stilbenoids are a group of phenolic compounds, and they have a number of benefits for human health. Sorghum can produce stilbenoid metabolites, but very limited research has been done on this. The total stilbenoid content is related to grain color. One study reported 0.4–1 mg/kg of trans-piceid in white sorghum and up to 0.2 mg/kg trans-resveratrol in red sorghum grains (Bröhan et al. 2011).

2.1.3 Tannins

The growing knowledge on the health benefits of tannins have led to tannins receiving more attention in the last years. They are heterogeneous polyphenolic polymers. Tannins in sorghum occur mainly in the pericarp and are polymerized products of flavan-3, 4-diols, and/or flavan-3-ols. Tannin content in sorghum bran is higher than in the flour. Bran from brown and red-brown sorghum varieties had significantly higher tannin content than bran of red and yellow varieties (Kumari et al. 2021). Tannins have many health benefits as they have antioxidant properties and are radical scavengers. They were also reported to improve immunity and have anticancer and anti-inflammatory properties. They are cardioprotective, are vasodilators and have antithrombotic effects (Queiroz et al. 2018). Sorghum is an excellent source of tannins compared to other cereals. Seyoum et al. (2016) reported tannin concentration ranging from 0.2 to 48.0 mg/g in sorghum, with grain with a black testa having the highest tannin level. A significant growing season effect was evident on the tannin content and activity in sorghum. Sorghum with a high tannin content generally has good resistance to birds, insects, and molds.

Tannins, unfortunately, also have anti-nutritional effects, as they form complexes with protein and iron. This impairs the digestibility of the protein and reduces iron absorption (Iyabo et al. 2018). Apart from their interaction with protein, tannins also affect carbohydrates, especially starch, hemicellulose, cellulose, and pectin, reducing their digestibility. Based on their biological and chemical characteristics, tannins have been grouped into hydrolyzable and condensed (known as proanthocyanidins) tannins. Hydrolyzable tannins are complex polymeric compounds classified as gallotannins, when derived from gallic acid, and ellagitannins, when derived from ellagic acid, which is a dimer of gallic acid. Many and varied hydrolyzable tannins have been isolated from edible and inedible plants. They have anticancer, anti-diabetic, and antibiotic effects. Condensed tannins are the main polyphenolic substances in sorghum (De Oliveira et al. 2017).

Condensed tannins are especially common in sorghum seeds with pigmented testas. The testa is a structure present between the pericarp and the endosperm of the grain of only some varieties. Sorghum varieties with a brown testa generally have more procyanidins than that of other seed colors. The presence of condensed tannins is one of the reasons why sorghum has higher levels of antioxidants than any other cereal (De Oliveira et al. 2017). Proanthocyanidins in sorghum grains have antioxidant, antitumor, and lipid-lowering activities, and they play a role in the prevention of cardiovascular diseases. They are powerful radical scavengers (Yu et al. 2018). The most basic unit of composition of condensed tannins is catechin, gallocatechin, allocatechin gallate, and afzelechin (Jiang et al. 2020). Testa color is used as a means to classify sorghum grains into three categories where type I is tannin-free sorghum without pigmented testa, type II sorghum has a pigmented testa layer (condensed tannins), and type III contain tannins in the testa and pericarp. The proanthocyanidins in sorghum form complexes and precipitate with proteins, probably causing both acidity and a bitter taste, which contribute to them being bird repellents (Li et al. 2021a). Procyanidins concentrations of between 10.6 and 40.0 mg/g have been reported in sorghum, depending on the genotype. This is higher than that of blueberries, known for their high procyanidin content (Yu et al. 2018). Sorghum grain extracts rich in tannin polyphenolics were shown to have *in vitro* and *in vivo* inhibitory effects against α -amylase and α -glucosidase enzymes. This could decrease hyperglycemia in diabetics (Links et al. 2015).

2.2 Carotenoids

Carotenoids have many beneficial effects on human health. The most researched carotenoids in sorghum include lutein and zeaxanthin, which are the xanthophylls, which are also the major carotenoids in sorghum, and β -carotene. Varied concentrations of carotenoids have been reported in different studies. This may be due to a genotype effect and different extraction and detection methods. In one study, the total carotenoid content in a hundred sorghum genotypes varied from 2.12 to 85.46 $\mu\text{g}/100\text{ g}$. Nine genes were reported to be involved in carotenoid synthesis or degradation, which is a contributing factor to the variability of levels of carotenoids in the grains of sorghum (Cardoso et al. 2015).

2.3 Lipids and Vitamin E

Sorghum grain has low lipid content, though higher than most other cereal grains, and its lipid profile is similar to that of maize. The oil from sorghum grain have higher levels of oleic and stearic acids and lower levels of linoleic, myristic, and palmitoleic acids, making it less saturated than maize grain oil. The fact that oil from sorghum grain is high in unsaturated fatty acids may provide the health benefit of lipid-lowering properties (Khalid et al. 2022).

Sorghum grain fatty acids are considered as bioactive compounds with health benefits, especially the phytosterols and policosanols. These fatty acids occur in the lipid fraction of the grain as the lipid group of triacylglycerols with a high linoleic, oleic, and palmitic fatty acid content. Phytosterols are steroids that originate from plants, and β -sitosterol is the main phytosterol that was found in sorghum. Campesterol and stigmasterol were also isolated. The concentration of phytosterols is largely influenced by the environment, genotype, and extraction methods. Policosanols (a class of high-molecular-weight aliphatic alcohols) also have different types of bioactivity. Different policosanols (C28, C30, and C32) were isolated from sorghum, of which C28 policosanol was the predominant one (Wongwaiwech et al. 2020).

The vitamin E contents in sorghum were reported to vary significantly, between 280.7 and 2962.4 $\mu\text{g}/100\text{ g}$ (wet basis). The α -, β -, γ -, and δ -tocopherols are the most studied in sorghum, with γ -tocopherol being the tocopherol most frequently found, followed by α -tocopherol. The genotype and the growing environment both have a significant influence on the vitamin E content in sorghum grains (Cardoso et al. 2015).

2.4 Amines

There are two classes of amines, which are the biogenic amines and polyamines, and they represent a class of low-molecular-mass nitrogenous bases. The polyamines represent 60–100% of the total amines. Amines are considered to be bioactive compounds, and sorghum was reported to be a main source of polyamines. Spermine and spermidine are the most prevalent amines, followed by putrescine and cadaverine (Paiva et al. 2015).

2.5 Carbohydrates

The major carbohydrate in sorghum grain is starch, present in a granular form in the endosperm. Starch consists of two polysaccharides, amylose and amylopectin, and can be classified as waxy and non-waxy. Sorghum starch classified as non-waxy is composed of amylose (25%) and amylopectin (75%), while waxy sorghum consists almost entirely of amylopectin. Waxy sorghum is highly digestible by enzymes, while non-waxy starch exhibits resistance to enzymatic digestion. Proteins, tannins, and starch granules interact in the sorghum grains, forming complexes, leading to very poor starch digestibility. Sorghum has the poorest starch digestibility of all cereals. Sorghum genotypes with high phenolic and tannin content were found to be

associated with enzyme inhibition and starch molecule interaction. This impairs the starch digestibility, which increases the resistance of starch, which leads to a lower glycemic index of foods in which the starch is included (Moraes et al. 2018).

2.6 Fibers, Vitamins, and Minerals

Dietary fibers are mainly non-starch polysaccharides, which have numerous health benefits, such as positive effects on diabetes, tumors, and atherosclerosis. The whole grain of sorghum consists 10–25% of bran, of which 35–48% is insoluble dietary fiber (Miafo et al. 2019). Sorghum, being a fiber-rich food, has a low glycemic index, causing a slower and lower rise in blood glucose level (Moraes et al. 2018). Sorghum grains contain many and various types of vitamins and minerals. Minerals include Ca, Fe, K, Mg, P, and Zn (Motlhaodi et al. 2018). The most prevalent vitamins are of the B-complex such as pyridoxine (vitamin B6), riboflavin (vitamin B2), and thiamine (vitamin B1), and the fat-soluble vitamins such as A, D, E, and K (Przybylska et al. 2019).

2.7 Sorghum Protein

Cereal grains generally contain up to 20% proteins, but unlike animal protein, they have a low essential amino acid content. However, cereals are widely available and are a staple in many regions throughout the world, making them an attractive protein source and a source from which peptide nutraceuticals can be released, which could eliminate reactive oxygen (Szerszunowicz and Kłobukowski 2020). A wide range of protein content in sorghum grains have been reported (6–18%), of which the biggest percentage is storage protein, or prolamins. Sorghum prolamins are called kafirins and are located in the protein bodies of the endosperm. Sorghum kafirins make up a large portion of the total protein in whole kernels (48–70%) and an even higher portion (up to 80%) in decorticated kernels, while the rest consists of the albumins and globulins (Espinosa-Ramirez and Serna-Saldivar 2016). Discrete kafirin protein bodies are formed, and in mature endosperm, these bodies form a tight matrix with starch granules. These matrices affect the processing quality of sorghum and contribute to grain hardness and digestibility. Sorghum storage proteins are classified according to solubility, structure, amino acid composition, and molecular mass. They consist mainly of albumins (water extractable fraction), globulins (dilute salt extractable fraction), and kafirins or prolamins (alkali soluble fraction) (Wong et al. 2009).

Kafirins are hydrophobic proteins and are divided into three groups based on their molecular weight: α -kafirin (23–27 kDa), β -kafirin (16, 18, and 20 kDa), and γ -kafirin (28 kDa) (Espinosa-Ramirez and Serna-Saldivar 2016). The most dominant kafirin proteins are the α -kafirins, accounting for 70–80% of total storage proteins in sorghum grain (66–71% in opaque endosperm). They occur mostly in outer layers of grain and a reduction of α -kafirins is associated with a loss of vitreous endosperm texture (Wu et al. 2013). Sodium dodecyl sulfate polyacrylamide gel electrophoresis

(SDS-PAGE) has been used frequently in the past to separate the kafirins. With this technique, the α -kafirins resolve into two distinct bands with a molecular weight of 23 and 25 kDa, although these molecular weights can range from 22 to 27 kDa depending on different laboratory protocols and conditions (Cremer et al. 2014).

The second largest group of kafirins is the γ -kafirins. They make up a large part of the total kafirins in opaque endosperm (19–21%) and a smaller part (9–12%) in vitreous endosperm. They have a molecular weight of about 20 kDa. The β -kafirins comprises about 7–8% of total kafirins in vitreous endosperm (10–13% in opaque endosperm). The outside layers of the protein bodies contain most of the β - and γ -kafirins, where cross-links are formed between proteins via disulfide bonds (Oria et al. 2000). The very minor δ -kafirins ($M_r = 15$ kDa) form the fourth group, accounting for less than 1% of the mature grain total seed storage protein. Sorghum protein is poorly digestible, and the improvement of protein digestibility is still a major research goal in the sorghum fraternity. Protein digestibility is strongly related with the overall grain α - and γ -kafirins content (Wu et al. 2013; Elkonin et al. 2016). Dense, spherically shaped protein bodies are formed in the seed endosperm during grain development from the accumulation of the different kafirin groups. Kafirin protein bodies are quite compact, and on the periphery of these bodies, cross-links are formed. This probably causes the structural barriers preventing enzymatic access, which reduce the digestibility of these proteins (Oria et al. 2000).

Amino acids in grains are largely obtained from storage proteins, so this is a major factor determining the nutritional quality of the grain as a source of food for humans and animals. Both maize and sorghum proteins are deficient in the essential amino acids lysine and tryptophan, reducing the nutritional value of these cereals, which is an even bigger problem in sorghum than maize. Sorghum grain has low amount of essential amino acids like methionine, lysine, and isoleucine. The main amino acids in sorghum protein are histidine, leucine, phenylalanine, tyrosine, threonine, tryptophan, and valine (Mohapatra et al. 2019).

The γ -kafirins are rich in proline, cysteine, and histidine. The β -kafirins are rich in cysteine. As the δ -kafirins are so small, it could be inconsequential in influencing sorghum grain quality traits. However, it is rich in the essential amino acid methionine, which is often deficient in cereal proteins, along with lysine, threonine, and tryptophan. If the δ -kafirins fraction can be increased, it could enhance the nutritional quality of sorghum proteins (Laidlaw et al. 2010).

2.8 Other Proteins in Sorghum

Various antifungal proteins have been identified in sorghum, including chitinase, glucanase, thionin, defensin, protease inhibitor, and ribosome-inactivating proteins. There are also several bioactive proteins in sorghum such as amylase and protease inhibitors, as well as glycine-rich RNA-binding proteins, protein kinases, and glutathione S-transferase isoenzymes. Proteins involved in lysine catabolism were also isolated from sorghum, which included lysine 2-oxoglutarate reductase and saccharopine dehydrogenase (Lin et al. 2013).

3 Genetic Improvement of Sorghum Nutraceutical Content

The nutraceutical content of sorghum can be improved in breeding programs if sufficient genetic variability is available to select for higher content. Natural genetic variation can be exploited, or genetic variability can be introduced from wild and close relatives, many of which are available from conserved genetic resources. Molecular techniques such as linkage and association mapping, and more recently, genome-wide association studies, can be used to understand the underlying genetic architecture of nutraceutical content. Sorghum mutant libraries are also proving very useful to use reverse genetics to determine traits that are useful for sorghum improvement, which could include nutraceutical content. Genetic engineering and genome editing techniques can now be applied to improve important traits, which already include some nutraceuticals.

3.1 Genetic Resources for Genetic Improvement of Nutraceuticals and Nutritional Value

The sorghum genus consists of *Sorghum bicolor*. This belongs to the subgenera of *Eu Sorghum*, within which there are three species: *Sorghum bicolor*, *Sorghum halepense* (Johnson grass), and *Sorghum propinquum*. The *S. bicolor* species consists of three subspecies: *bicolor*, *verticilliflorum*, and *drummondii* (Sudan grass). The *bicolor* subspecies has 5 races (bicolor, caudatum, durra, guinea, and kafir) and 10 intermediate races (Ananda et al. 2020). There are numerous global genetic resources of conserved sorghum worldwide and over 240,000 accessions are conserved in ex-situ gene banks. Most of these are cultivated accessions (98.3%) and a small percentage are wild weedy relatives (1.7%) (Upadhyaya et al. 2016), although there may be duplicates which yet have to be identified and managed. There are at least 20 sorghum gene banks in the world. The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India have a collection of 37,949 accessions from 92 countries (ICRISAT 2021). The United States Department of Agriculture (USDA) National Plant Germplasm System (NPGS) at the Plant Genetic Resources Conservation Unit in Griffin, GA, maintains a collection of over 45,000 accessions (USDA 2021). The Institute of Crop Science, Chinese Academy of Agricultural Sciences (ICS-CAAS), China, holds a collection of 18,263 accessions, and the National Bureau of Plant Genetic Resources (NBPGR) of India conserves 20,221 accessions (NBPGR 2021). Wild relatives are widely used in cereal breeding programs and likewise has potential in sorghum breeding to enhance yield potential and genetic diversity.

3.2 Subset Collections as Sources for Marker-Assisted Breeding

Significant research has been done on accessions collected from Ethiopia. Usually a high number of accessions is screened in different growing environments, and then a subset of this is genotyped using a genotyping by sequencing (GBS) approach. A

core subset is then identified which best represents the genetic diversity present in the whole germplasm collection. A subset of 374 accessions followed by another subset of 1425 Ethiopian landrace accessions were analyzed (Girma et al. 2019). After that, another subset of 387 Ethiopian germplasm accessions were comprehensively phenotyped and genomic characterization was done (Girma et al. 2020). The natural variation and genetic structure of these populations were determined. Single nucleotide sequence polymorphisms (SNP) associated with important traits can be used as molecular markers in breeding programs (Girma et al. 2019, 2020).

3.3 Sorghum Linkage and Association Mapping Resources

Natural variation in populations has been shaped by natural and artificial selection, and a combination of the two. Variation is probably linked to specific agronomic adaptive traits that have ecophysiological relevance. Ecogeographical and historical information also largely determine natural variation, and analysis of this variation can shed light on the evolutionary processes that led to the genetic variation patterns. Sorghum germplasm is known to have high levels of natural variation (Boyles et al. 2019) and is very suited for genetic analysis. In linkage mapping, recombination is generated by crossing parents, while in association mapping, historical recombination events are used. Linkage and association mapping are widely used to characterize the genetic architecture of traits. This includes information of how many loci are involved and how they are distributed, the gene action determining the traits, and linkage and allele frequency. With this knowledge, scientists can hypothesize on the genes that underlie variation in studied traits.

3.3.1 Linkage Mapping Resources

Many genetic traits in sorghum have been investigated through linkage mapping in the past 25 years, such as plant and flowering traits, pigmentation, drought and cold tolerance, and disease resistance (Xin et al. 2021). Almost no research has been reported on seed composition, including nutraceutical content, with the exception of kafirins.

3.3.2 Association Mapping Resources

There has been significant developments in the methods that can be used to generate high-density genome-wide markers. This has caused a shift towards genome-wide association studies (GWAS) from linkage mapping and gene association. The sorghum association panel is currently the most widely used sorghum GWAS resource. This panel was designed to capture and represent global plant diversity, and plant function and end uses. This association panel has genotyping by sequencing (GBS) marker data (Hu et al. 2019) and is available from the Germplasm Resources Information Network (GRIN). All the sorghum association panel accessions flower in temperate latitudes. This panel has been used for GWAS studies of many traits (Mural et al. 2020). There is also a bioenergy association panel, which is another global diversity panel that is available from GRIN, where the majority of the

accessions are tropical and photoperiod-sensitive. Although this panel consists mainly of tropical accessions, sweet and forage sorghums are also included. The bioenergy association panel has associated SNP data from GBS and whole-genome resequencing (Bellis et al. 2020; Lozano et al. 2021). A sorghum collection has been made available for phenotyping and association studies, consisting of 2000 georeferenced accessions, which were genotyped with GBS. Most of these accessions are also available from GRIN (Bellis et al. 2020). A large number of landrace accessions have been collected from Africa, for which GBS SNP data are available. There are entire GRIN collections from Niger, Nigeria, Senegal, Ethiopia, and Sudan, and for sweet sorghum.

3.3.3 Multi-parent Mapping Resources

Techniques such as linkage mapping are powerful and sensitive, while association mapping has the strength of power and sensitivity, which created an interest to combine the best of these techniques in different genetic approaches. This has led to the use of multi-parent mapping approaches, which include nested association mapping (NAM), backcross NAM, and multi-parent advanced generation intercross (MAGIC), all which have been applied in sorghum research (Boyles et al. 2019). The NAM approach is based on that which was developed for maize and now applied in many crops where multiple recombinant inbred line (RIL) families are developed that have diverse founders but share a common parent (Gage et al. 2020). RTx430 was selected as the common founder line for the global grain sorghum NAM resource (primarily because it has for decades been the most important public pollinator line), together with 10 diverse global founders (all part of the sorghum association panel), from which a total of more than 2200 RILs in 10 families were developed. The common germplasm between the association panel and the NAM allows the validation and comparison of the data from these two sources (Olatoye et al. 2020).

In the past, mainly biparental linkage families were used in research, but the aim of the MAGIC approach is to increase allelic diversity and at the same time to increase the power and specificity of quantitative trait loci (QTL) detection relative to GWAS. All founder lines contribute equally to the MAGIC in order to balance allele frequencies. One sorghum MAGIC resource has been developed and is currently available from the developers (Ongom and Ejeta 2018). There is also a backcross NAM (BCNAM) approach, where backcrosses are made and selection done to recover the common elite parent phenotype, which may have more of a breeding advantage (Jordan et al. 2011).

3.4 The Sorghum Mutant Library

Xin et al. (2008) developed a pedigree mutant library for functional genomic studies by using a systematic approach in such a way that all mutations are captured and preserved. Individual seeds from an elite inbred line, BTx623 (used to generate the first sorghum reference genome), were treated with the chemical mutagen ethyl

methane sulfonate (EMS). Seeds were soaked in various concentrations of EMS to optimize the chances to obtain desirable mutations. This mutant library is a permanent resource, which can be used by scientists to test mutants under different growing conditions, to enhance breeding efforts. Currently, 6400 independent seed pools are available in this library to breeders and scientists. These are an excellent resource for sorghum breeding, as mutants with potentially useful traits can be selected for sorghum improvement programs (Xin et al. 2021). Purdue University has followed a similar approach to generate approximately 10,000 pedigreed seed pools (Addo-Quaye et al. 2018). A wide range of phenotypes that could be used for sorghum improvement have been identified and selected within this pedigreed mutant library. The traits and phenotypes identified in this mutant library have the potential to be used in sorghum breeding programs and in genomic studies. Both forward and reverse genetic resources have been developed that can be used to explore the mutant library traits (Wang et al. 2021).

Next-generation sequencing techniques have been developing rapidly, providing large numbers of DNA markers, which are also financially affordable. The output of high-quality DNA sequences has been increasing, and next-generation sequencing has become much cheaper. Whole genome sequencing of pedigreed mutants has become a resource for reverse genetics by searching gene mutations online. Next-generation sequencing has been used to annotate SNP markers. Mutant phenotypes of interest can be studied to map and identify causal mutations through mapping-by-sequencing or next-generation mapping (Hartwig et al. 2012).

Addo-Quaye et al. (2018) introduced mutations with EMS and then sequenced 586 mutants, where they identified 1,275,872 homozygous and 477,531 heterozygous mutations. In the sorghum genome, as is the case in other crops, the mutations often have deleterious effects. The sequenced data of the mutant collections can be a very useful resource for reverse genetics. These collections can be used together with other methods, such as genome-wide association and biparental mapping of QTLs of important agronomic traits, to validate candidate genes (Xin et al. 2021). Although very limited research has been done on the genetics of nutraceuticals so far (with the exception of kafirins) compared to the adaptive and yield-related traits, the resources are there, and they should be used to generate data for genetic improvement.

3.5 Genetic Engineering Approaches for Improving Nutritional Composition of Sorghum Grain

Recent years have been marked by significant progress in sorghum genetic transformation technologies that have made it feasible to use genetic engineering and genome editing methods to improve the nutritional properties of sorghum grain. Such progress is based on improved technologies of plant regeneration in sorghum tissue culture, in particular, the use of nutrient media containing increased concentrations of phosphate, proline, and asparagine, which reduces the release of phenolic pigments characteristic of cultivated sorghum tissues and increase embryogenic potential (Elkonin and Pakhomova 2000), as well as elevated levels of copper

ions, which improve the development of the root system of regenerants. Efficient media for plant regeneration allowed substantial improvement of genetic transformation methods of sorghum either through biolistic DNA delivery (Belide et al. 2017) or *Agrobacterium*-mediated genetic transformation (Do et al. 2016). In the latter case, significant improvements are due to the use of “hypervirulent” *A. tumefaciens* strain NTL₄ or another *Agrobacterium* strain containing specially designed Ti-plasmids with additional copies of *vir*-genes (“superbinary” vectors) or containing “helper” plasmids with additional *vir*-genes that enhance the transfer of T-DNA from *Agrobacterium* cells to sorghum cells. An important factor that also contributed to the increase in the efficiency of *Agrobacterium*-mediated genetic transformation in sorghum was the obtaining of an auxotrophic mutant LBA4404 Thy – incapable of growth on a media without thymidine, the use of which made it possible to significantly simplify the transformation procedure. The creation of binary vectors carrying the genes of morphogenetic regulators BABY BOOM and WUSCHEL, which promote the direct development of embryoids from scutellum cells of immature embryos, and thereby increase the number of regenerants and the frequency of transgenic plants, also had a significant effect (Che et al. 2022).

The main goals of genetic engineering approaches for improving nutritional value of sorghum grain are improvement of the digestibility of sorghum storage proteins, modification of starch content, enrichment of sorghum grain with essential amino acids and carotene, and decreasing the level of phytate that reduces bioavailability of minerals and phosphate. Some of these goals are being successfully realized currently, while the achievement of other goals is the task of future research.

3.5.1 Improvement of Kafirin Digestibility

Kafirins are resistant to proteolytic digestion, which is one of the main causes of reduced nutritional value of sorghum grain. This resistance to proteolytic digestion not only reduces their digestibility by animals and humans but also reduces the digestibility of grain starch, since undigested kafirins prevent the amylolytic cleavage of starch granules (Duressa et al. 2018).

Significant research has been devoted to the study of the factors that cause protein indigestibility. The reasons for kafirin resistance to protease digestion are multifactorial. It is assumed that kafirin resistance to protease digestion is caused by the chemical structure of kafirins, which are abundant with sulfur-containing amino acids capable to form S–S bonds that result in the formation of kafirin oligo- and polymers resistant to protease digestion; interactions of kafirins with non-kafirin proteins and nonprotein components such as polyphenols and polysaccharides; and the organization of different kafirins in the protein bodies, with the β - and γ -kafirin located at the periphery of the protein body, while the most abundant α -kafirin is located within the protein body. Such a spatial arrangement of kafirins in the protein bodies of sorghum, apparently, is a consequence of the prolamin synthesis pattern in the process of kernel development. In maize, it was found that starting from 4 to 5 days after fertilization, protein bodies accumulate γ - and β -prolamins, while later on, intensively accumulating α -prolamins moved γ - and β -prolamins to the periphery of the protein body. Due to the peripheral occurrence of γ -kafirin (which is

considered as the most stable to protease digestion) in the protein body, it is generally accepted that it reduces digestibility of α -kafirin, which comprises up to 80% of total endosperm kafirins (Duressa et al. 2018). In addition, γ -kafirin effectively forms oligo- or polymers of high molecular weight, which has high resistance to protease digestion.

This hypothesis was confirmed by the study of a mutant, P721Q, with a high level of lysine and improved digestibility of kafirins, obtained by chemical mutagenesis. In this mutant, the protein bodies had irregular shape, having deep invaginations, while normally they have a round shape. The γ -kafirin is located only in the bottom of these invaginations and does not form a continuous layer, which prevents digestion of α -kafirin in normal sorghum. It was suggested that such a structure of the protein bodies determines the high digestibility of proteins in the P721Q mutant and lines derived from it (Oria et al. 2000). Floury type endosperm grains are formed due to this mutation and lysine content is increased. It was therefore denoted with the symbol *hdhl* (*high digestibility high lysine*). Further investigation of this mutation using two-dimensional gel electrophoresis and mass spectrometry showed an increase of non-kafirin proteins (such as cytoskeleton and chaperone proteins, and the proteins involved in amino acids and carbohydrates synthesis) and a decrease in kafirin content in *hdhl* endosperm. The overexpression of chaperone proteins, which are probably involved in the repair of protein misfolding that was caused by the mutation, is in part responsible for increased lysine content of the P721-opaque sorghum (Benmoussa et al. 2015).

Studies have been undertaken on the chromosome localization of this mutation. In the research of Winn et al. (2009), with the analysis of a hybrid population obtained by crossing highly digestible line P850029 (derived from P721Q) and the wild type line Sureno, two QTLs were identified. These QTLs (both from the high-digestibility parent) were located on chromosome 1 in genomic regions within 20 cM from each other. They were in repulsion phase, meaning one QTL (locus 1 from the HD parent) unfavorably affects digestibility and one QTL (locus 2 from the HD parent) favorably affects digestibility. Protein digestibility can be increased if this linkage in repulsion is broken, which will allow the recombination of favorable alleles. It is noteworthy that another research group found that the proteinase inhibitor gene is located on chromosome 1 (Duressa et al. 2018). Perhaps QTLs identified by Winn et al. (2009) are linked to this gene.

In another investigation of chromosome localization of the *hdhl* mutation, one major QTL was identified on chromosome 5, in the 58 Mb region that overlaps with the genomic loci of the 17-kafirin gene cluster. In this research, the F₂ population generated from a cross between a normal line (BTx623) and a high-digestible mutant, P721Q was used (Massafaro et al. 2016). The F₂ plants with highly digestible protein displayed the unique protein body structure of P721Q. The results of this investigation confirms the data obtained by Wu et al. (2013) which showed that the *hdhl* phenotype in the P721Q mutant and lines derived from it is a consequence of a point mutation in one of the genes from the *k1C* family located on chromosome 5, encoding the 22 kDa α -kafirin. Sequencing revealed a point mutation in the nucleotide sequence that encodes the signal polypeptide responsible

for the packaging of α -kafirin inside the protein body (substitution at position 61 G→A) (Wu et al. 2013). This mutation resulted in the formation of a missense codon at the last amino acid of the signal sequence. Despite the fact that the mutation was found in only one of the genes of this cluster, its presence was sufficient to cause significant changes in the structure of protein bodies and the digestibility of kafirins. It was hypothesized that this mutation decreases the accumulation of α -kafirin in protein bodies that leads to a change in their ultrastructure and increases their sensitivity to the action of proteases (Wu et al. 2013).

In addition, a comparative analysis of the nucleotide sequences of all 27 kafirin genes in sorghum samples with high and low digestibility revealed four α -kafirin alleles localized on chromosome 5, which are closely associated with digestibility. Three alleles were associated with high digestibility (*Sobic.005G185600*, *Sobic.005G188800*, and *Sobic.005G189000*) and one with low digestibility (*Sobic.005G192801*). In silico predictive analysis showed the variants cause missense change in the amino acid sequences of the corresponding proteins (Duressa et al. 2020).

3.5.2 RNA Interference Technology

The use of biotechnology methods, in particular, RNA interference technology, opens up much broader prospects for obtaining mutants with improved digestibility of kafirins. This technology has been widely used for a long time to modify the synthesis of storage proteins as well as starch and other nutrients of endosperm in a number of cereals (Elkonin et al. 2016).

A number of research groups have been doing studies on the induction of RNA silencing of kafirin genes (Da Silva et al. 2011a, b; Kumar et al. 2012; Grootboom et al. 2014). RNA silencing was induced by genetic constructs carrying inverted repeats of several kafirin genes (α 1, δ 2, γ 1, and γ 2) separated by the *ADH1* gene intron sequence. mRNA transcribed from this construct forms a double-stranded hairpin and underwent enzymatic degradation. The constructs were driven by the 19-kDa maize α -zein promoter (Da Silva et al. 2011a, b). In another study (Kumar et al. 2012), the genetic construct used to induce γ -kafirin silencing consisted of the complete γ -*KAFIRIN* gene sequence under the control of its own promoter. As a terminator, the sequence of the tobacco mosaic virus ribozyme gene was used. The expression of this gene should destroy γ -kafirin mRNA. In another construct, to induce α -kafirin silencing, α -kafirin inverted repeats separated by the intron sequence of the *Arabidopsis* gene encoding the D1 spliceosome protein were used; this construct was driven by the α -kafirin promoter (Kumar et al. 2012). Later, another complex construct consisting of inverted fragments of γ 1-, γ 2-, and δ -kafirin genes was used to obtain transgenic sorghum plants with silencing of kafirins (Grootboom et al. 2014).

The transgenic plants obtained in these experiments had a modified protein body structure which resembled that of the P721Q mutant, and an increased digestibility of kafirins, which was observed when both raw and cooked sorghum flour was treated with pepsin. One such example is the transgenic plants of Tx430, carrying a genetic construct for silencing α - and γ -kafirin. They were identified as having

improved in vitro protein digestibility (78% and 61% for the raw and cooked flour, respectively), while in the non-transgenic control, the digestibility varied in the ranges of 40–50% and 34–40%, respectively (Da Silva et al. 2011a). The genetic construct to suppress only δ - and γ -kafirins also improved the digestibility of raw flour but did not affect the digestibility of cooked flour. In the experiments of Kumar et al. (2012), the cooked flour from transgenic kernels carrying a genetic construct for γ -kafirin suppression did not differ from the non-transgenic control, while α -kafirin suppression improved the digestibility of such flour.

Based on the hypothesis that the surface location of γ -kafirin in protein bodies is the reason for the low digestibility of sorghum protein bodies, a genetic construct (NRKAFSIL) was created that should prevent the accumulation of the γ -kafirin without affecting the accumulation of other kafirins (Elkonin et al. 2016). This construct contained the fragment of the nucleotide sequence of γ -kafirin gene (GenBank Accession No: M73688) in direct and inverted orientation, separated by the maize ubiquitin intron sequence. The construct was driven by the 35S promoter. Using *Agrobacterium*-mediated genetic transformation, this construct was introduced into two sorghum cultivars: Zheltozernoe 10 (Zh10) and Avans. Electrophoretic spectra of endosperm proteins were compared before and after pepsin digestion, showing that in the transgenic plant, the amount of undigested kafirin monomers and total undigested protein was significantly lower than in the original non-transgenic line (by 1.7–1.9 times) when quantitative SDS-PAGE analysis was done. The level of digestibility reached 85–92%, while in the original line, this value was about 60%. The effect of increased digestibility of kafirins was traced to the T₄ generation; however, in some cases, it disappeared, possibly due to the instability of the introduced genetic construct or due to its silencing.

It should be noted that RNAi silencing of the γ -kafirin gene also caused a decrease in the content of α -kafirins. This phenomenon was especially pronounced in the Avans-1/18 transgenic line (Elkonin et al. 2021). At the same time, the functioning of the construct for the γ -kafirin gene silencing did not lead to a significant decrease of the total protein content of the grain compared to the original non-transgenic cultivar (14.3% vs. 15.5%). This fact, apparently, is a consequence of the balancing of protein synthesis in the kernels.

Another feature of transgenic sorghum lines with silencing of γ - and α -kafirins is the modification of the endosperm texture, in particular, the reduction of vitreous endosperm layer and formation of kernels containing floury endosperm characteristic to the P721Q mutant (Da Silva et al. 2011b; Kumar et al. 2012; Grootboom et al. 2014). Apparently, expression of RNAi genetic constructs also affects the formation of the protein-carbohydrate matrix, which caused formation of the vitreous endosperm layer. In experiments with the genetic construct for γ -kafirin gene silencing in variety Zh10, transgenic plants were obtained that had kernels with different endosperm types (Fig. 1).

Normal endosperm with a thick or thin vitreous layer, floury endosperm, or a modified type of endosperm in which the vitreous layer developed as sectors or spots surrounded by floury endosperm were seen (Elkonin et al. 2016). In these experiments, the kernels of some transgenic plants that exhibited a thick vitreous

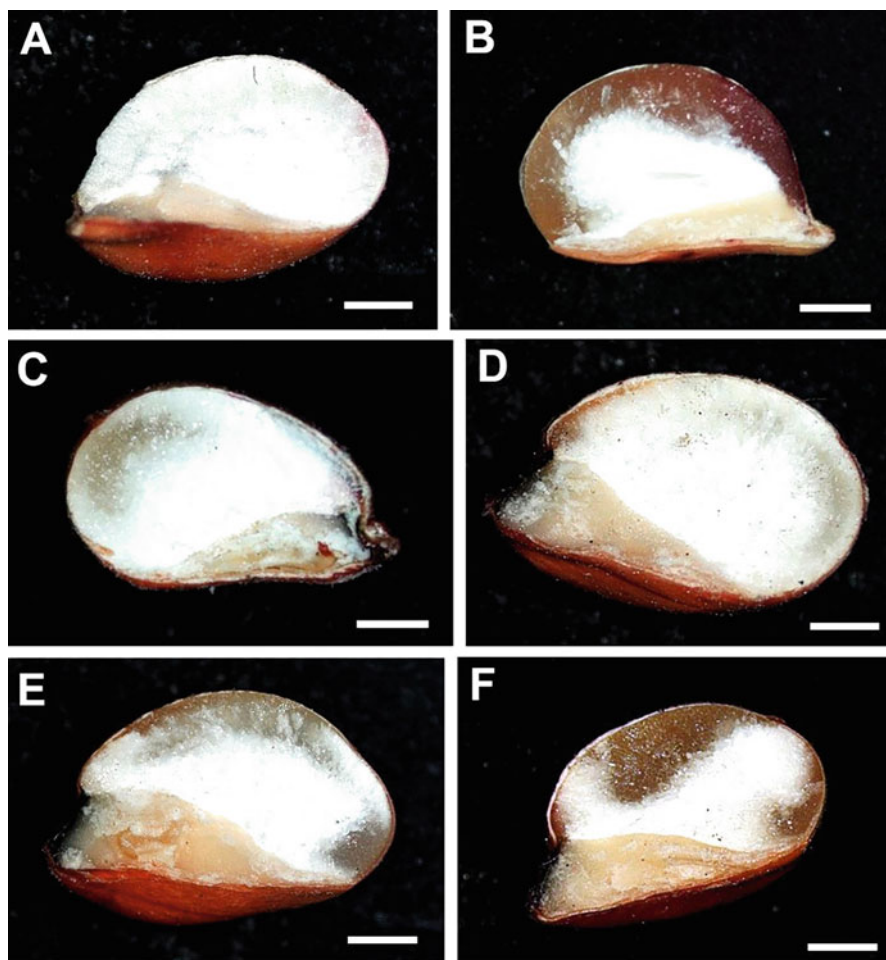


Fig. 1 Cross sections of kernels of transgenic sorghum plants with genetic construct for silencing of the γ -kafirin gene (Elkonin et al. 2016). (a) Kernel with floury endosperm set on the plant from T₃ generation; (b) kernel of original non-transgenic line Zh10 with thick vitreous endosperm (marked by arrows); (c–e) modified endosperm type with blurs and sectors of vitreous endosperm observed in kernels of different T₁ and T₂ plants; and (f–h) irregularly developed vitreous endosperm developed in kernels of T₂ and T₃ plants. Bar = 1 mm (© 2017 Elkonin LA, Italyanskaya JV, Panin VM, Selivanov NYu. Originally published in “Plant Engineering”, InTech, Zagreb (Croatia) under CC BY 3.0 license. Available from: <https://doi.org/10.5772/intechopen.69973>)

endosperm had a digestibility level as high as 92%, while the amount of undigested monomers was reduced by 17.5 times, and the amount of total undigested protein was 4.7 times less compared to the original lines. Previously, transgenic plants with similar endosperm texture were also observed in the Tx430 transgenic line with a construct for silencing α - and γ -kafirins.

It seems that the different types of endosperm that form are due to the particular expression of genetic constructs in the genome of the recipient line. In transgenic line Avans-1/18, in the T₁ generation, a revertant with vitreous endosperm and significantly reduced digestibility of kafirins was found (Elkonin et al. 2021). This revertant contained part of the NRKAFSIL genetic construct determining resistance to the selective agent (*bar* gene). At the same time, this revertant has a deletion in the *ubi1* intron, which is a part of the construct for silencing of the γ -*KAFIRIN* gene, that, apparently, caused inefficiency of silencing and loss of characteristic features of mutation. These data indicate that the 588 bp *ubi1* intron sequence can be used as a molecular marker in the screening of plants with high protein digestibility during hybridization of the Avans-1/18 mutant with various sorghum varieties.

The well-known correlation between the high digestibility of kafirins and floury endosperm was confirmed by these findings (Duressa et al. 2018). Such a correlation could hardly be explained by impaired synthesis of γ -prolamines (γ -kafirins), which are believed to play an important role in the interaction of protein bodies with starch granules, since mutants with impaired synthesis of α -kafirins also have a floury endosperm type (Da Silva et al. 2011b; Kumar et al. 2012; Grootboom et al. 2014). This correlation is possibly due to a violation of the formation of non-kafirins in mutants with suppressed synthesis of α - or γ -kafirins, which form a protein matrix characteristic of vitreous endosperm. In mutants where kafirins are silenced, an increase in essential amino acid contents of lysine and threonine is seen, which appears to be due to increased synthesis of other proteins, including those with a higher content of essential amino acids. Thus, a significant increase of lysine content (an increase of 1.2 g/100 g protein compared to the non-transgenic control) was found in transgenic sorghum plants where α -, γ -, δ -kafirins genes were silenced due to the presence of complex genetic constructs for RNAi silencing and the presence of the lysine ketoglutarate reductase gene (which controls catabolism of free lysine) (Da Silva 2012).

In the transgenic plants of variety Zh10 containing the genetic construct for γ -kafirins silencing, the proportion of lysine increased 1.6–1.7 times (Elkonin et al. 2016); in the Avans-1/18 line, the increase was 75%, from 0.36% in the original line to 0.63%. It seems that the increase was caused by a decrease in amount of α -kafirins, which are poor in lysine and threonine, while the synthesis of other proteins was not affected. This caused an increase in the relative proportions of lysine and threonine. The suppression of γ -kafirins synthesis probably has no effect on the synthesis of proteins that are rich in lysine and threonine, but it prevents the accumulation of α -kafirins.

The use of RNA-interference technology, however, has a number of significant limitations: the genetic construct for silencing may undergo partial or complete destruction in the recipient genome (Elkonin et al. 2021), or its expression itself can undergo silencing. In addition, the functioning of the RNA silencing mechanism is an epigenetic process, and epigenetic processes in plants largely depend on environmental conditions (such as temperature, soil, and air humidity) (von Born et al. 2018). In the transgenic line Zh10, carrying the genetic construct for silencing

γ -kafirin, the level of digestibility of kafirins in plants grown in a field plot under natural moisture conditions was lower than in plants grown in outdoor vessels (Elkonin and Italyanskaya 2017). In some cases, off-target, nontarget, and unintended effects of genetic constructs for RNAi were observed (Christiaens et al. 2018). Besides these scientific problems, public concerns on using transgenic plants are a serious problem, which limits widespread use of sorghum lines carrying genetic constructs for RNA silencing for food and feed purposes.

3.5.3 Genome Editing Technologies

Site-directed mutagenesis using genetic constructs carrying the CRISPR/Cas system is one of the most effective technologies that are actively used to solve a variety of problems in genetics and plant breeding (Zhu et al. 2020). This approach allows the changing of the structure of the genes of plants, therefore, without introducing foreign genetic information, to change the plant metabolism in the necessary direction (Song et al. 2016). At the same time, in the offspring of mutants, due to recombination, it is possible to select plants that carry the induced mutation but are free from the genetic construct that induced it. As a result, the resulting mutants practically do not carry foreign genetic information; therefore, they are not transgenic organisms.

The CRISPR/Cas9 system includes Cas9 endonuclease and guide RNA (gRNA), which directs Cas9 endonuclease to the target nucleotide sequence (Song et al. 2016). The classical variant of the Cas9 nuclease recognizes the NGG-3' Protospacer Adjacent Motif (PAM) sequence adjacent to the target (protospacer), which thus serves as the identification mark of the target in the edited genomic DNA. Cas9 endonuclease produces double-strand breaks in the target DNA located three nucleotides upstream of the PAM sequence. These breaks result in insertions or deletions at the target site, which can lead to frameshifts and null mutations.

The CRISPR-Cas system was successfully used to induce mutations in the nucleotide sequence of the 22 kDa signal polypeptide of α -kafirin (Li et al. 2018). Mutations were deletions ranging in size from 1 to 33 nucleotides, and, more rarely, insertions ranging in size from 1 to 16 nucleotides. In the kernels of T₁ and T₂ plants, a reduced level of α -kafirin and an altered structure of protein bodies were observed; some T₂ plants had higher protein digestibility and increased lysine levels. These results indicate the promise of using genome editing techniques to improve the nutritional value of sorghum grain.

Edited sorghum plants with mutations in the nucleotide sequences of the β - and γ -kafirin gene have also been reported (Massel et al. 2022). In this study, high efficiency of endogenous *U6* promoters (in particular, *SbU62.3* promoter) was shown to improve gene editing efficiency in sorghum of up to 90% of experimental plants.

A series of binary vectors (pC1-pC4) were created for site-directed mutagenesis of genes encoding α - and γ -kafirins (Gerashchenkov et al. 2021). These vectors contain the *Cas9* endonuclease gene under the control of the *ubi1* promoter and gRNA nucleotide sequences that are complementary to target sites encoding signal polypeptides of α - and γ -kafirins (the *k1C5* and *gKAF1* genes, respectively). Using *Agrobacterium*-mediated genetic transformation, the constructs for site-directed mutagenesis of these genes were introduced into the genome of cv. Avans. Using

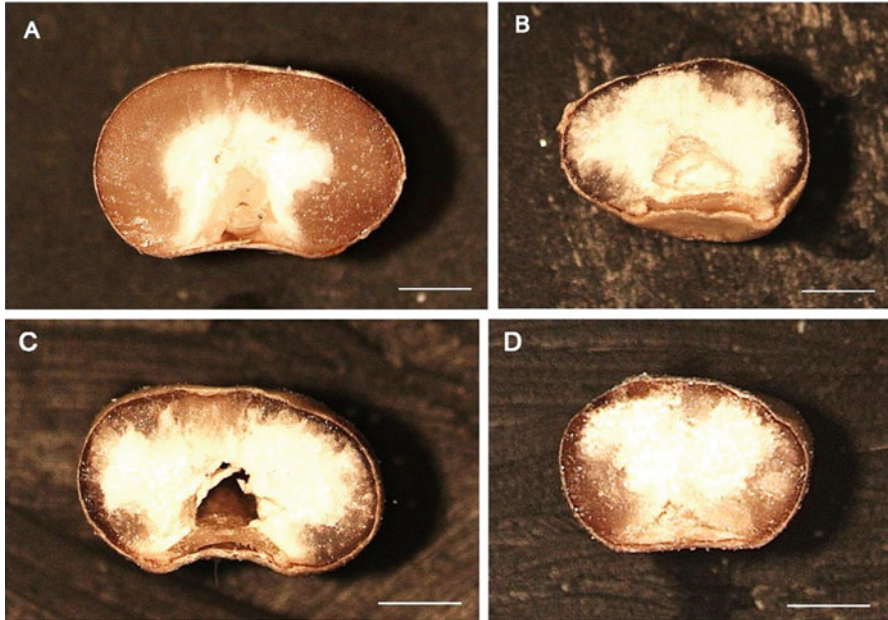


Fig. 2 Cross sections of kernels set on the panicle of the sorghum plant carrying a genetic construct for α -KAFIRIN (*k1C5*) gene editing target ((b), (c), (d)); ((a) – kernel of original cv. Avans). Bar 1 mm (Amer J Plant Sci, 2021, 12: 1276–1287, with permission)

the pC2 vector (to induce mutations in the *k1C5* gene), regenerants (T_0) with kernels with modified endosperm texture were obtained, in which a significant reduction in the vitreous endosperm was observed (Fig. 2).

This demonstrates the efficiency of genome editing approaches for improvement of nutritional value of sorghum grain. The successful results obtained in these studies demonstrate the validity of the choice of kafirin signaling polypeptides as targets for improving the nutritional value of sorghum grain. However, it is likely that this approach is only one of the ways to use the CRISPR/Cas method to improve the nutritional value of sorghum grain, and other approaches will be used to solve this problem in future experiments.

3.5.4 Synthetic Biology Approaches

With the help of synthetic biology approaches, transgenic lines with increased protein digestibility and increased protein content were obtained that contain an artificially synthesized β -kafirin gene (Liu et al. 2019). This gene encoded modified β -kafirin protein with additional proteolytic cleavage sites that should improve its digestibility. Some of the resulting transgenic lines had higher protein content in the seeds (by 11–37%) and a higher digestibility of kafirins (by 11–21%) compared to the non-transgenic original variety. The protein bodies had an irregular shape with invaginations characteristic for highly digestible sorghum lines.

In order to increase the content of lysine in the grain, transgenic sorghum lines carrying the gene encoding the high-lysine protein of barley, *gordotinin*, were obtained. *Agrobacterium*-mediated genetic transformation was used to introduce the gene encoding the high lysine analog (HT12 protein) of the *Hordeum vulgare* α -hordothionin protein under the control of the 27 kDa maize γ -zein promoter and terminator into the genome of two sorghum lines, P898012 and PHI391. The *A. tumefaciens* strain LBA4404 contained a “super-binary” vector with two unlinked T-DNA cassettes. The one cassette contained the lysine-rich *HT12* gene with the second cassette containing a herbicide-resistant *bar* gene as a selectable marker. The two different T-DNA cassettes in the co-transformation vector allows the segregation of the marker and trait genes in the progeny of the primary transformants. This also allows the elimination of the marker gene and consequently marker-free transgenic plants are obtained. Three high levels of the HT12 protein were expressed in the grain of the five independent transgenic events that were co-transformed with both genes, with a 40–60% increase of lysine.

Provitamin A is very important to human health, but sorghum grain has low levels of β -carotene or provitamin A. In order to increase β -carotene content in sorghum grain, the genetic construct encoding a number of enzymes involved in the carotenoid biosynthesis pathway has been introduced into the genome of sorghum line Tx430 (Che et al. 2016). These enzymes are: 1-deoxyxylulose 5-phosphate synthase, the precursor for carotenoid biosynthesis, *Zea mays* phytoene synthase 1, and the *Pantoea ananatis* carotene desaturase, involved in synthesis of phytoene and lycopene, respectively, which are the β -carotene precursors. This introduction resulted in an increase of β -carotene levels in mature seeds of transgenic plants of up to 9.1 $\mu\text{g/g}$ (compared to 0.5 $\mu\text{g/g}$ in non-transgenic control seeds). The β -carotene in plants can degrade during storage due to oxidation. To counter this effect, the barley *HGGT* gene encoding homogentisate geranylgeranyl transferase was introduced into the same genetic construct. Homogentisate geranylgeranyl transferase is involved in synthesis of vitamin E, and this vitamin has strong antioxidant effects. It was found that co-expression of the *HGGT* gene stacked with carotenoid biosynthesis genes, enhanced all-trans- β -carotene accumulation, and reduced β -carotene oxidative degradation. This led to stable provitamin A levels in sorghum seeds. Field trials of transgenic plants with increased carotene content showed that the all-trans- β -carotene levels were increased by about 20-fold in transgenic lines compared to the non-transgenic controls (Che et al. 2019). This is an example of how effective genetic engineering can be for modifying plant metabolism to meet human needs.

Thus, development of genetic transformation techniques allowed the development of sorghum lines with significantly improved protein digestibility, increased content of lysine and other essential amino acids, and β -carotene. However, there still remains a number of unsolved problems. Future research should focus on the development of lines with reduced phytate content, which lowers bioavailability of Fe, Zn, and phosphate, as well as lines which have high protein digestibility of seed storage proteins, and vitreous endosperm. Starch digestibility is also an important question because improving the caloric value of staple food is of great importance. Identification of the naturally occurring allele of pullulanase (*SbPUL-RA*) – starch

debranching enzyme – that confers significantly higher *in vitro* starch digestibility (Gilding et al. 2013) shows that this trait may be improved using a genome editing approach. These innovations, coupled with development of genomics of nutritive traits, will significantly improve the gene pool of existing varieties and hybrids of sorghum and make this important crop more in demand in world agricultural production.

4 Conclusion

Sorghum is an underutilized cereal, well adapted to changing climatic conditions, which are causing increased temperatures and an increased frequency of drought spells. The inclusion of especially whole grain sorghum in the diet may help avoid chronic lifestyle illnesses. The inclusion of sorghum grain as a regular part of the human diet has the potential to reduce the risk of cardiovascular diseases, some types of cancer, and type II diabetes. Sorghum has highly resistant starch, a high fiber content, high levels of bioactive compounds, and kafirin protein with potential benefits. The bioactivity of sorghum grain is influenced by phenolic compounds such as phenolic acid, flavonoids, stilbenes, and tannins, and it is rich in pro-cyanidins (condensed tannins) and 3-deoxyantocyanidins. Sorghum tannins, also called proanthocyanidins, have shown anti-inflammation and anticancer properties. Sorghum also contains B-complex vitamins, fat-soluble A, D, E, and K vitamins, and minerals such as potassium, phosphorus, magnesium, and zinc. The generally high antioxidant activity of sorghum grain could contribute towards combating diseases associated with oxidative stress. Sorghum grain also has cholesterol-lowering and antimicrobial properties. It also was shown to improve glucose metabolism, which could positively affect diabetes. Sorghum is gluten-free with a high fiber content, which could benefit celiac disease patients. Sorghum could therefore be included in functional foods as a source of bioactive ingredients. Of the bioactive compounds in sorghum, phenolic compounds have been the main focus for research, so there is still a lot that is not known about the other compounds. With developing technology, new bioactive compounds may also be detected.

All the studies reported so far found significant natural genetic variation for measured bioactive compounds, although the growing environment also consistently had a large influence on the expression of compounds. As was reported above, there are many untapped genetic resources available in the form of germplasm collections, which can be screened for naturally high occurrence of bioactive compounds. Significant research has been done on the genetics of kafirin protein indigestibility and to a lesser extent on compounds like carotenoids. There is, however, a large gap in the knowledge of the genetics underlying the expression of other bioactive compounds in sorghum. Linkage and association mapping, GWAS and mutant libraries have been used to research the genetic architecture of especially adaptive traits, and yield and related traits, and genetic engineering has been applied to contribute to improved protein indigestibility and carotene levels in sorghum. There is, however, scant data on the genetics of seed composition, bioactive compounds, and traits linked to the

nutritional value of sorghum grain in general. When more is known about the genes linked to nutraceuticals and the gene action involved in expression of nutraceuticals, conventional breeding can be used to select and cross parents with good nutraceutical content, markers can be developed for marker-assisted selection, and genomics and genetic engineering could be used to improve nutraceutical content. Sorghum research is lagging behind that of other commercial crops such as maize, rice, and wheat. Likewise, research on the nutritional value of grains is lagging behind research on yield and yield-related traits, adaptive traits, and disease resistance. Databases should be generated to integrate all available information on seed nutraceutical and other nutritional components, which can be used by the sorghum community for sorghum improvement and genetic studies.

References

- Addo-Quaye C, Tuinstra M, Carraro N, Weil C, Dilkes BP (2018) Whole-genome sequence accuracy is improved by replication in a population of mutagenized sorghum. *G3* 8:1079–1094. <https://doi.org/10.1534/g3.117.300301>
- Ananda GKS, Myrans H, Norton SL, Gleadow R, Furtado A, Henry RJ (2020) Wild sorghum as a promising resource for crop improvement. *Front Plant Sci* 11:1108. <https://doi.org/10.3389/fpls.2020.01108>
- Belide S, Vanhercke T, Petrie JR, Singh SP (2017) Robust genetic transformation of sorghum (*Sorghum bicolor* L.) using differentiating embryogenic callus induced from immature embryos. *Plant Methods* 13:109. <https://doi.org/10.1186/s13007-017-0260-9>
- Bellis ES, Kelly EA, Lorts CM, Gao H, DeLeo VL, Rouhan G, Budden A, Bhaskara GB, Hu Z, Muscarella R, Timko MP, Nebie B, Runo SM, Chilcoat ND, Juenger TE, Morris GP, dePamphilis CW, Lasky JR (2020) Genomics of sorghum local adaptation to a parasitic plant. *Proc Natl Acad Sci U S A* 117:4243–4251. <https://doi.org/10.1073/pnas.1908707117>
- Benmoussa M, Chandrashekar A, Ejeta G, Hamaker BR (2015) Cellular response to the high protein digestibility/high-lysine (*hdhl*) sorghum mutation. *Planta* 241:70–77. <https://doi.org/10.1016/j.plantsci.2015.08.025>
- Bouargalne Y, Ben Mrid R, Bouchmaa N, Zouaoui Z, Benmrid B, Kchikich A, El Omari R, Kabach I, Mohamed N (2022) Genetic diversity for agromorphological traits, phytochemical profile, and antioxidant activity in Moroccan sorghum ecotypes. *Sci Rep* 12:5895. <https://doi.org/10.1038/s41598-022-09810-9>
- Boyles RE, Brenton ZW, Kresovich S (2019) Genetic and genomic resources of sorghum to connect genotype with phenotype in contrasting environments. *Plant J* 97:19–39. <https://doi.org/10.1111/tpj.14113>
- Bröhan M, Jerkovic V, Collin S (2011) Potentiality of red sorghum for producing stilbenoid-enriched beers with high antioxidant activity. *J Agric Food Chem* 59:4088–4094. <https://doi.org/10.1021/jf1047755>
- Burdette A, Gamer PL, Mayer EP, Hargrove JL, Hartle DK, Greenspan P (2010) Anti-inflammatory activity of select sorghum (*Sorghum bicolor*). *Brans J Med Food* 13:879–887. <https://doi.org/10.1089/jmf.2009.0147>
- Cabrera AH, Luzardo I, Ramirez AK, Morales E, Campos R, Gaytan (2020) Effect of the nixtamalization process on the protein bioaccessibility of white and red sorghum flours during in vitro gastrointestinal digestion. *Food Res Int* 134:1–13. <https://doi.org/10.1016/j.foodres.2020.109234>
- Cardoso L d M, Pinheiro SS, da Silva LL, de Menezes CB, de Carvalho CW, Tardin FD, Queiroz VA, Martino HS, Pinheiro-Sant’Ana HM (2015) Tocochromanols and carotenoids in sorghum

- (*Sorghum bicolor* L.): diversity and stability to the heat treatment. *Food Chem* 172:900–908. <https://doi.org/10.1016/j.foodchem.2014.09.117>
- Chaudhari SP, Powar PV, Pratapwar MN (2017) Nutraceuticals: a review. *World J Pharm Sci* 6:681–739. <https://doi.org/10.1007/s13555-018-0221-x>
- Che P, Zhao Z-Y, Glassman K, Dolde D, Hu TX, Jones TJ, Gruis DF, Obukosia S, Wambugu F, Albertson MC (2016) Elevated vitamin E content improves all-trans beta-carotene accumulation and stability in biofortified sorghum. *Proc Natl Acad Sci U S A* 113:11040–11045. <https://doi.org/10.1073/pnas.1605689113>
- Che P, Zhao Z-Y, Hinds M, Rinehart K, Glassman K, Albertsen M (2019) Evaluation of agronomic performance of β -carotene elevated sorghum in confined field conditions. In: Zhao Z-Y, Dahlberg J (eds) *Sorghum: methods and protocols, Methods in molecular biology*, vol 1931. Springer Science+Business Media, LLC, pp 209–220
- Che P, Wu E, Simon MK, Anand A, Lowe L, Yang M, Albertsen MC, Gordon-Kamm W, Jones TJ (2022) *Wuschel2* enables highly efficient CRISPR/Cas-targeted genome editing during rapid de novo shoot regeneration in sorghum. *Commun Biol* 5:344. <https://doi.org/10.1038/s42003-022-03308-w>
- Choi SC, Kim JM, Lee YG, Kim C (2019) Antioxidant activity and contents of total phenolic compounds and anthocyanins according to grain colour in several varieties of *Sorghum bicolor* (L.) Moench. *Cereal Res Commun* 47:228–238. <https://doi.org/10.1556/0806.47.2019.014>
- Christiaens O, Dzhambazova T, Kostov K, Arpaia S, Reddy Joga M, Urru I, Sweet J, Smaghe G (2018) Literature review of baseline information on RNAi to support the environmental risk assessment of RNAi-based GM plants. EFSA Supporting Publication 2018:EN-1424. <https://doi.org/10.2903/sp.efsa.2018.EN-1424>
- Cox S, Noronha L, Herald T, Bean S, Lee SH, Perumal R, Wang W, Smolensky D (2019) Evaluation of ethanol-based extraction conditions of sorghum bran bioactive compounds with downstream anti-proliferative properties in human cancer cells. *Heliyon* 5:e015891–e015897. <https://doi.org/10.1016/j.heliyon.2019.e01589>
- Cremer JE, Godwin ID, Liu L, Wang D, Bean SR, Tilley M, Wilson JD, Kaufman RC, Vu TH, Gilding EK, Godwin ID, Wang D (2014) Impacts of kafirin allelic diversity, starch content, and protein digestibility on ethanol conversion efficiency in grain sorghum. *Cereal Chem* 91:218–227. <https://doi.org/10.1094/cchem-04-13-0068-r>
- Da Silva LS, Taylor J, Taylor JR (2011a) Transgenic sorghum with altered kafirin synthesis: kafirin solubility, polymerization, and protein digestion. *J Agric Food Chem* 59:9265–9270. <https://doi.org/10.1021/jf201878p>
- Da Silva LS, Jung R, Zhao ZY, Glassman K, Grootboom AW, Taylor J, Taylor JRN (2011b) Effect of suppressing the synthesis of different kafirin sub-classes on grain endosperm texture, protein body structure and protein nutritional quality in improved sorghum lines. *J Cereal Sci* 54:160–167. <https://doi.org/10.1016/j.jcs.2011.04.009>
- De Oliveira KG, Vieira VA, Lanamar C, de Morais L, Pinheiro HM, Anunciacao PC, Beserra C, da Silva EC, Barros F (2017) Effect of the storage time and temperature on phenolic compounds of sorghum grain and flour. *Food Chem* 216:390–398. <https://doi.org/10.1016/j.foodchem.2016.08.047>
- Da Silva LS (2012) Transgenic sorghum: Effects of altered kafirin synthesis on kafirin polymerisation, protein quality, protein body structure and endosperm texture. [thesis]. Pretoria, South Africa: Department of Food Science, Faculty of Natural and Agricultural Sciences, University of Pretoria, South Africa
- Do PT, Lee H, Mookkan M, Folk WR, Zhang ZJ (2016) Rapid and efficient *Agrobacterium*-mediated transformation of sorghum (*Sorghum bicolor*) employing standard binary vectors and bar gene as a selectable marker. *Plant Cell Rep* 35:2065–2076. <https://doi.org/10.1007/s00299-016-2019-6>
- Duressa D, Weerasoriya D, Bean SR, Tilley M, Tesso T (2018) Genetic basis of protein digestibility in grain sorghum. *Crop Sci* 58:2183–2199. <https://doi.org/10.2135/cropsci2018.01.0038>

- Duessa D, Bean S, Amand PS, Tesso T (2020) Identification of variant α -kafirin alleles associated with protein digestibility in grain sorghum. *Crop Sci* 60:2467–2478. <https://doi.org/10.1002/csc2.20198>
- Elkonin LA, Italiyanskaya YV (2017) *In vitro* digestibility of storage endosperm proteins of transgenic sorghum plants carrying genetic construct for silencing of the gamma-kafirin gene. *Adv Cur Nat Sci* 12:96–100. [In Russian]
- Elkonin LA, Pakhomova NV (2000) Influence of nitrogen and phosphorus on induction embryogenic callus of sorghum. *Plant Cell Tissue Organ Cult* 61:115–123. <https://doi.org/10.1023/A:1006472418218>
- Elkonin LA, Italiyanskaya JV, Domanina IV, Selivanov NY, Rakitin AL, Ravin NV (2016) Transgenic sorghum with improved digestibility of storage proteins obtained by *Agrobacterium*-mediated transformation. *Russ J Plant Physiol* 63:678–689
- Elkonin LA, Panin VM, Kenzhegulov OA, Sarsenova SK (2021) RNAi-mutants of *Sorghum bicolor* (L.) Moench with improved digestibility of seed storage proteins. In: Jimenez-Lopez JC (ed) Grain and seed proteins functionality. IntechOpen Ltd., London, UK. <https://doi.org/10.5772/intechopen.96204>
- Espinosa-Ramirez J, Serna-Saldivar SO (2016) Functionality and characterization of kafirin-rich protein extracts from different whole and decorticated sorghum genotypes. *J Cereal Sci* 70:57–65. <https://doi.org/10.1016/j.jcs.2016.05.023>
- Espitia-Hernández P, Chávez González ML, Ascacio-Valdés JA, Dávila-Medina D, Flores-Naveda A, Silva T, Chacón XR, Sepúlveda L (2022) Sorghum (*Sorghum bicolor* L.) as a potential source of bioactive substances and their biological properties. *Crit Rev Food Sci Nutr* 62:2269–2280. <https://doi.org/10.1080/10408398.2020.1852389>
- FAOSTAT (2019) Food and agriculture data. Food and Agriculture Organization of the United Nations. <http://www.fao.org/faostat/en>
- Gage JL, Monier B, Giri A, Buckler ES (2020) Ten years of the maize nested association mapping population: impact, limitations, and future directions. *Plant Cell* 32:2083–2093. <https://doi.org/10.1105/tpc.19.00951>
- Gerashchenkov G, Elkonin L, Gerashchenkov K, Rozhnova N, Hiekel S, Kumlehn J, Chemeris AV (2021) Binary vector construction for site-directed mutagenesis of kafirin genes in sorghum. *Am J Plant Sci* 12:1276–1287. <https://www.scirp.org/journal/ajps>
- Gharaati S (2019) Extraction techniques of phenolic compounds from plants. In: Soto-Hernandez-M, Garcia-Mateos R, Palma-Tenango M (eds) Plant physiological aspects of phenolic compounds. IntechOpen. <https://doi.org/10.5772/intechopen.84705>
- Gilding EK, Frère CH, Cruickshank A, Rada AK, Prentis PJ, Mudge AM, Mace ES, Jordan DR, Godwin ID (2013) Allelic variation at a single gene increases food value in a drought-tolerant staple cereal. *Nat Commun* 4:1483. <https://doi.org/10.1038/ncomms2450>
- Girma G, Nida H, Seyoum A, Mekonen M, Nega A, Lule D, Dessalegn K, Bekele A, Gebreyohannes A, Adeyanju A, Tirfessa A, Ayana G, Taddese T, Mekbib F, Belete K, Tesso T, Ejeta G, Mengiste T (2019) A large-scale genome-wide association analyses of Ethiopian sorghum landrace collection reveal loci associated with important traits. *Front Plant Sci* 10:691. <https://doi.org/10.3389/fpls.2019.00691>
- Girma G, Nida H, Tirfessa A, Lule D, Bejiga T, Seyoum A, Mekonen M, Nega A, Dessalegn K, Birhanu C, Bekele A, Gebreyohannes A, Ayana G, Tesso T, Ejeta G, Mengiste T (2020) A comprehensive phenotypic and genomic characterization of Ethiopian sorghum germplasm defines core collection and reveals rich genetic potential in adaptive traits. *Plant Genome* 13: e20055. <https://doi.org/10.1002/tpg2.20055>
- Grootboom AW, Mkhonza NL, Mbambo Z, O’Kennedy MM, da Silva LS, Taylor J, Taylor JRN, Chikwamba R, Mehlo L (2014) Co-suppression of synthesis of major α -kafirin sub-class together with γ -kafirin-1 and γ -kafirin-2 required for substantially improved protein digestibility in transgenic sorghum. *Plant Cell Rep* 33:521–537. <https://doi.org/10.1007/s00299-013-1556-5>

- Hartwig B, James GV, Konrad K, Schneeberger K, Turck F (2012) Fast isogenic mapping-by-sequencing of ethyl methanesulfonate induced mutant bulks. *Plant Physiol* 160:591–600. <https://doi.org/10.1104/pp.112.200311>
- Hu Z, Olatoye MO, Marla S, Morris GP (2019) An integrated genotyping-by-sequencing polymorphism map for over 10,000 sorghum genotypes. *Plant Genome* 12:18044. <https://doi.org/10.3835/plantgenome2018.06.0044>
- Hu J, Zhang L, Lin W, Tang W, Chan FKL, Ng SC (2021) Review article: probiotics, prebiotics and dietary approaches during COVID-19 pandemic. *Trends Food Sci Technol* 108:187–196. <https://doi.org/10.1016/j.tifs.2020.12.009>
- ICRISAT (2021) Genetic resources. International Crops Research Institute for the Semi-Arid Tropics. <http://exploreit.icrisat.org/profile/genetic%20resources/67>
- Iyabo OO, Ibiyinka O, Abimbola O (2018) Comparative study of nutritional, functional and anti-nutritional properties of white *Sorghum bicolor* (sorghum) and *Pennisetum glaucum* (pearl millet). *Int J Eng Techn Manag Res* 5:151–158. <https://doi.org/10.29121/ijetmr.v5.i3.2018.187>
- Jiang Y, Zhang H, Qi X, Wu G (2020) Structural characterization and antioxidant activity of condensed tannins fractionated from sorghum grain. *J Cereal Sci* 92:102918. <https://doi.org/10.1016/j.jcs.2020.102918>
- Jordan DR, Mace ES, Cruickshank AW, Hunt CH, Henzell RG (2011) Exploring and exploiting genetic variation from unadapted sorghum germplasm in a breeding program. *Crop Sci* 51:1444–1457. <https://doi.org/10.2135/cropsci2010.06.0326>
- Kadri F, Messaoud B, Noureddine B, Bernard W (2017) Free radical scavenging activity correlated with phenolic contents in aqueous methanolic extracts of three Algerian cultivars of sorghum: *Sorghum bicolor* (L.) Moench. *EIWahat J Res Stud* 10:166–180. <https://doi.org/10.54246/1548-010-002-007>
- Khalid W, Ali A, Arshad MS, Afzal F, Akram R, Siddeeg A, Kousar S, Rahim MA, Aziz A, Maqbool Z, Saeed S (2022) Nutrients and bioactive compounds of *Sorghum bicolor* L. used to prepare functional foods: a review on the efficacy against different chronic disorders. *Int J Food Prop* 25:1045–1062. <https://doi.org/10.1080/10942912.2022.2071293>
- Kumar T, Dweikat I, Sato S, Ge Z, Nersesian N, Chen H, Elthon T, Bean S, Ioerger BP, Tilley M, Clemente T (2012) Modulation of kernel storage proteins in grain sorghum (*Sorghum bicolor* (L.) Moench). *Plant Biotechnol J* 10:533–544. <https://doi.org/10.1111/j.1467-7652.2012.00685.x>
- Kumari PK, Umakanth AV, Narsaiah TB, Uma A (2021) Exploring anthocyanins, antioxidant capacity and α -glucosidase inhibition in bran and flour extracts of selected sorghum genotypes. *Food Biosci* 41:100979. <https://doi.org/10.1016/j.fbio.2021.100979>
- Laidlaw H, Mace E, Williams S, Sakrewski K, Mudge A, Prentis P, Jordan D, Godwin I (2010) Allelic variation of the β -, γ - and δ -kafirin genes in diverse sorghum genotypes. *Theor Appl Genet* 121:1227–1237. <https://doi.org/10.1007/s00122-010-1383-9>
- Li A, Jia S, Yobi A, Ge Z, Sato SJ, Zhang C, Angelovic R, Clemente TE, Holding DR (2018) Editing of an alpha-kafirin gene family increases digestibility and protein quality in sorghum. *Plant Physiol* 177:1425–1438. <https://doi.org/10.1104/pp.18.00200>
- Li Z, Zhao X, Zhang X, Liu H (2021a) Bioactive compounds and biological activities of sorghum grains. *Foods* 10:2868. <https://doi.org/10.3390/foods10112868>
- Li M, Xu T, Zheng W, Gao B, Zhu H, Xu R, Deng H, Wang B, Wu Y, Sun X, Zhang Y, Yu LL (2021b) Triacylglycerols compositions, soluble and bound phenolics of red sorghums, and their radical scavenging and anti-inflammatory activities. *Food Chem* 340:128123. <https://doi.org/10.1016/j.foodchem.2020.128123>
- Lin P, Ho J, Bun T, Sai V, Ho M, Xia L (2013) A sorghum xylanase inhibitor-like protein with highly potent antifungal, antitumor and HIV-1 reverse transcriptase inhibitory activities. *Food Chem* 141:2916–2922. <https://doi.org/10.1016/j.foodchem.2013.04.013>
- Links MR, Taylor J, Kruger MC, Taylor JRN (2015) Sorghum condensed tannins encapsulated in kafirin microparticles as a nutraceutical for inhibition of amylases during digestion to attenuate hyperglycaemia. *J Funct Foods* 12:55–63. <https://doi.org/10.1016/j.jff.2014.11.003>

- Liu G, Gilding EK, Kerr ED, Schulz BL, Tabet B, Hamaker BR, Godwin ID (2019) Increasing protein content and digestibility in sorghum grain with a synthetic biology approach. *J Cereal Sci* 85:27–34. <https://doi.org/10.1016/j.jcs.2018.11.001>
- Lozano R, Gazave E, Dos Santos JPR, Stetter MG, Valluru R, Bandillo N, Fernandes SB, Brown PJ, Shakoor N, Mockler TC, Cooper EA, Taylor Perkins M, Buckler ES, Ross-Ibarra J, Gore MA (2021) Comparative evolutionary genetics of deleterious load in sorghum and maize. *Nat Plants* 7:17–24. <https://doi.org/10.1038/s41477-020-00834-5>
- Luo M, Hou F, Dong L, Huang F, Zhang R, Su D (2020) Comparison of microwave and high-pressure processing on bound phenolic composition and antioxidant activities of sorghum hull. *Int J Food Sci Technol* 55:3190–3202. <https://doi.org/10.1111/ijfs.14583>
- Massafaro M, Thompson A, Tuinstra M, Dilkes B, Weil CF (2016) Mapping the increased protein digestibility trait in the high-lysine sorghum mutant P721Q. *Crop Sci* 56:2647–2651. <https://doi.org/10.2135/cropsci2016.03.0188>
- Massel K, Lam Y, Hintzsche J, Lester N, Botella JR, Godwin ID (2022) Endogenous U6 promoters improve CRISPR/Cas9 editing efficiencies in *Sorghum bicolor* and show potential for applications in other cereals. *Plant Cell Rep* 41:489–492. <https://doi.org/10.1007/s00299-021-02816-z>
- Miafo APT, Koubala BB, Kansci G, Muralikrishna G (2019) Free sugars and non-starch polysaccharides–phenolic acid complexes from bran, spent grain and sorghum seeds. *J Cereal Sci* 87:124–131. <https://doi.org/10.1016/j.jcs.2019.02.002>
- Mohapatra D, Patel AS, Kar A, Deshpande SS, Kumar M (2019) Effect of different processing conditions on proximate composition, anti-oxidants, anti-nutrients and amino acid profile of grain sorghum. *Food Chem* 271:129–135. <https://doi.org/10.1016/j.foodchem.2018.07.196>
- Moraes E, da Silva R, Alves S, Vieira VA, Ludemann R, Borck PC, Magalhaes E, Marostica MR (2018) Whole sorghum flour improves glucose tolerance, insulin resistance and preserved pancreatic islets function in obesity diet-induced rats. *J Funct Foods* 45:530–540. <https://doi.org/10.1016/j.jff.2017.03.047>
- Motlhaodi T, Bryngelsson T, Chite S, Fatih M, Ortiz R, Geleta M (2018) Nutritional variation in sorghum [*Sorghum bicolor* (L.) Moench] accessions from southern Africa revealed by protein and mineral composition. *J Cereal Sci* 83:123–129. <https://doi.org/10.1016/j.jcs.2018.08.010>
- Mural RV, Grzybowski M, Miao C, Damke A, Sapkota S, Boyles RE, Fernandez MGS, Schnable PS, Sigmon B, Kresovich S, Schnable JC (2020) Meta-analysis identifies pleiotropic loci controlling phenotypic trade-offs in sorghum. *Genetics* 218:iyab087. <https://doi.org/10.1093/genetics/iyab087>
- NBPGR (2021) National Bureau of Plant Genetic Resources. <http://exploreit.icrisat.org/profile/genetic%20resources/67>
- Ofofu FK, Elahi F, Daliri EB, Tyagi A, Chen XQ, Chelliah R, Kim JH, Han SI, Oh DH (2021) UHPLC-ESI-QTOF-MS/MS characterization, antioxidant and antidiabetic properties of sorghum grains. *Food Chem* 337:127788. <https://doi.org/10.1016/j.foodchem.2020.127788>
- Olatoye MO, Marla SR, Hu Z, Bouchet S, Perumal R, Morris GP (2020) Dissecting adaptive traits with nested association mapping: genetic architecture of inflorescence morphology in sorghum. *G3* 10:1785–1796. <https://doi.org/10.1534/g3.119.400658>
- Omrani M, Keshavarz M, Nejad Ebrahimi S, Mehrabi M, McGaw LJ, Ali Abdalla M, Mehrbod P (2020) Potential natural products against respiratory viruses: a perspective to develop anti-COVID-19 medicines. *Front Pharmacol* 11:586993. <https://doi.org/10.3389/fphar.2020.586993>
- Ongom PO, Ejeta G (2018) Mating design and genetic structure of a multi-parent advanced generation intercross (MAGIC) population of sorghum (*Sorghum bicolor* (L.) Moench). *G3* 8:331–341. <https://doi.org/10.1534/g3.117.300248>
- Oria MP, Hamaker BR, Axtell JD, Huang CP (2000) A highly digestible sorghum mutant cultivar exhibits a unique folded structure of endosperm protein bodies. *Proc Natl Acad Sci U S A* 97:5065–5070. <https://doi.org/10.1073/pnas.080076297>
- Paiva CL, Evangelista WP, Queiroz VA, Gloria MB (2015) Bioactive amines in sorghum: method optimisation and influence of line, tannin and hydric stress. *Food Chem* 173:224–230. <https://doi.org/10.1016/j.foodchem.2014.10.039>
- Przybylska-Balcerak A, Frankowski J, Stuper-Szablewska K (2018) Bioactive compounds in sorghum. *Eur Food Res Technol* 245: 1075–1080. <https://doi.org/10.1007/s00217-018-3207-0>

- Przybylska A, Frankowski J, Stuper K (2019) Bioactive compounds in sorghum. *Eur Food Res Technol* 245:1075–1080. <https://doi.org/10.1007/s00217-018-3207-0>
- Queiroz VAV, da Silva A, Beserra C, de Carvalho CWP, Liboreiro C, Costa P, Pereira RR (2018) A low calorie and nutritive sorghum powdered drink mix: influence of tannin on the sensorial and functional properties. *J Cereal Sci* 79:43–49. <https://doi.org/10.1016/j.jcs.2017.10.001>
- Seyoum Y, Negussie R, Baye K (2016) Nutrient retention and fate of iron-binding phenolic compounds during the Injera processing of tannin-free and high-tannin sorghum. *J Sci Food Agric* 96:1541–1547. <https://doi.org/10.1002/jsfa.7246>
- Shen S, Huang R, Li C, Wu W, Chen H, Shi J, Chen S, Ye X (2018) Phenolic compositions and antioxidant activities differ significantly among sorghum grains with different applications. *Molecules* 23:1203–1215. <https://doi.org/10.3390/molecules23051203>
- Song G, Jia M, Chen K, Kong X, Khattak B, Xie C, Li A, Mao L (2016) CRISPR/Cas9: a powerful tool for crop genome editing. *Crop J* 4:75–82. <https://doi.org/10.1016/j.cj.2015.12.002>
- Szerszunowicz I, Klobukowski J (2020) Characteristics of potential protein nutraceuticals of plant origin with antioxidant activity. *Molecules* 25:1621. <https://doi.org/10.3390/molecules25071621>
- Upadhyaya HD, Vetriventhan M, Deshpande S (2016) Sorghum germplasm resources characterization and trait mapping. In: Rakshit S, Wang Y (eds) *The sorghum genome*. Springer, pp 77–94
- USDA (2021) Agricultural Research Service, National Plant Germplasm System, Germplasm Resources Information Network GRIN-Taxonomy. National Germplasm Resources Laboratory. <https://npgsweb.ars-grin.gov/gringlobal/taxon/taxonomysimple>
- Von Born P, Bernardo-Faura M, Rubio-Somoza I (2018) An artificial miRNA system reveals that relative contribution of translational inhibition to miRNA-mediated regulation depends on environmental and developmental factors in *Arabidopsis thaliana*. *PLoS ONE* 13: e0192984. <https://doi.org/10.1371/journal.pone.0192984>
- Wang L, Lu Z, Regulski M, Jiao Y, Chen J, Ware D, Xin Z (2021) BSAseq: an interactive and integrated web-based workflow for identification of causal mutations in bulked F2 populations. *Bioinformatics* 37: 382–387. <https://doi.org/10.1093/bioinformatics/btaa709>
- Winn JA, Mason RE, Robbins AL, Rooney WL, Hays DB (2009) QTL mapping of a high protein digestibility trait in *Sorghum bicolor*. *Int J Plant Genomics* 2009:471853. <https://doi.org/10.1155/2009/471853>
- Wongwaiwech D, Weerawatanakorn M, Boonnoun P (2020) Subcritical dimethyl ether extraction as a simple method to extract nutraceuticals from byproducts from rice bran oil manufacture. *Sci Rep* 10:21007. <https://doi.org/10.1038/s41598-020-78011-z>
- Wong JH, Lau T, Cai N, Singh J, Pedersen JF, William, Vensel H, Hurkman WJ, Wilson JD, Lemaux PG, Buchanan BB (2009) Digestibility of protein and starch from sorghum (*Sorghum bicolor*) is linked to biochemical and structural features of grain endosperm. *J Cereal Sci* 49:73–82. <https://doi.org/10.1016/j.jcs.2008.07.013>
- Wu Y, Yuan L, Guo X, Holding DR, Messing J (2013) Mutation in the seed storage protein kafirin creates a high-value food trait in sorghum. *Nat Commun* 4:2217. <https://doi.org/10.1038/ncomms3217>
- Wu G, Bornman JF, Bennett SJ, Clarke MW, Fang Z, Johnson SK (2017) Individual polyphenolic profiles and antioxidant activity in sorghum grains are influenced by very low and high solar UV radiation and genotype. *J Cereal Sci* 77:17–23. <https://doi.org/10.1016/j.jcs.2017.07.014>
- Xin Z, Wang ML, Barkley NA, Burow G, Franks C, Pederson G, Burke J (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. *BMC Plant Biol* 8:103. <https://doi.org/10.1186/1471-2229-8-103>
- Xin Z, Wang M, Cuevas HE, Chen J, Harrison M, Pugh NA, Morris G (2021) Sorghum genetic, genomic, and breeding resources. *Planta* 254:114. <https://doi.org/10.1007/s00425-021-03742-w>
- Yu J, Yan F, Lu Q, Liu R (2018) Interaction between sorghum procyanidin tetramers and the catalytic region of glucosyltransferases-I from *Streptococcus mutans* UA159. *Food Res Int* 112: 152–159. <https://doi.org/10.1016/j.foodres.2018.06.027>
- Zhu H, Li C, Gao C (2020) Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nat Rev Mol Cell Biol* 21:661–677. <https://doi.org/10.1038/s41580-020-00288-9>