From the Institute of Agronomy and Plant Breeding I

Professorship for Plant Breeding

of the Justus Liebig University of Giessen

ELUCIDATION OF THE GENETIC REGULATION OF VICINE AND CONVICINE IN FABA BEAN (*VICIA FABA* L.)

DISSERTATION

for the award of the doctorate degree (Dr. agr.)

in the Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management at the Justus Liebig University of Giessen

Submitted by

Samson Ugwuanyi

from Ikpamodo, Nigeria

Giessen, 2025

With the permission of the Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management of the
Justus Liebig University Giessen

Board of Examiners:

1. reviewer: Prof. Dr. Rod J. Snowdon

2nd reviewer: Prof. Dr. Sarah Schiessl-Weidenweber

Examiner: Prof. Dr. Andreas Stahl

Examiner: Prof. Dr. John Clifton-Brown

Chairperson: Prof. Dr. Jan Siemens

Date of the Disputation: 20/03/2026

Table of Contents

1 Introduction

1.1 General introduction of faba bean (*Vicia faba L.*)

Faba bean (*Vicia faba* L.), also known as broad bean, is a globally important grain legume with high agronomic and nutritional value. Recognized for its high yield potential, it plays a key role in diversifying cropping systems and providing a rich source of vegetable protein, particularly in temperate climates (Khazaei et al., 2017). Its long history of cultivation dates back to the Neolithic era, making it one of the earliest domesticated food legumes. Archaeological evidence from the southern Levant suggests its seeds were in use as early as 14,000 years ago (Caracuta et al., 2016). The crop is believed to have originated in the Near East and Mediterranean regions, from where it spread across the globe (Duc et al., 2010).

Faba bean is widely cultivated across multiple continents due to its adaptability to diverse climatic conditions, particularly in cool-season environments such as Mediterranean winters and northern European summers, where other protein crops perform poorly (Duc et al., 2010). *V. faba* is rich in protein and other valuable nutrients, making it a good source of food for both humans and animals. Beyond its role as a high-protein food source, faba bean significantly contributes to sustainable agriculture. Its strong nitrogen-fixing ability enhances soil fertility, making it a valuable component in mixed cropping and crop rotation systems (Jensen et al., 2010). The United Nations' Intergovernmental Panel on Climate Change (IPCC) has emphasized that adopting plant-based diets can substantially reduce carbon emissions, making the broader cultivation of locally adapted, protein-rich crops like faba bean essential in reaching this goal (IPCC, 2019). Despite its many potentials, faba bean contains several antinutritional compounds, including vicine and convicine (v-c), that hinder its widespread acceptance and use (Hendawey and Younes, 2013; Multari et al., 2015). Addressing these antinutritional factors is therefore crucial to increasing the crop's consumption and utilization, thereby maximizing its potential as a sustainable food and feed source.

1.2 Global production of faba bean

Faba bean ranks as the sixth most widely cultivated pulse worldwide. Despite its agronomic advantages, it occupies only 1.2% of the 200 million hectares allocated to annual grain legume production, placing it seventh among major legumes, behind soybean (*Glycine max* L.), groundnut (*Arachis hypogaea* L.) and common bean (*Phaseolus vulgaris* L.) (Adhikari et al., 2021).

In Europe, faba bean is the second most cultivated grain legume behind pea, covering about one-quarter of the 1.6 million hectares under legume production (Duc et al., 2015). Notably, the

area under faba bean cultivation in Europe has more than doubled over the past decade, driven by growing recognition of its value as a sustainable domestic protein source for animal feed and for its ecological benefits in cropping systems. Globally, the total area under faba bean cultivation has declined significantly, from 5.4 million hectares in 1961 to 2.8 million hectares in 2023 (Figure 1). However, during this period, productivity has improved significantly, increasing from 0.9 t/ha to 2.1 t/ha (FAOSTAT, 2025). In recent years, global production has also risen, reaching an all-time high of approximately 6.2 million tonnes in 2022, representing a more than 25% increase over the past decade (Figure 2).

Europe is currently the leading faba bean producer, a position it has held for the past four years, contributing around 30% of global production. Asia and Africa follow, with 29% and 26% shares, respectively (FAOSTAT, 2025). The top producing countries include China, Ethiopia, the United Kingdom, Australia, Germany, France, Sudan, Egypt, Italy, and Morocco (Figure 3).

Despite recent production increases, the long-term decline in cultivation area is largely attributed to competition from other crops, yield instability caused by biotic and abiotic stresses, and the persistence of antinutritional factors, which continue to limit its broader utilization (Adhikari et al., 2021).

1.3 Nutritional and ecological significance of faba bean

Faba bean is a valuable protein source for both human and animal nutrition, containing between 25% and 37% protein, with nearly half of its grain composed of starch (Duc et al., 2010; Ali et al., 2016; Warsame et al., 2020). It is also rich in essential minerals such as potassium, calcium, magnesium, iron, and zinc, which further enhances its potential for global human nutrition (Lizarazo et al., 2015; Longobardi et al., 2015; Neme et al., 2015; Karkanis et al., 2018). The crop is particularly significant in Africa, Asia, and Mediterranean countries, where it serves as a staple protein source for humans. In Europe, especially Germany, faba bean is increasingly promoted as a sustainable alternative to imported plant proteins, contributing to improvement in soil fertility and reducing reliance on synthetic fertilizers (Köpke and Nemecek, 2010; Link et al., 2010). The feed industry is another major consumer in this region, with dry faba bean seeds widely used in livestock diets, particularly for pigs and poultry, while its straw and silage occasionally supplement ruminant feed (Duc et al., 2010).

In addition to its nutritional value, faba bean plays a critical role in sustainable agriculture due to its exceptional nitrogen-fixing ability. The crop can fix up to 200 kg N ha^{-1} through symbiotic nitrogen fixation, making it one of the most efficient nitrogen-fixing legumes (Herridge et al.,

2008; Liu et al., 2019; Karkanis et al., 2018; Desoky et al., 2021). This process reduces the need for synthetic nitrogen fertilizers in subsequent cropping seasons, enhancing soil health while minimizing environmental impact (Adhikari et al., 2021). This feature makes faba bean an integral component of various agricultural systems, where it is incorporated to leverage its nitrogen-fixing potential in crop production. Notably, crop rotation and intercropping systems commonly integrate faba bean to maximize these nitrogen-fixing benefits (Jensen et al., 2010; Karkanis et al., 2018). In Mediterranean agriculture, for instance, rotation with faba bean supplies over three-quarters of the required nitrogen for subsequent crops (Denton et al., 2017).

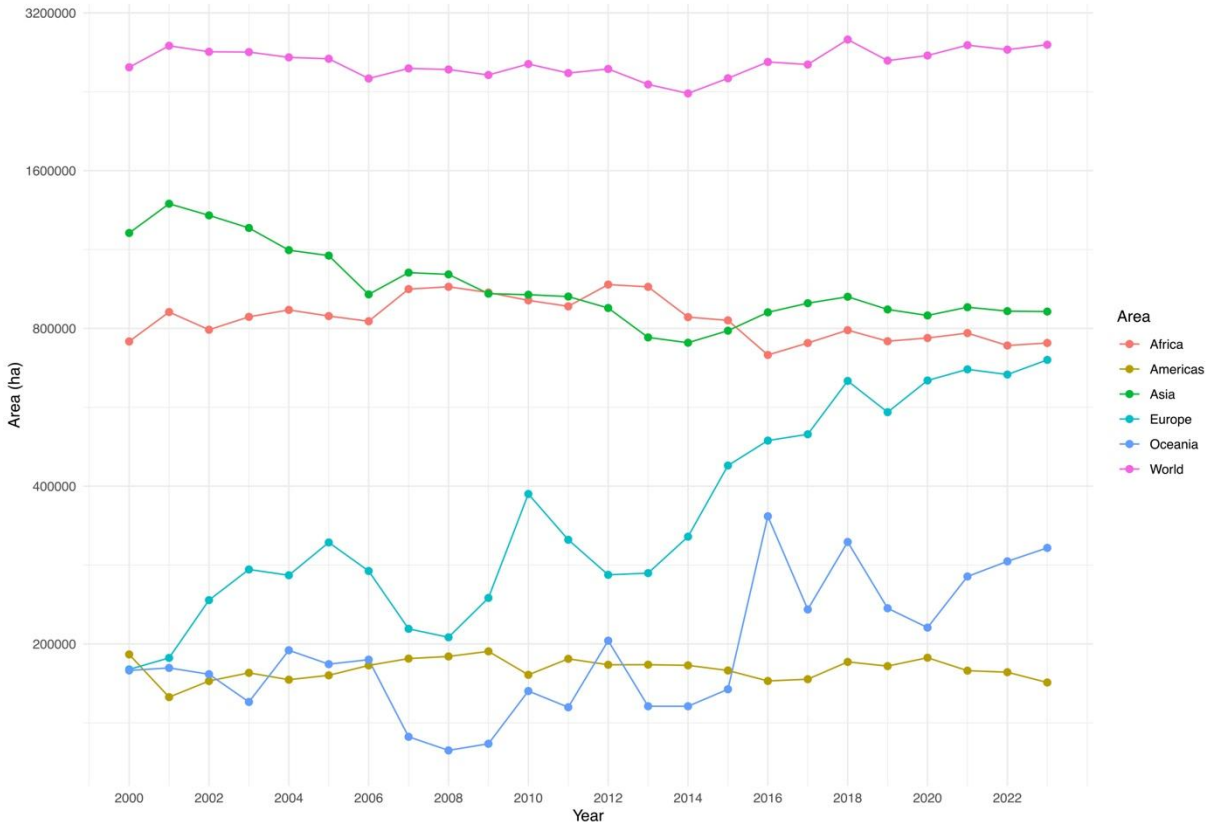


Figure 1: Global cropping area of faba bean. Total harvested area (in hectares) of faba bean across the world from 2000 to 2023.

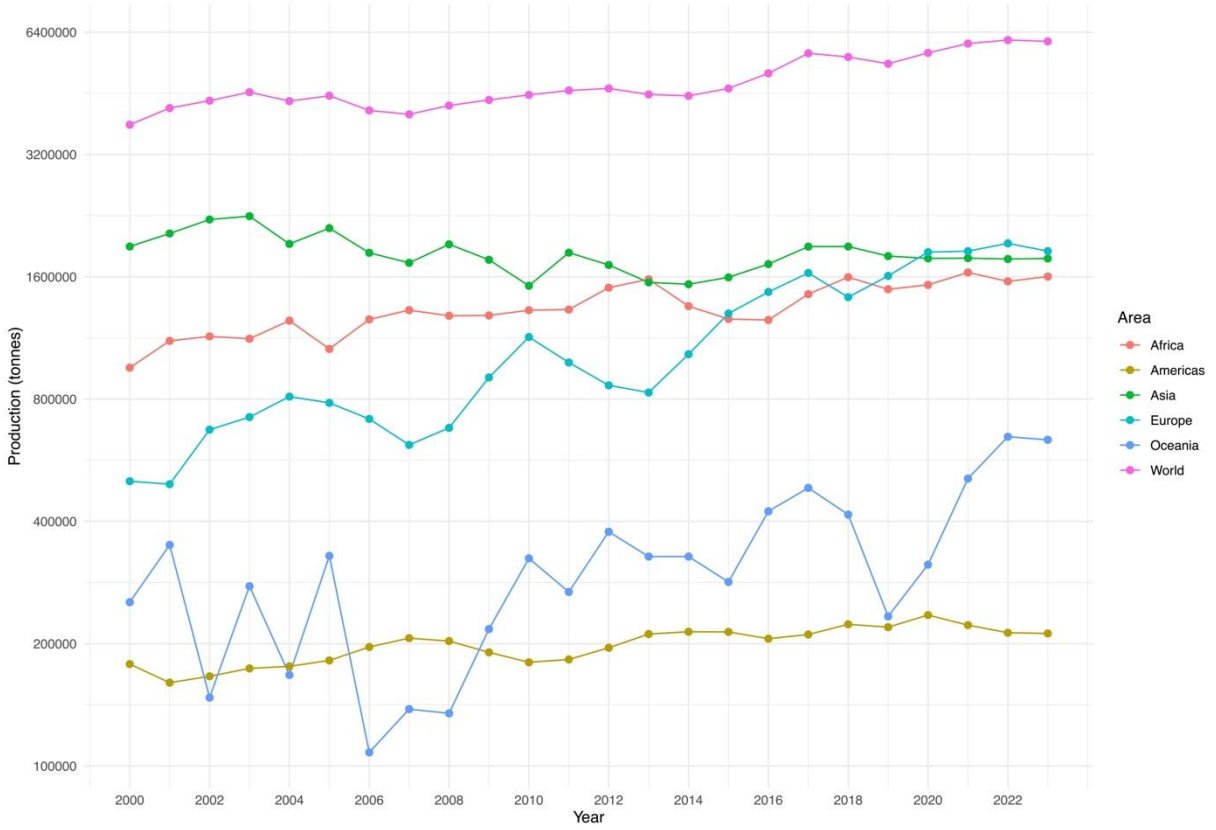


Figure 2: Global production of faba bean. Total grain yield (in tonnes) of faba bean across the world from 2000 to 2023.

1.4 Antinutritional factors in faba bean

While faba bean offers significant nutritional benefits, it also contains several antinutritional compounds. Legumes are known to contain bioactive compounds traditionally classified as antinutritional factors (ANFs) (Gemede, 2014). Some of the antinutrients present in faba bean include vicine, convicine, phenolic compounds, phytic acid, oligosaccharides, protease inhibitors, saponins, lectins and oxalates (Hendawey and Younes, 2013; Multari et al., 2015). These naturally occurring compounds in plants can hinder the absorption of nutrients in the human body when consumed (Gemede, 2014). Their impact varies depending on factors such as plant source, chemical composition, concentration and dietary habits. Some interfere with nutrient absorption, while others pose direct health risks, causing adverse effects in sensitive individuals (Gulewicz et al., 2014; Shi et al., 2024). For example, plant-based diets rich in grain legumes often show low mineral bioavailability, particularly for calcium, iron, and zinc, due to absorption inhibitors such as phytates and polyphenols (Sandberg, 2002; Tako et al., 2014), while in faba bean, vicine-convicine, the most problematic antinutritional factors, are associated with favism, a severe haemolytic condition in individuals with glucose-6-phosphate dehydrogenase (G6PD) deficiency (Khazaei et al., 2019). Interestingly, despite their

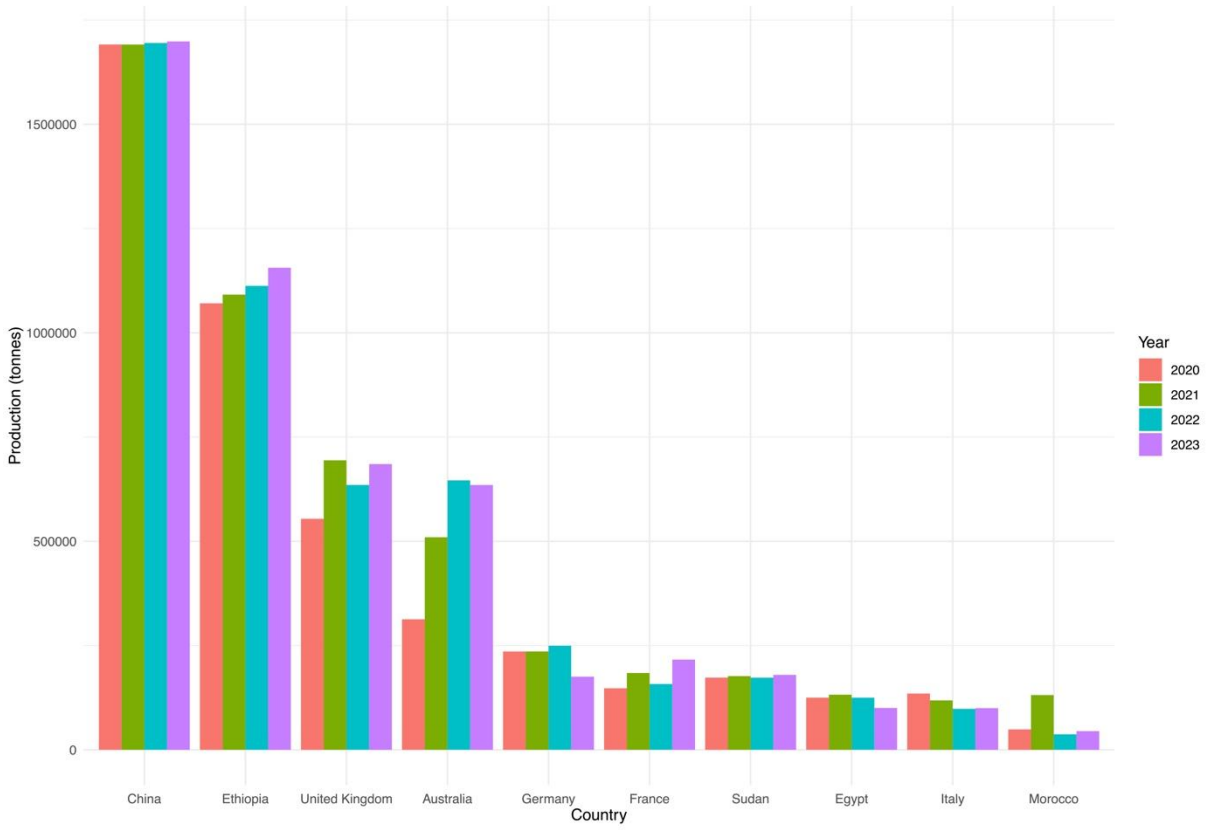


Figure 3: Top faba bean producers. Total grain yield (in tonnes) across the top 10 producing countries over the past four years.

antinutritional properties, some ANFs have also been recognized for their bioactive potential, offering health benefits that may help prevent chronic diseases (Shi et al., 2024).

1.4.1 Vicine and convicine

Vicine and convicine are two structurally related pyrimidine glycosides; while vicine contains an amino group, convicine has a hydroxy group. The ratio of vicine to convicine in faba bean seeds is approximately 2:1 and both are present in all parts of the plant (Goyoaga et al., 2008). These compounds, which constitute 0.3% to 1.5% of the faba bean seed dry matter, are almost unique to the genus Vicia, with bitter melon (*Momordica charantia*) being the only non-Vicia species known to synthesize vicine (Goyoaga et al., 2008; Gauttam and Kalia, 2013; Khazaei et al., 2019). Vicine and convicine are the primary antinutritional compounds in faba beans. They are responsible for haemolytic anemia (favism), a condition that significantly limits the broader use of faba beans (Luzzatto and Arese, 2018).

1.4.2 Phenolic compounds

Phenolic compounds are water-soluble molecules with hydroxyl groups attached to aromatic rings. They act as antioxidants by chelating metal ions, offering potential protection against diseases linked to oxidative stress, such as cancer, diabetes and inflammation (Dai and Mumper,

2010; Pereira et al., 2009; Turco et al., 2016). However, their ability to bind proteins reduces protein digestibility and nutrient availability (Bai et al., 2018). Dark-colored faba bean varieties contain high levels of condensed tannins (CTs), which make up 70-80% of the total phenolic content (TPC) in the seed coat (Nasar-Abbas et al., 2009; Sharan et al., 2021). Tannins, a type of polymeric flavonoid, form insoluble complexes with proteins and enzymes, and may also limit vitamins A and B12 absorption (Gulewicz et al., 2014; Wang et al., 1998). Studies show that TPC and tannin levels vary significantly by genotype and growing location, with low-tannin cultivars showing reduced concentrations (Oomah et al., 2011; Chaieb et al., 2011; Martineau-Côté et al., 2023), while zero-tannin cultivars also exist (Gutierrez et al., 2008; Crépon et al., 2010; Woyengo and Nyachoti, 2012).

1.4.3 Phytic acid

Phytic acid, or phytate, is the main phosphorus reserve in plant seeds and grains, comprising 60-90% of total phosphorus in cereals and legumes. It strongly chelates minerals like iron, zinc, calcium, magnesium, and copper, reducing their bioavailability since humans lack phytase to release them. Phytates hinder mineral absorption by forming insoluble, indigestible complexes in the gastrointestinal tract (Sandstrom, 1997). When consumed in excess, they can reduce zinc and iron uptake, leading to deficiencies, especially in legume-heavy diets (Labba et al., 2021). Phytic acid also binds proteins, impairing enzyme function and nutrient absorption (Shi et al., 2024). Among legumes, faba beans contain 21.7 mg/g db of phytic acid, higher than kidney beans (15.9 mg/g db), while chickpea and field pea contain 32.0 and 12.0 mg/g db, respectively. Studies report average faba bean phytic acid levels around 17.5 mg/g db, though content varies significantly by genotype and growing region (Shi et al., 2024).

1.4.4 Oligosaccharides

Oligosaccharides are indigestible, fermentable carbohydrates that act as dietary fiber and prebiotics. However, their fermentation in the colon produces gases (CO_2, H_2, CH_4), causing flatulence, especially in individuals with intestinal disorders. Because of this, they are considered ANFs (Shi et al., 2024). In legumes, the main oligosaccharides are raffinose, stachyose, and verbascose, members of the raffinose family oligosaccharides (RFOs), formed by sequential galactosyl additions to sucrose. In faba beans, reported RFO contents are 8.67-8.79 mg/g (raffinose), 20.38-25.53 mg/g (stachyose), and 37.54-64.00 mg/g (verbascose). Levels are generally consistent across cultivars but vary with environment. Raffinose is usually the least abundant, as it is synthesized earlier in seed development and later converted into stachyose and verbascose during maturation (Shi et al., 2024).

1.4.5 Protease inhibitors

Protein digestibility in legumes is strongly affected by protease inhibitors, which impair protein breakdown and can consequently induce pancreatic hypertrophy (Sharma and Sehgal, 1992). Faba beans contain protease inhibitors that mainly target trypsin and chymotrypsin, binding to specific amino acid residues of these enzymes and reducing digestion efficiency. In faba beans, trypsin inhibitory activity (TIA) is lower than in chickpeas and soybeans but higher than in peas and lentils, while chymotrypsin inhibitory activity (CIA) is among the lowest. Reported values for raw faba beans are 5.96 TIU/mg (TIA) and 1.1 CIU/mg (CIA). Split beans show higher activity than whole beans because inhibitors are concentrated in the cotyledons (Shi et al., 2024).

1.4.6 Saponins

Saponins are heat-stable, water-soluble compounds with surface activity that impart bitterness and reduce the bioavailability of some minerals and vitamins. They are reported to lower plasma cholesterol in humans (Singh et al., 2017) and may contribute to cancer prevention according to epidemiological studies. However, their bitterness can make them undesirable in food products (Shi et al., 2004; Labba et al., 2021). In faba beans, total saponin content is approximately 757 μg/g in whole flour and 480 μg/g in split flour, levels considerably lower than those in peas, chickpeas and lentils (1367-7572 μg/g) (Shi et al., 2024).

1.4.7 Lectins

Lectins are sugar- and protein-binding compounds that can agglutinate red blood cells and impair nutrient absorption in the intestine. In faba beans, lectin activity is around 5.5 HU/mg (db) in both whole and split seeds, comparable to peas but lower than in beans, lentils and soybeans (Shi et al., 2024).

1.4.8 Oxalates

Oxalates are organic acids that chelate minerals such as calcium, lowering their bioavailability and, at high intake, promoting calcium oxalate crystal formation associated with kidney stones. Oxalate levels are around 241.5 mg/100 g (db) in whole faba bean flour and 291.4 mg/100 g (db) in split flour, with the higher values in split beans reflecting greater concentrations in the cotyledons (Shi et al., 2024).

1.5 Processing methods for reducing the negative effects of ANFs in faba beans

While the biosynthesis of some antinutrients is well understood, enabling the development of faba bean cultivars with reduced or eliminated content (Gutierrez et al., 2008; Woyengo and Nyachoti, 2012), complete removal of others remains a challenge. Nonetheless, several

processing techniques can significantly reduce or eliminate ANFs, thereby mitigating their adverse effects on human health (see Table 1). Depending on their stability, some being heat-sensitive (thermolabile) and others heat-resistant (thermostable), common processing methods include cooking, dehulling, soaking, autoclaving, germination, extrusion, fermentation, enzymes treatment and roasting (Luo and Xie, 2013; Patterson et al., 2017; Shi et al., 2017).

1.5.1 Cooking

Faba beans are commonly cooked for long periods to improve softness, palatability and protein digestibility (Bakr, 1996; Hamza et al., 1987). Cooking decreases tannins, total phenolics, phytic acid, v-c, trypsin inhibitors, α-amylase inhibitors, hemagglutinin activity, oligosaccharides, lectins and oxalates (Stone et al., 2021; Shi et al., 2018; Khalil, 1995). It also improves in vitro protein digestibility and reduces bitterness from phenolics and v-c (Shi et al., 2024).

1.5.2 Dehulling

Dehulling removes the seed coat, which contains most tannins and phenolic compounds, thereby reducing their levels and improving sensory and protein quality (Alonso et al., 2000; Fernando, 2021; Dhull et al., 2022). Removing the hull increases protein and amino acid content and improves in vitro protein and starch digestibility, likely due to partial loss of phytic acid, condensed tannins, and polyphenols (Dhull et al., 2022). However, phytic acid may increase in the remaining cotyledon fraction due to its localization in protein bodies (Luo and Xie, 2013). Dehulling has minimal effect on oligosaccharides, v-c or protease inhibitors.

1.5.3 Soaking

Soaking faba beans in water (or alkaline solutions) reduces soluble antinutrients such as phytic acid, trypsin inhibitors, oligosaccharides (raffinose, stachyose, verbascose), v-c, oxalates, tannins and lectins (Jamalian, 1999; Luo and Xie, 2013; Shi et al., 2017; Shi et al., 2018; Stone et al., 2021). It can also slightly increase protein content (Anderson et al., 1994; Bakr, 1996; Alonso et al., 2000; Dhull et al., 2022). Soaking leaches some phenolic compounds and saponins, depending on conditions (Stone et al., 2021; Sharma and Sehgal, 1992; Shi et al., 2024). Vicine and convicine levels can be reduced further by soaking in water or mild acid solutions before cooking (Jamalian and Ghorbani, 2005). Water absorption efficiency depends on factors such as temperature, seed size and sodium bicarbonate concentration (Kader, 1995).

1.5.4 Autoclaving / high-pressure heat treatments

Autoclaving or high-pressure heat processing effectively reduces antinutritional compounds such as trypsin inhibitors, chymotrypsin inhibitors, tannins, total phenolics, v-c, hemagglutinin

and lectins, while improving in vitro protein digestibility and protein efficiency ratio (Sharma and Sehgal, 1992; Hall and Moraru, 2021; Khalil, 1995; Osman et al., 2014; Shi et al., 2018; Dhull et al., 2022). Autoclaving for longer periods enhances protein digestibility more effectively. Certain amino acids (e.g., leucine, tyrosine, threonine, histidine) may increase after treatment (Sharma and Sehgal, 1992). Boiling and dehulling also reduce v-c substantially (Cardador-Martínez et al., 2012).

1.5.5 Germination

Germination enhances the nutritional quality of faba beans while reducing antinutrients. It slightly decreases starch (~15%) but significantly reduces phytic acid (~45%) and α-galactosides (~94%), improves dietary fiber, and increases calcium bioavailability due to reduced phytic acid and hemicellulose (Alonso et al., 2000; Hamza et al., 1987; Vidal-Valverde et al., 1998). Protein content and in vitro starch digestibility increase during 24-72 h of germination as starch is utilized for energy. Trypsin, chymotrypsin and α-amylase inhibitors, as well as polyphenols and tannins, are reduced, though hemagglutinating activity remains largely unchanged. Folate content also rises significantly (>40%) (Beleia et al., 1993; Hefni et al., 2015). Germination also decreases v-c levels (Coda et al., 2015; Dhull et al., 2022).

1.5.6 Extrusion

Extrusion processing improves digestibility and eliminates several antinutritional factors without significantly affecting protein, lipids or ash content (Adamidou et al., 2011; Nosworthy et al., 2018). Preconditioning and high-temperature extrusion reduce nonstarch polysaccharides, gelatinize starch and eliminate trypsin, chymotrypsin, α-amylase inhibitors and hemagglutinating activity. Phytic acid and tannins are also reduced, though mineral bioavailability may not improve substantially (Francis et al., 2001; Adamidou et al., 2011). Extruded faba bean flour retains higher protein and fiber than other legume flours, making it useful for plant-based protein products (Gu et al., 2020; Liu et al., 2022; Osman et al., 2014; Rosa-Sibakov et al., 2016; Dhull et al., 2022).

1.5.7 Fermentation

Microbial fermentation (e.g., lactic acid bacteria or *Rhizopus oligosporus*) reduces tannins, phenolics, v-c, oligosaccharides and protease inhibitors, sometimes achieving complete degradation in certain cultivars (Verni et al., 2019; Polanowska et al., 2020; Coda et al., 2015).

1.5.8 Wet extraction

Wet extraction for protein isolates completely removes v-c and significantly reduces oligosaccharides and protease inhibitors, while retaining protein content (Vogelsang-O'Dwyer et al., 2020). However, phytic acid may remain concentrated in isolates (Shi et al., 2022).

1.5.9 Roasting

Roasting or dry heating is traditionally used to produce faba bean flour. Higher roasting temperatures preserve protein content better, while low temperatures may reduce it (Dhull et al., 2022). Roasting effectively decreases trypsin inhibitors, phytic acid, α-galactosides, tannins and some phenolics. Roasting effectively decreases trypsin inhibitors, phytic acid, α-galactosides, tannins, and some phenolics. Prolonged roasting at 150 °C can generate new phenolic compounds that enhance antioxidant capacity, though initial antioxidant activity decreases (Anderson et al., 1994; Vidal-Valverde et al., 1998; Siah et al., 2014).

1.6 Vicine and convicine as major antinutritional factors limiting faba bean utilization

Faba bean has had a complex history, particularly in its center of diversity, the Mediterranean region. Despite being one of the earliest domesticated crops in Near Eastern agriculture, it was considered undesirable in some ancient cultures, leading to periods where its cultivation was banned (Caracuta et al., 2016). It is recorded that the Greek philosopher Pythagoras famously forbade his followers from consuming faba beans, and this aversion reportedly caused his life to end tragically when he refused to cross a faba bean field while fleeing pursuers (Meletis and Konstantopoulos, 2004; Khazaei et al., 2019). This historical aversion to faba bean is now understood to be linked to the presence of pyrimidine glycosides, vicine and convicine, which can induce a severe haemolytic response in certain individuals.

Vicine and convicine are the primary antinutritional compounds limiting the use of faba bean in food and feed. Their aglycones are potent oxidants that can trigger acute haemolytic anaemia, also known as favism, in individuals with glucose-6-phosphate dehydrogenase deficiency (Duc et al. 1989; Crépon et al. 2010; Khamassi et al., 2013). This genetic condition, the most common enzyme deficiency worldwide and more prevalent in regions historically affected by malaria, makes red blood cells vulnerable to oxidative damage, potentially leading to life-threatening complications. Globally, over 400 million people are at risk of favism when consuming faba bean with high v-c content, due to the approximately 4% global prevalence of G6PD deficiency. This condition is more common and severe in men because of its X-linked inheritance pattern

Table 1: Processing methods for reducing antinutritional factors in faba beans.

Processing method	Affected ANFs	Comments / references
Cooking/ boiling	Tannins, total phenolics, phytic acid, vicine, convicine, trypsin and α-amylase inhibitors, hemagglutinin, oligosaccharides, lectins, oxalates	Improves protein digestibility; reduces bitterness; some studies report minimal effect on phytic acid (Alonso et al., 2000; Osman et al., 2014; Cardador-Martínez et al., 2012; Stone et al., 2021; Shi et al., 2018; Khalil, 1995; Martineau-Côté et al., 2024)
Dehulling	Tannins, polyphenols; phytic acid	Effective due to hull removal; improves protein quality; may increase phytic acid in cotyledons (Dhull et al., 2022; Fernando, 2021; Luo and Xie, 2013)
Soaking	Phytic acid, trypsin inhibitors, oligosaccharides (RFOs), vicine, convicine, oxalates, tannins, lectins, some phenolics, saponins	Leaching of soluble compounds; alkaline solutions more effective for vicine/convicine (Shi et al., 2018; Stone et al., 2021; Jamalian, 1999; Sharma and Sehgal, 1992)
Autoclaving	Trypsin inhibitors, chymotrypsin inhibitors, tannins, total phenolics, vicine, convicine, hemagglutinin, lectins	Improves digestibility while maintaining protein content (Sharma and Sehgal, 1992; Hall and Moraru, 2021; Shi et al., 2018; Khalil, 1995)
Germination	Phytic acid, tannins, polyphenols, α-galactosides, oligosaccharides (RFOs), protease inhibitors, vicine, convicine	Improves protein/starch digestibility, dietary fiber, calcium bioavailability, and folate; hemagglutinin largely unchanged (Alonso et al., 2000; Wei et al., 2022; Beleia et al., 1993; Coda et al., 2015; Hefni et al., 2015; Dhull et al., 2022)
Extrusion	Trypsin inhibitors, chymotrypsin inhibitors, α-amylase inhibitors, hemagglutinin, tannins, total phenolics, phytic acid, oligosaccharides	Starch gelatinization improves digestibility; protein retained (El-Hady and Habiba, 2003; Alonso et al., 2000; Adamidou et al., 2011; Nosworthy et al., 2018; Gu et al., 2020)
Fermentation	Tannins, total phenolics, vicine, convicine, oligosaccharides, protease inhibitors	Can achieve complete degradation in some cultivars; improves digestibility (Verni et al., 2019; Polanowska et al., 2020; Coda et al., 2015)

Processing method	Affected ANFs	Comments / references
Wet extraction /protein isolation	Vicine, convicine, oligosaccharides, protease inhibitors	ANFs significantly reduced; phytic acid may concentrate; protein retained (Vogelsang-O'Dwyer et al., 2020; Shi et al., 2022)
Roasting	Trypsin inhibitors, phytic acid, α-galactosides, tannins, phenolics	High temperatures can generate new antioxidant phenolics; partial protein degradation to smaller subunits (Anderson et al., 1994; Vidal-Valverde et al., 1998; Siah et al., 2014)

(Nkhoma et al., 2009). Favism occurs because these compounds are metabolized into divicine and isouramil, which oxidize intracellular glutathione. G6PD-deficient individuals struggle to regenerate oxidized glutathione due to their reliance on the pentose phosphate pathway for nicotinamide adenine dinucleotide phosphate (NADPH), leading to oxidative damage and red blood cell aggregation (Winterbourn, 1989; Khazaei et al., 2019). Interestingly, G6PD deficiency has an economic significance as it offers some protection against malaria. The malaria parasite increases oxidation in mutant red blood cells, which the immune system then identifies and removes before the parasite can reproduce, reducing the severity of the disease (Cappellini and Fiorelli, 2008).

Since no treatment exists for favism, prevention relies on early diagnosis of G6PD deficiency and strict avoidance of foods rich in v-c or the use of oxidative drugs (Arese et al., 2012). An alternative approach to mitigate favism risk involves developing v-c-free faba bean cultivars or breeding cultivars with reduced v-c content, as studies have confirmed that consuming low v-c faba bean is safe for G6PD-deficient individuals (Gallo et al., 2018). The presence of v-c also has dietary disadvantages when used as feed for livestock. In monogastric animals for example, they reduce feed conversion efficiency and can cause erythrocyte rupture, particularly in broiler chickens, while lowering productivity in egg-laying hens (Crepon et al. 2010; Khazaei et al., 2015).

1.7 Elimination or reduction of vicine and convicine content: A key objective in faba bean breeding

The need to eliminate or reduce v-c content has been a major breeding goal to enhance faba bean's suitability for food and feed. Initial attempts to identify natural low v-c accessions were largely unsuccessful. However, after years of concerted effort, Duc et al. (1989) successfully

identified the first low v-c genotype (Line 1268) by screening over 900 accessions from the Department of Plant Genetic Resources in Radzików, Poland. This low v-c trait also designated "vc-" and associated with near-zero v-c content, follows a recessive Mendelian inheritance pattern (Khazaei et al., 2019). Despite this knowledge and apparently simple inheritance, eliminating v-c from faba bean remains a challenge, as the complete elucidation of the more complex genetic mechanisms underlying their biosynthesis are not yet fully understood. This has also hindered the establishment of efficient molecular breeding tools for marker-assisted selection (MAS). This limitation has slowed breeding progress and increased the cost of developing low v-c cultivars.

1.8 Advances towards understanding the genetic control of vicine and convicine biosynthesis

Vicia faba is a partially outcrossing diploid legume with six pairs of exceptionally large chromosomes. Its haploid genome is approximately 13 giga base pairs (Gbp), three times larger than lentil and pea genomes and 16 times the size of the chickpea genome (Khazaei et al., 2021). Within diploid field crops, it ranks among the largest, with over 85% of its genome consisting of repetitive sequences (Soltis et al., 2003; Novak et al., 2020). This enormous genome size has historically hindered the rapid advancement of genomic tools, thereby complicating genetic studies of key agronomic traits.

Early genetic studies of v-c content were challenged by the absence of a well-annotated reference genome for this crop which implied that candidate gene search relied on comparative genomics approach with other closely related legume species. The biosynthesis of v-c is reported to be a qualitative trait regulated by a few genes (Duc et al., 1989; Ramsay et al., 1995; Gutierrez et al., 2006; Khazaei et al., 2015). Initial mapping efforts utilized polymerase chain reaction (PCR)-based markers such as random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), and isozymes to analyze v-c levels in BCF2 backcross and F6 recombinant inbred line (RIL) populations of faba bean (Ramsay et al., 1995). This study identified three genomic regions associated with low v-c content, including one major locus and two minor quantitative trait loci (QTL). Subsequent study by Gutiérrez et al. (2006) further confirmed the association of two additional RAPD markers to the low v-c trait in an F2 segregating population.

Building on these findings, Khazaei et al. (2015) employed a transcript-based approach to develop single nucleotide polymorphism (SNP) markers, enabling comparative mapping between *Medicago truncatula* and *V. faba* in an F5 RIL population derived from a cross between

the cultivars Mélodie/2 and ILB 938/2. This study reinforced the bimodal inheritance of v-c content and identified a major QTL at the previously reported "vc-" locus on chromosome 1, which accounted for 76% of phenotypic variation. This major effect QTL was later mapped to the distal end of faba bean chromosome 1 (Ramsay et al., 1995; Gutiérrez et al., 2006) and narrowed down to an interval flanked by molecular markers (Khazaei et al., 2015).

Furthermore, Ray et al. (2015) identified six candidate sequences potentially involved in the v-c biosynthetic pathway. Notably, BLAST analysis of the contig 4518 suggested homology to Medtr2g009270, a gene on chromosome 2 of *M. truncatula* encoding 3,4-dihydroxy-2-butanone 4-phosphate synthase (DHBPS), a component of the bifunctional riboflavin biosynthesis protein *RIBA1*, known to be conserved across legumes including soybean and chickpea. However, it was only recently that Björnsdotter et al. (2020) provided strong experimental evidence linking *RIBA1* to v-c biosynthesis. Their work demonstrated that the GTP cyclohydrolase II (GCHII) function of the *RIBA1* gene, designated as *VICINE-CONVICINE* 1 (*VC1)*, catalyzes the first step of pyrimidine glucoside synthesis from GTP.

Despite this breakthrough, the complete biosynthetic pathway of v-c remains unresolved. Björnsdotter et al. (2020) proposed that v-c biosynthesis involves three additional enzymatic steps, with the first likely catalyzed by an N-glycosidase, as suggested by Frelin et al. (2015). Most importantly, they discovered that a two-base pair frameshift mutation in *VC1* inactivates the gene's GTP cyclohydrolase II domain, significantly reducing v-c biosynthesis by up to 95%. However, faba bean genotypes carrying this mutation still retain residual v-c content, indicating the possible involvement of additional genes in the pathway.

1.9 Scope and objectives

The global human population is projected to reach approximately 9 billion in the coming decades, necessitating a significant increase in food production. Concurrently, over half of the world's population suffers from micronutrient deficiencies, often referred to as "hidden hunger", due to insufficient intake of essential minerals, vitamins and proteins (Zhu et al., 2010). In developing countries, reliance on a single staple crop exacerbates these nutritional deficiencies (Graham and Gregorio, 2001), necessitating the cultivation of high-yielding, nutrient-rich crops, with a particular emphasis on sustainable plant-based protein sources such as faba bean (Augustin and Cole, 2022).

Faba bean is a highly nutritious legume with immense potential for improving food security and combating hidden hunger. It is an environmentally sustainable crop with a high nitrogen-

fixing capacity, reducing dependence on synthetic fertilizers (Köpke and Nemecek, 2010). Compared to other legumes, it enhances soil fertility, supports crop rotations, and minimizes environmental impact (Mesfin, 2020). However, despite these advantages, faba bean remains underutilized due to the presence of v-c antinutritional compounds. The reduction or elimination of v-c content is therefore critical to expanding its use as a protein-rich food source. Breeding efforts to develop low v-c cultivars have been ongoing since the identification of the low v-c trait in a Greek landrace by Duc et al. (1989), with seed v-c levels reduced to approximately 5% of normal concentrations. Since then, breeding programs have focused on introgressing the low v-c allele into elite cultivars for safer human consumption and improved livestock feed quality (Duc et al., 2004; Khazaei et al., 2017). However, progress has been constrained by the incomplete understanding of the genetic mechanisms regulating v-c biosynthesis and the lack of reliable genomic tools.

Although *VC1* has been identified as a key gene in the pathway, its inactivation through a 2-bp mutation does not eliminate v-c completely, suggesting the involvement of additional minor-effect loci. While v-c biosynthesis has traditionally been considered a monogenic trait, studies indicate a more complex genetic architecture. For instance, a genome-wide association study identified SNP and AFLP markers linked to v-c content in a diverse panel of faba bean accessions (Puspitasari, 2017). Although expected associations were found near the *VC1* locus, Webb et al. (2016) identified a marker on chromosome 5 that was strongly linked to a minor v-c QTL. This region overlaps with a second QTL accounting for approximately 15% of phenotypic variation reported by Khazaei et al. (2015), corresponding to a syntenic region on chromosome 7 of *M. truncatula*. These findings suggest that further minor QTL may be influencing v-c biosynthesis, necessitating further investigation.

Furthermore, efforts to implement marker-assisted selection for v-c breeding have met with challenges. Several molecular markers have been developed for breeding, including cleaved amplified polymorphic sequence (CAPs) and especially SNP markers primarily located within the *VC1* gene (Khazaei et al., 2017; Tacke et al., 2021). However, these markers often exhibit low predictive accuracy, often resulting in incorrect genotype classification and false heterozygous clusters for homozygous individuals. This suggests the possible existence of multiple copies of the *VC1* gene or highly similar sequences within the faba bean genome, emphasizing the need for further genetic studies to enable the development of molecular tools and breeding strategies that can facilitate the efficient implementation of MAS in breeding for low v-c.

Recently, high-quality reference genome assemblies for the high and low v-c faba bean cultivars, Hedin2 and Tiffany, respectively, were released (Jayakodi et al., 2023). These assemblies provide essential genomic resources that enable detailed investigations of important agronomic traits, such as v-c content. Given the incomplete understanding of v-c biosynthesis and the persistence of residual v-c content in low v-c genotypes, this thesis therefore aims to leverage these reference genome assemblies to comprehensively investigate the genetic control of v-c biosynthesis. The underlying hypothesis is that, beyond the previously identified *VC1* gene, additional active genes or minor-effect loci contribute to v-c biosynthesis in faba bean. To investigate this, a homology-based approach was employed to map the *RIBOFLAVIN BIOSYNTHESIS PROTEIN 1* (*RIBA1*) to the reference genome assemblies. By identifying all homologous gene models, gene function analysis was performed using a genetically diverse panel of faba bean accessions exhibiting a wide range of seed v-c content. This approach allowed an exploration of the potential involvement of additional genes in the v-c biosynthetic pathway, thereby advancing the understanding of v-c biosynthesis at the genomic level and providing robust molecular tools to support the development of improved faba bean varieties with minimal v-c content. The specific objectives of this thesis included:

1. Functional analysis of all bifunctional riboflavin genes in the faba bean genome using a genetically diverse panel to identify potential minor-effect loci responsible for residual v-c content in low v-c genotypes.
2. Identification of functional polymorphisms associated with variations in v-c content, followed by the development and validation of efficient molecular markers to facilitate the breeding of low v-c faba bean cultivars.
3. Elucidation of the molecular mechanisms underlying maternal inheritance of seed v-c content.

2 Multiple copy number variants of *VC1* gene reveal single-copy expression as a key determinant of vicine content

Ugwuanyi, S., Makhoul, M., Golicz, A.A., Obermeier, C. and Snowdon, R.J. (2025). *Frontiers in Plant Science*. 16:1565210. doi: 10.3389/fpls.2025.1565210

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 13 June 2025
DOI 10.3389/fpls.2025.1565210

OPEN ACCESS

EDITED BY
Elisabetta Mazzucotelli,
Council for Agricultural Research and Economics- Research Centre for Genomics and Bioinformatics, Italy

REVIEWED BY
Ragupathi Nagarajan,
CTC Genomics, United States
Ana Luisa Garcia-Oliveira,
Alliance Bioversity & CIAT, Colombia

*CORRESPONDENCE
Samson Ugwuanyi
samson.ugwuanyi@agrar.uni-giessen.de

RECEIVED 22 January 2025
ACCEPTED 19 May 2025
PUBLISHED 13 June 2025

CITATION
Ugwuanyi S, Makhoul M, Golicz AA, Obermeier C and Snowdon RJ (2025) Multiple copy number variants of *VC1* gene reveal single-copy expression as a key determinant of vicine content. *Front. Plant Sci.* 16:1565210. doi: 10.3389/fpls.2025.1565210

Multiple copy number variants of *VC1* gene reveal single-copy expression as a key determinant of vicine content

Samson Ugwuanyi*, Manar Makhoul, Agnieszka A. Golicz, Christian Obermeier and Rod J. Snowdon

Department of Plant Breeding, Justus Liebig University, Giessen, Germany

Faba bean is a valuable legume crop desired globally for its high nutritional composition. However, the seed vicine and convicine (v-c) content reduces the nutritional quality of faba bean protein and can induce favism in susceptible individuals. Recently, *VC1* gene, encoding a bi-functional riboflavin protein, was reported to initiate the v-c biosynthetic pathway in *V. faba*. In low v-c cultivars, a 2 bp insertion in this gene disrupts its function by causing a frameshift and premature stop codon. However, because v-c biosynthesis is only partially reduced, this suggests that additional genes may also be involved in the pathway. Here, we identify and investigate multiple tandem gene duplications at the *VC1* locus. Our findings reveal that *VC1* exhibits multiple structural variants and copy number variations, but its expression is independent of copy number. Low v-c genotypes carry both variants of the gene — with and without the 2 bp insertion — but only the variant with the insertion is expressed. In contrast, high v-c genotypes consistently express the variant lacking the insertion. Although some high v-c genotypes also carry the insertion, it is found in a non-expressed variant, while the expressed variant lacks the insertion, resulting in the high v-c phenotype. We also report a novel diverging *VC1* homolog, *RIBA2*, which shares expression domains with *VC1*. This homologous gene encodes GTP cyclohydrolase II, a critical enzyme in the v-c pathway. Expression of this gene contributes ~5-10% of riboflavin gene transcripts in immature seeds suggesting it as a minor-effect candidate locus in v-c biosynthesis. Moreover, two SNPs within the coding sequence of *RIBA2* segregated with v-c content, offering a reliable alternative for marker-assisted selection in faba bean breeding. In conclusion, this study contributes to the elucidation of the complex genetic regulation of v-c biosynthesis and provides valuable insights to facilitate further efforts in its reduction in faba bean.

KEYWORDS

VC1 gene, structural variation, KASP markers, *Vicia faba*, convicine

1 Introduction

Faba bean (*Vicia faba* L.) is a grain legume which is globally important for its highly nutritious, protein-rich seeds. The global production of faba bean stands at 5.7 million tons in 2020 (FAOSTAT, 2022), making it the sixth most produced pulse crop and the highest yielding legume after soybean (Adhikari et al., 2021). It is reported to have originated from the Mediterranean basin and subsequently spread to be cultivated across nearly all continents worldwide (Duc et al., 2010; Caracuta et al., 2016). It has become increasingly popular, especially in cool-season climates where other protein crops perform poorly. Faba bean seeds contain up to 37% protein and are rich in micronutrients, making them a suitable source of food for humans and feed for their livestock (Duc et al., 1999; Karkanis et al., 2018; Warsame et al., 2020). In addition to its nutritional benefits, faba bean can improve soil fertility in association with rhizobium bacteria by fixing a significant quantity of nitrogen in the soil, which reduces the need for application of inorganic nitrogen fertilizer in subsequent seasons (Karkanis et al., 2018; Adhikari et al., 2021). This feature is leveraged in many agricultural systems by incorporating faba bean into crop rotation or mixed cropping with others crops such as cereals (Stagnari et al., 2017; Carrillo-Perdomo et al., 2020). This ecological importance of faba bean and its roles in food and feed has increased its global reputation significantly.

However, the agronomic relevance of faba bean is limited by the presence of significant quantities of vicine and convicine (v-c) in all parts of the plants (Luzzatto and Arese, 2018; Choudhary and Mishra, 2019; Badjona et al., 2023). The metabolic products of vicine and convicine—divicine and isouramil—release free radicals that cause oxidative damage to red blood cells in people with glucose-6-phosphate dehydrogenase (G6PD) deficiency, leading to acute hemolytic anemia, a condition also known as favism (Badjona et al., 2023; Björnsdotter et al., 2021). Therefore, to enhance faba bean usage and its general acceptability, reducing the v-c content to the barest minimum safe for food and feed is essential. However, the genetic basis of v-c accumulation remains to be fully elucidated, and all low v-c cultivars still carry baseline v-c levels.

Previous studies have highlighted a bimodal pattern in the v-c phenotype, primarily influenced by a major quantitative trait locus (QTL) on chromosome 1 of the faba bean genome (Duc et al., 1989; Ramsay et al., 1995; Gutierrez et al., 2006; Khazaei et al., 2015). Molecular markers flanking this locus were identified by Khazaei et al. (2015); however, the gene *RIBA1* was only recently discovered within this region by Björnsdotter et al. (2021). *RIBA1*, named *VC1*, encodes a bi-functional riboflavin protein responsible for catalyzing the pivotal step in the v-c biosynthetic pathway. The gene has two functional domains, RibA and RibB, encoding GTP cyclohydrolase II (GCHII) and 3,4-dihydroxy-2-butanone-4-phosphate synthase (DHBPS), respectively. However, it is the GTP cyclohydrolase II domain that directly catalyzes v-c biosynthesis, via conversion of purine nucleoside triphosphate GTP into the unstable intermediates leading to v-c (Björnsdotter et al., 2021). A 2 bp insertion in the GCHII domain leads to loss of function in low v-c cultivars. This frameshift mutation inactivates *VC1* by causing a premature stop codon, hindering the correct synthesis of GTP cyclohydrolase II. However, this mutation does not eliminate v-c completely, although it causes a significant reduction. This suggests a potential involvement of other genes or gene copies, necessitating further research to comprehensively understand the genetic factors that control v-c biosynthesis in faba bean. Recently completed *V. faba* genome assemblies and gene annotations for low and high v-c cultivars (Jayakodi et al., 2023) provide an excellent basis to advance knowledge in this regard. Therefore, the objectives of this study were to elucidate the genetic regulation of v-c content by identifying active bi-functional riboflavin genes and polymorphisms corresponding to changes in v-c content. Due to the residual v-c content in seeds of low v-c genotypes, we hypothesize that at least one additional locus controls v-c biosynthesis. To study this, we identified RIBA protein-coding genes by homology mapping of the RIBA1 protein sequence to two faba bean reference genomes with high and low v-c content, respectively. All similar protein-encoding genes were identified and functionally analyzed across a set of well-characterized low and high v-c cultivars using bioinformatic predictions, gene expression and transcript analyses, and phenotype correlation analysis.

Abbreviations: cDNA, Complementary DNA; CNV, Copy number variation; DHBPS, 3,4-Dihydroxy-2-Butanone-4-Phosphate Synthase; DNA, Deoxyribonucleic Acid; ESF, Early Seed Filling; G6PD, Glucose-6-Phosphate Dehydrogenase; GCHII, Guanosine Triphosphate Cyclohydrolase II; gDNA, Genomic DNA; GTP, Guanosine Triphosphate; HVC, High Vicine-Convicine; KASP, Kompetitive Allele-Specific PCR; LSF, Late Seed Filling; LVC, Low Vicine-Convicine; miRNA, MicroRNA; NCBI, National Center for Biotechnology Information; PCR, Polymerase Chain Reaction; qPCR, Quantitative Polymerase Chain Reaction; QTL, Quantitative Trait Locus; RNA, Ribonucleic Acid; RT-qPCR, Reverse Transcription Quantitative Polymerase Chain Reaction; SNP, Single Nucleotide Polymorphism; v-c, Vicine and Convicine.

2 Materials and methods

2.1 Plant materials and cultivation

The study utilized a diverse set of well-characterized faba bean lines consisting of nine low v-c and nine high v-c genotypes with different genetic backgrounds (Supplementary Table S1). These lines represent a set of carefully selected genotypes with known characteristics and performances. They were obtained from Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ, Hohenlieth, Germany) and comprise both commercial cultivars and lines inbred beyond the fifth generation of selfing. Vicine content for each genotype was measured from a sample of two faba bean seeds using spectrophotometry as described by Sixdenier et al. (1996).

Two seeds from each genotype were planted in 4-liter plastic pots within a pollinator-proof chamber in the greenhouse. Fresh leaves at 21 days after planting were collected from the first fully opened leaves for DNA isolation. For RNA extraction for gene expression and transcript analyses, a subset of ten genotypes was selected, including five with low and five with high seed v-c content.

The plants were manually tripped during flowering to ensure pod-set (Kambal et al., 1976). Fresh and immature seeds were collected from this subset for RNA isolation at two stages during seed development: early seed filling stage (ESF) (stage four) and late seed filling stage (LSF) (stage six). Specifically, seeds aged 13–16 days and 20–25 days after tripping were selected, respectively. At the ESF stage, the cleft between the cotyledons is broad, with a spherical chalazal chamber. The embryo has a butterfly-shaped appearance and is surrounded by endosperm. In contrast, at LSF, the intense green cotyledons are closely positioned and have a curved axis (Borisjuk et al., 1995).

2.2 DNA extraction

Fresh leaves collected from each genotype were immediately flash-frozen in liquid nitrogen and lysed using Qiagen Tissue Lyser II (Qiagen, Düsseldorf, Germany). Genomic DNA was isolated from the leaf powder following the Doyle and Doyle (1990) method.

2.3 RNA extraction and cDNA synthesis

Immature seeds were immediately flash-frozen in liquid nitrogen and then manually ground into powder using a mortar and pestle. RNA was isolated from 100 mg of the powdered samples using the Zymo RNA miniprep kit (Zymo Research, Freiburg, Germany) following the manufacturer's manual. DNaseI treatment was performed to remove genomic DNA, following the procedure outlined in the Zymo RNA kit manual. RNA concentration and quality were determined using Qubit RNA assay kit (ThermoFisher Scientific, Germany) and agarose gel (1%), respectively. First-strand cDNA was synthesized using the RevertAid cDNA synthesis kit (ThermoFisher Scientific, Germany). Initially, 1 µl of Random Hexamer primer was added to 1 µg of RNA, incubated for 5 minutes at 65°C, and subsequently cooled on ice. The cDNA reaction master mix was prepared by adding 4 µl of Reaction buffer (5x), 2 µl of dNTP mix (10mM), 1 µl of Ribolock RNase inhibitor (20 U/µl), and 1 µl of RevertAid H Minus Reverse Transcriptase (200 U/µl). Reactions were carried out in a thermal cycler at the following temperature conditions: 25°C for 5 minutes, 42°C for 60 minutes, 70°C for 5 minutes, and then held at 4°C. The resulting cDNA was utilized for gene expression analysis and Sanger sequencing of *VC1* and *RIBA2* genes.

2.4 Homology-based gene identification

Custom databases for faba bean were established from the reference assemblies of the high v-c cultivar, Hedin, and the low v-c cultivar, Tiffany (Jayakodi et al., 2023), using ncbi-blast 2.12.0+. RIBA1 protein sequence was aligned to these databases using the tblastn function to identify all RIBA genes in both genomes. Genes exhibiting sequence similarity were identified and filtered to retain only those with over 85% protein sequence similarity to RIBA1, and were then functionally analyzed in this study.

2.5 Phylogenetic and sequence analysis

Alignments were performed using MAFFT tool in Jalview (version 2.11.3.0). Phylogenetic analysis was conducted using the phylogeny.fr platform. The maximum likelihood method, implemented in the PhyML program (v3.1/3.0 aLRT), was employed to reconstruct the phylogenetic tree while TreeDyn (v198.3) was used for tree rendering. Amino acid sequences of RIBA proteins for chickpea (XP_004485599), lupin (KAE9591829), grass pea (CAK6822722), lotus (XP_057429064), Medicago (XP_003593237) and pea (XP_050880829) were downloaded from NCBI database (https://www.ncbi.nlm.nih.gov/). Functional domains of RIBA proteins were predicted using PsiPred workbench (http://bioinf.cs.ucl.ac.uk/psipred/).

2.6 Primer design and synthesis

All primers for the experiments were designed using Primer3 plus (https://www.primer3plus.com/index.html) and subsequently synthesized by Microsynth AG (Balgach, Switzerland). For each primer, gene sequences from the two reference genomes, Hedin and Tiffany (Jayakodi et al., 2023), were aligned to identify conserved regions with priming efficiency, as predicted by Primer3 Plus.

2.7 PCR Validation of *VC1* and *RIBA2* genes in faba bean

All *VC1* variants and *RIBA2* were validated in faba bean genotypes using selective PCR amplification. PCR reactions were set up in a final volume of 25 µl, including 12.5 µl of GoTaq Hot Start Green Master Mix, (Promega, Madison, WI, United States), 1.25 µl of 10 µM forward and reverse primers (Supplementary Table S2), 1.5 µl of genomic DNA and 8.5 µl of MilliQ water. The reactions were carried out in a T100 Thermal Cycler (Bio-Rad Laboratories, Hercules, CA, United States) with the following conditions: 94°C for 2 minutes, 35 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds and extension at 72°C for 40 to 90 seconds depending on the size of amplicon, followed by final extension for 5 minutes at 72°C. Amplicons were separated on a 1% agarose gel and visualized under UV light.

2.8 *VC1* copy number determination by quantitative PCR

VC1 copy number was determined by quantitative PCR in 10 µl final volume, containing 5 µl 2x SYBR Green master mix (ThermoFisher Scientific, Germany), 1 µl 10 µM forward primer and 1 µl 10 µM reverse primer (Supplementary Table S2), 1 µl of DNA and 2 µl of MilliQ water. ELF1A was used as the reference gene for normalization (Gutierrez et al., 2011). There were two biological and three technical replicates for each genotype, as well as triplicates of water samples serving as no-template controls.

Quantitative PCR was done using StepOneplus (ThermoFisher Scientific, Germany) with the following temperature conditions: 95°C for 10 minutes, 40 cycles of denaturation at 95°C for 15 seconds followed by 60°C for 1 minute. Relative quantification was determined by delta-delta CT (ΔΔCt) method (Livak and Schmittgen, 2001).

2.9 Relative quantification of *VC1* and *RIBA2* expression levels by reverse-transcription quantitative PCR

The expression levels of *VC1* and *RIBA2* were determined from cDNA synthesized from the previous step. The reaction mixture consisted of 5 µl of 2x SYBR Green master mix (ThermoFisher Scientific, Germany), 1 µl of 10 µM forward primer and 1 µl of 10 µM reverse primer (Supplementary Table S2), 1 µl of cDNA, and 2 µl of MilliQ water. The reference genes, CYP2 and ELF1A, were used for normalization as they were previously reported to be stably expressed in faba bean (Gutierrez et al., 2011). As controls, three water samples were included as no-template controls and each sample had two biological and three technical replicates. Quantitative PCR was done using StepOneplus (ThermoFisher Scientific, Germany) as described above for copy number quantification and relative transcript levels were determined by delta-delta CT (ΔΔCt) method (Livak and Schmittgen, 2001).

2.10 *VC1* and *RIBA2* cDNA sequencing

The cDNA samples, synthesized as described above, were used for selective amplification of *VC1* and *RIBA2*. Forward and reverse primer pairs specific to *VC1* and *RIBA2* were used (Supplementary Table S2). The PCR conditions were as described previously in the gene validation section, except that the annealing temperature was set at 58°C. Subsequently, the resulting amplicons were sent for Sanger sequencing at Microsynth AG (Balgach, Switzerland). For each genotype, samples from the two developmental stages were sequenced.

2.11 KASP marker assay development and SNP genotyping

Genotyping was performed using Kompetitive Allele-Specific PCR (KASP) assay technique for polymorphisms within *VC1* gene. *VC1* genes from the two reference assemblies were aligned to identify possible single nucleotide polymorphisms. For identified variants, allele-specific primers were designed. To detect the presence of 2 bp insertion in *VC1* (SNP08) within our faba bean set, primers were designed where A1 binds to the wild-type variant, and A2 binds to the mutant variant. The common primer used was C1. Detailed information on all *KASP* markers can be found in Supplementary Table S3. The KASP marker assay procedure was conducted according to the methodology outlined in the study by Makhoul and Obermeier (2022).

2.12 Data analysis

All experiments, including PCR, sequencing, copy number quantification, gene expression analysis, and KASP genotyping, were repeated at least twice. The resulting qPCR data from copy number quantification and gene expressions were analyzed using Microsoft Excel and R (version 4.3.2). The library packages ggpubr and ggplot2 were used to generate plots using R studio. Statistical differences were inferred using t-test for two groups or one-way ANOVA for more than two groups.

3 Results

3.1 Multiple gene models encode RIBA proteins in the *Vicia faba* genome

In this study, we investigated the role of the different genes that encode bifunctional RIBA enzymes responsible for catalyzing the initial step in the v-c biosynthetic pathway. We employed a homology-based method to align RIBA1 protein to the recently assembled faba bean reference genomes, Hedin— a high v-c cultivar — and Tiffany— a low v-c cultivar (Jayakodi et al., 2023). Our analysis revealed the presence of multiple gene models encoding RIBA protein in the faba bean genome. In Hedin, we identified four such genes on chromosome 1 and contig_8341, namely Vfaba.Hedin2.R1.1g485480, Vfaba.Hedin2.R1.1g485520, Vfaba.Hedin2.R1.1g485560, and Vfaba.Hedin2.R1.Ung108560. Tiffany, on the other hand, had five genes on chromosome 1, including Vfaba.Tiffany.R1.1g399960, Vfaba.Tiffany.R1.1g400000, Vfaba.Tiffany.R1.1g400040, Vfaba.Tiffany.R1.1g400120, and Vfaba.Tiffany.R1.1g400280 (Figure 1a; Supplementary Table S4). Comparative analysis indicated that three of these genes from Hedin and four from Tiffany shared significant identity (>96%) with the *RIBA1* (*VC1)* gene previously reported by Björnsdotter et al. (2021) (Figure 1b). Therefore, we classified these genes as genetic variants of *VC1*. These *VC1* variants exhibited contrasting structural variations that grouped into three major variants, defined by their structural characteristics. *VC1A* had a tandem duplication of a 65 bp in intron 3 and *VC1B* had a partial deletion in this intron, while *VC1C* lacked the entire intron (Figure 1c). There were three copies of *VC1A* in the Hedin assembly and two copies each of *VC1B* and *VC1C* in Tiffany. As expected, one copy each of *VC1B* and *VC1C* carried a 2 bp frameshift insertion in exon 6, as reported by Björnsdotter et al. (2021), which renders the gene product non-functional. This insertion was absent in *VC1A*. The modification in intron 3 is associated with a three-nucleotide difference in exon 3 and can distinguish *VC1C/vc1c* from *VC1A* and *VC1B/vc1b*. Beside these variations, *vc1b* and *VC1C* share similar alleles at most SNP positions. Additionally, there are SNPs within exon 4 and 5 that can distinguish other *VC1* variants (Supplementary Figure S1).

The fourth and fifth gene models identified in Hedin and Tiffany, respectively, shared only approximately 65% sequence similarity with *VC1* and 99% sequence similarity with each other, suggesting that they represent a newly discovered *VC1*-like gene.

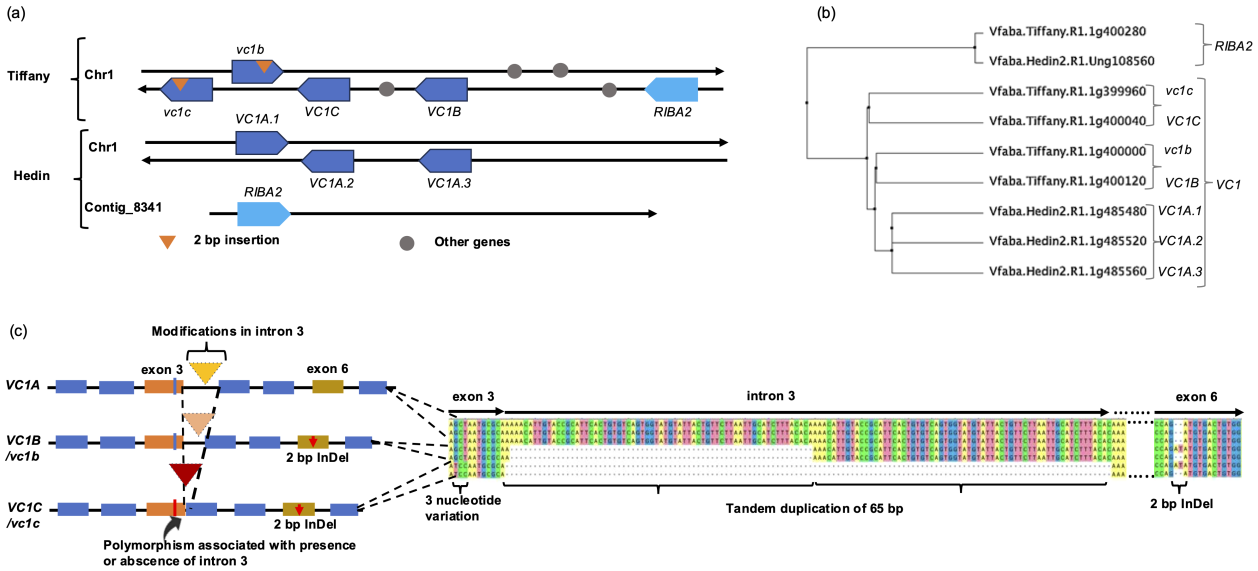


FIGURE 1
Identification and analysis of RIBA genes. **(a)** Configurations of RIBA genes from two reference assemblies. The direction of arrows shows the orientation of the genes on the top and bottom strand of chromosome 1 and Contig_8341. **(b)** Phylogenetic tree showing the genetic relationship among RIBA genes in *V. faba*. **(c)** Major structural variations in *VC1* genes. *VC1* variants were structurally differentiated by modifications in intron 3. These modifications are associated with a 3-nucleotide variation in exon 3.

While there were structural variations between *VC1* and this homolog, these differences occurred primarily in non-coding regions, with the coding regions displaying a high degree of similarity (>87%). As a result, they shared over 90% similarity in their predicted protein sequences.

Analysis of the protein sequence of this homologous gene showed presence of two functional domains, RibA and RibB, as in other bifunctional riboflavin proteins, and highly identical to *VC1* protein domains. Subsequently, we analyzed bifunctional RIBA protein homologs from other legume crops in the Fabaceae family including chickpea, lupin, grass pea, lotus, Medicago and pea. Comparison of the amino acid sequence encoded by this newly identified *V. faba* gene to those of *VC1* and RIBA genes from other legumes revealed high similarity among these RIBA proteins and conservation of all key amino acid residues required for catalytic activity. All the proteins shared a high identity in the two catalytic domains, except *vc1b* and *vc1c* which differed significantly in the second domain due to a 2 bp insertion, altering the reading frame and causing a premature stop codon (Figures 2a, b). Hence, the high degree of similarity among these proteins, especially within their functional domains, suggests that their roles as RIBA proteins are well conserved. We therefore denominated this novel homolog as *RIBA2*, since it represents a second RIBA locus in *V. faba*.

We validated the existence of all the genes in a set of 18 faba bean genotypes. First, we selectively amplified the genes using PCR. The results indicated that *RIBA2* was present in all faba bean genotypes. *VC1A* was present in most high v-c genotypes, while *VC1B* and *VC1C* were found in all low v-c genotypes but were also carried by some high v-c genotypes (Figure 3a). It was also observed that all three *VC1* variants could be present in a single genotype, as observed in the low v-c genotype, NPZ-FB-73, and the high v-c genotype, NP-FB-143. Secondly, the presence of active and inactive copies of *VC1B* and *VC1C* in the Tiffany genome assembly suggests that low v-c genotypes may still carry functional copies of *VC1*. To confirm the hypothesis that both active and inactive copies of *VC1* are carried by low v-c genotypes, we employed a KASP assay targeting the 2 bp insertion in exon 6. KASP primers were designed such that allele 1-specific primers bind to the active variants while allele 2-specific primers bind to the mutated variants. Subsequently, it was observed that most high v-c genotypes (7 out of 9) were homozygous for the assay indicating the presence of only active variants, whereas all nine genotypes with low v-c and two high v-c genotypes were heterozygous, indicating the presence of both active and mutated variants (Figures 3b, c). There was no homozygous genotype call for only the mutant, revealing that no faba bean genotype carried only the mutant gene copies. However, since a set of 18 faba bean genotypes is too small to draw definitive conclusions, we extended our analysis to a larger and more diverse panel of 97 genotypes to test this hypothesis. This diversity panel included breeding lines and cultivars exhibiting a wide range of seed v-c content, obtained from Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ, Hohenlieth, Germany). We observed that all low v-c genotypes, as well as some high v-c genotypes, carried a functional *VC1* variant in addition to the non-functional *vc1* allele (Figure 3d; Supplementary Table S5), confirming the hypothesis that all low v-c faba beans carry multiple *VC1* variants, including gene variants with and without the inactivating insertion. Additionally, some high v-c faba beans exhibit the same characteristic.

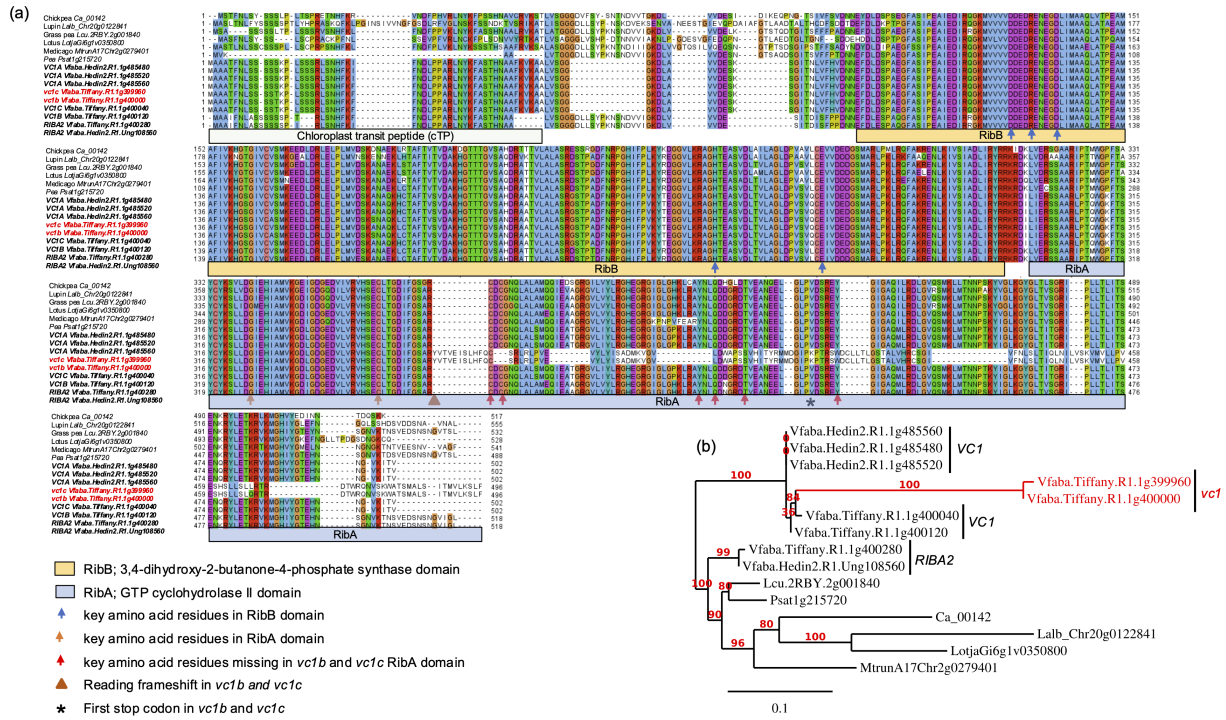


FIGURE 2
Comparison of amino acid sequences and phylogeny of RIBA proteins. **(a)** Alignment of amino acid sequences and **(b)** phylogenetic tree of RIBA proteins from *V. faba* and related legume crops in Fabaceae family, including chickpea, lupin, grass pea, lotus, Medicago and pea. Major variations lie within the domain encoding chloroplast transit peptide (cTP). RibB and RibA domains of the proteins show high similarity, except the protein sequences encoded by *vc1b* and *vc1c* which show variations in the RibA domain due to a frameshift mutation. The functional domains are shown underneath the alignment as predicted using PsiPred workbench. Predicted key residues based on Björnsdotter et al. (2021) are indicated with different colors of arrows.

3.2 *VC1* exhibits copy number variations and is dosage-insensitive

The presence of multiple copies of *VC1* in faba bean reference genomes suggests a potential variable copy number for the gene. As a result, we determined *VC1* copy number through relative quantification by qPCR. The result confirmed that *VC1* shows copy number variations, with gene copies ranging from 2 to 5 across 18 faba bean genotypes (Figure 4a). Unexpectedly, low v-c genotypes tended to carry higher copy numbers than high v-c genotypes. This observation could mean that some of the gene variants are not functional or not expressed as was observed for the faba bean hilum color locus (Jayakodi et al., 2023). Therefore, we assessed the transcription activity of *VC1* and *RIBA2* genes by quantifying their relative expression levels using RT-qPCR in a subset of ten faba bean cultivars with varying v-c contents. These cultivars, selected to reflect the diversity in seed v-c levels, included genotypes representing all known *VC1* haplotypes. Furthermore, our focus was on whole seeds at the early stage of seed development, in which *VC1* expression is known to be the highest (Björnsdotter et al., 2021).

Expression analysis revealed significantly ($P < 0.001$) higher *VC1* expression levels in high v-c genotypes at early and late seed filling stages compared to low v-c genotypes (Figure 4b). High v-c genotypes showed about 5-fold and 4-fold more expression ($P < 0.001$) than low v-c genotypes at ESF and LSF, respectively (Figure 4c). In contrast to *VC1* differential expression, *RIBA2* displayed relatively consistent expression levels between high and low v-c genotype groups at both stages (Figure 4d). However, *VC1* expression was significantly higher than *RIBA2* expression. Within high v-c genotypes, *VC1* exhibited up to 11-fold more expression than *RIBA2* (Figure 4e), compared to 2.6-fold within low v-c genotypes (Figure 4f).

Importantly, despite fewer copy numbers in high v-c genotypes, the elevated expression of *VC1* suggests that it may not be sensitive to dosage. To validate this, we conducted a correlation analysis between *VC1* gene copy number and gene expression. Subsequently, we observed a negative correlation (Supplementary Figure S2A), indicating that higher copy number correlated with lower expression level. This negative correlation seems to be attributed to a shared genetic origin among low v-c genotypes, rather than interactive effects of variants resulting from antisense regulation. To verify this observation, we used a SNP genotyping dataset to trace the lineage of our low v-c lines across a broader panel of 347 faba bean genotypes. These genotypes were genotyped using the faba bean 50k Affymetrix chip. After filtering for missing data, minimum allele frequency, and heterozygosity, 13k polymorphic SNPs remained and were used for population structure analysis. The principal component analysis plot depicted a close clustering of the low v-c lines (Supplementary Figure S2B). This clustering pattern confirms a common genetic origin among low v-c genotypes,

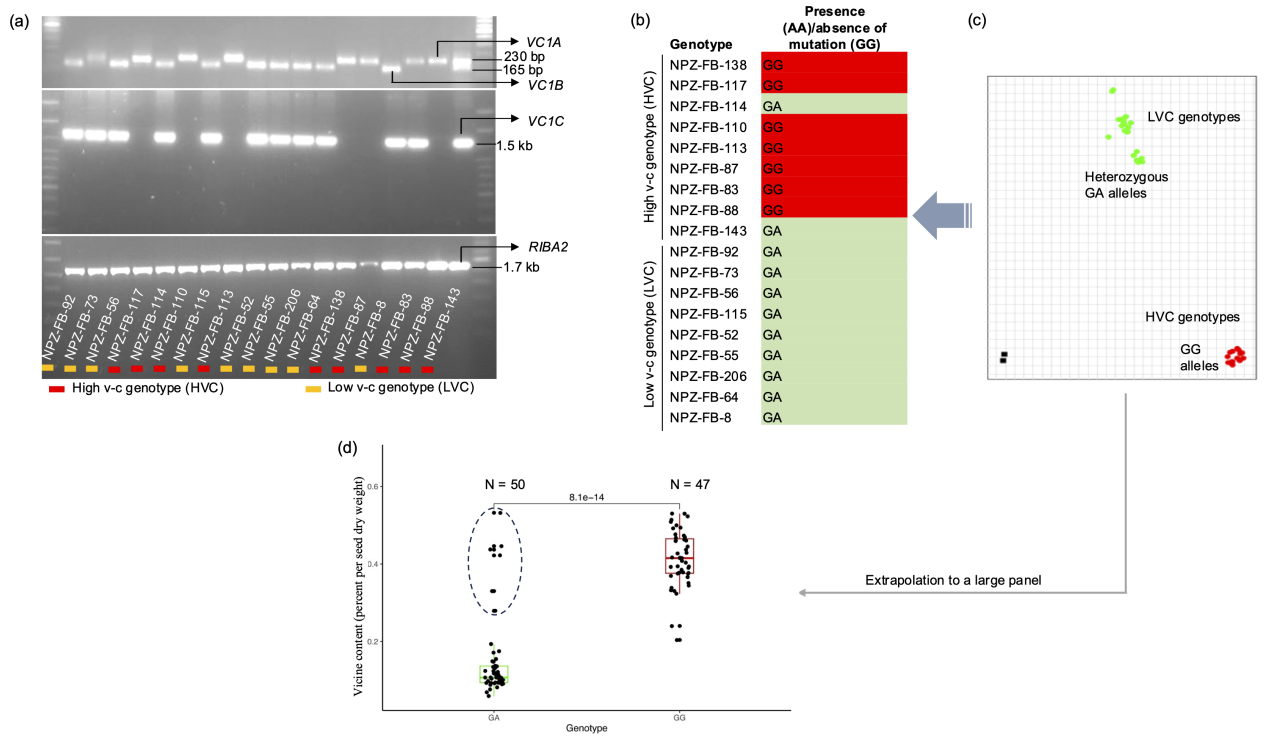


FIGURE 3
Validation of the presence of RIBA genes and the mutation in *VC1*. **(a)** Selective amplification by PCR showing the presence of *VC1* and *RIBA2* variants in a set of 18 faba bean lines. **(b)** Allele calls and **(c)** allelic discrimination plot for the 2 bp insertion in exon 6 of *VC1*. Allele GG represents the absence of insertion (wild type variants), while allele AA represents the presence of insertion (mutant). Heterozygous cluster (GA) indicates the presence of both the wild type and mutant variants. **(d)** Boxplot illustrating the v-c contents for two groups: *VC1* wildtypes (GG) and genotypes carrying both the mutant and wildtype (GA) gene variants in a panel of 97 genotypes. High v-c genotypes carrying GA alleles are enclosed within a circle.

elucidating the observed negative correlation between *VC1* gene copy number and expression.

However, this observed correlation between high copy number and low expression suggests that not all *VC1* copies are expressed, potentially due to some regulatory mechanisms. To identify the functional variants, we sequenced cDNA fragments of *VC1*. As *VC1* variants can be differentiated by the mutations in exon 3 and 6, and SNPs within exon 4 and 5, we designed primers flanking this region and subsequently sequenced amplicons from four low v-c and five high v-c genotypes.

The analysis of cDNA sequences showed that three *VC1* gene variants were expressed across genotypes, namely, high v-c genotypes can express *VC1A* or *VC1C*, while *vc1b* was detected in low v-c genotypes (Figures 5a, b; Supplementary Table S6). Surprisingly, only one gene variant was expressed per genotype, regardless of the total number of variants present. For instance, low v-c genotype NPZ-FB-92 carried four *VC1* gene copies, two each of *VC1B* and *VC1C*, but only *vc1b* mutant was expressed, as in all other low v-c genotypes. Similarly, high v-c genotype NPZ-FB-143 carried multiple gene copies including all three variants but expressed only *VC1A*. This demonstrates and explains why *VC1* was found dosage insensitive.

Additionally, we constructed a phylogenetic tree using *VC1* cDNA sequences for the sequenced subset. The genotypes clustered into two major groups (Figure 5c). The first group comprised only high v-c genotypes expressing *VC1A*, while the second group comprised two subgroups: low v-c genotypes expressing *vc1b* and high v-c genotypes expressing *VC1C*. This distinction within the second subgroup was primarily attributed to the presence or absence of the 2 bp insertion, suggesting it as a functional polymorphism within *VC1* capable of distinguishing genotypes based on v-c contents. We additionally sequenced *RIBA2* cDNA which overlaps with this *VC1* region, particularly within the GCHII domain. *RIBA2* cDNA sequences from both low v-c and high v-c genotypes did not carry any inactivating insertion (Supplementary Table S6).

3.3 *RIBA2* gene is a candidate v-c locus with localized SNPs that segregate with v-c phenotypes

It is evident that *VC1* is not the only locus controlling v-c biosynthesis in faba bean. We identified a diverging homolog of *VC1* which may also be involved in v-c regulation. As earlier noted, analysis of protein sequences of this homologous gene revealed two functional domains, RibA and RibB, similar to the functional *VC1* protein domains, which are conserved across related legumes like chickpea, lupin, and pea. This high similarity and conservation of

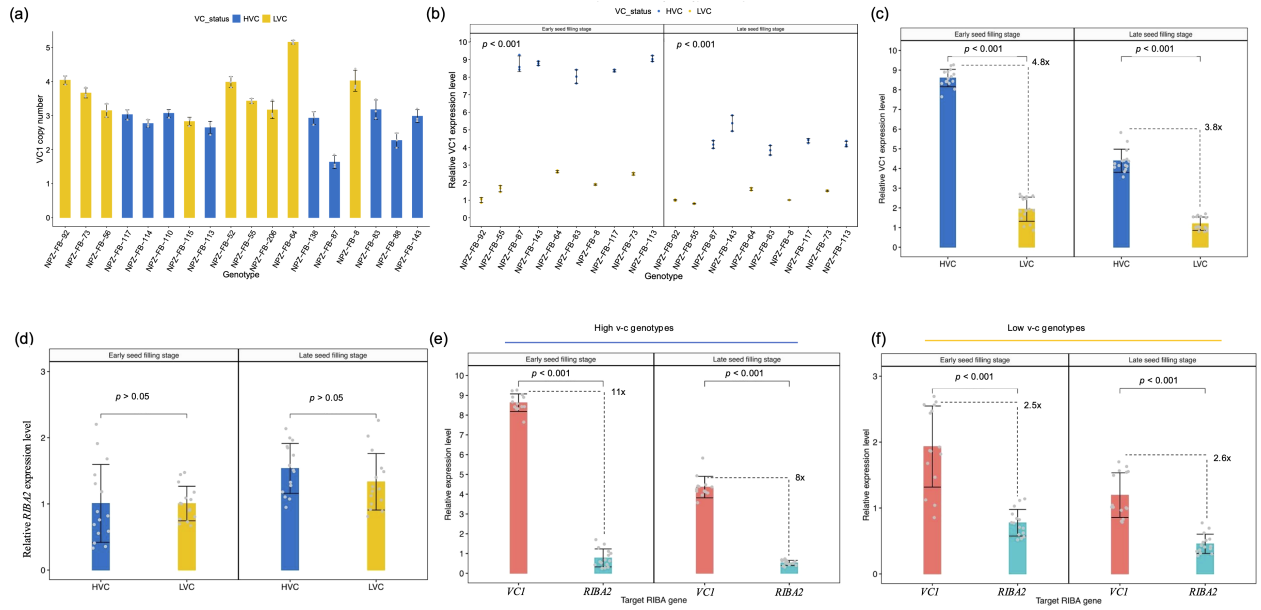


FIGURE 4
Copy number variation and transcription activities of RIBA genes. **(a)** *VC1* copy numbers across 18 faba bean genotypes, comprising high (HVC) and low vicine-convicine (LVC) lines (means ± SD, n = 3). **(b)** *VC1* relative expression levels during seed development (means ± SD, n = 3). **(c)** Comparisons of the relative expression levels between high and low v-c genotypes at early and late seed filling stages (means ± SD, n = 15). **(d)** Comparisons of relative levels of *RIBA2* transcripts between high and low v-c (means ± SD, n = 15). **(e)** Comparisons of *VC1* and *RIBA2* expression levels within high v-c genotypes (means ± SD, n = 15). **(f)** Comparison of *VC1* and *RIBA2* expression levels within low v-c genotypes (means ± SD, n = 15). Significance for comparisons between groups were determined using t-test.

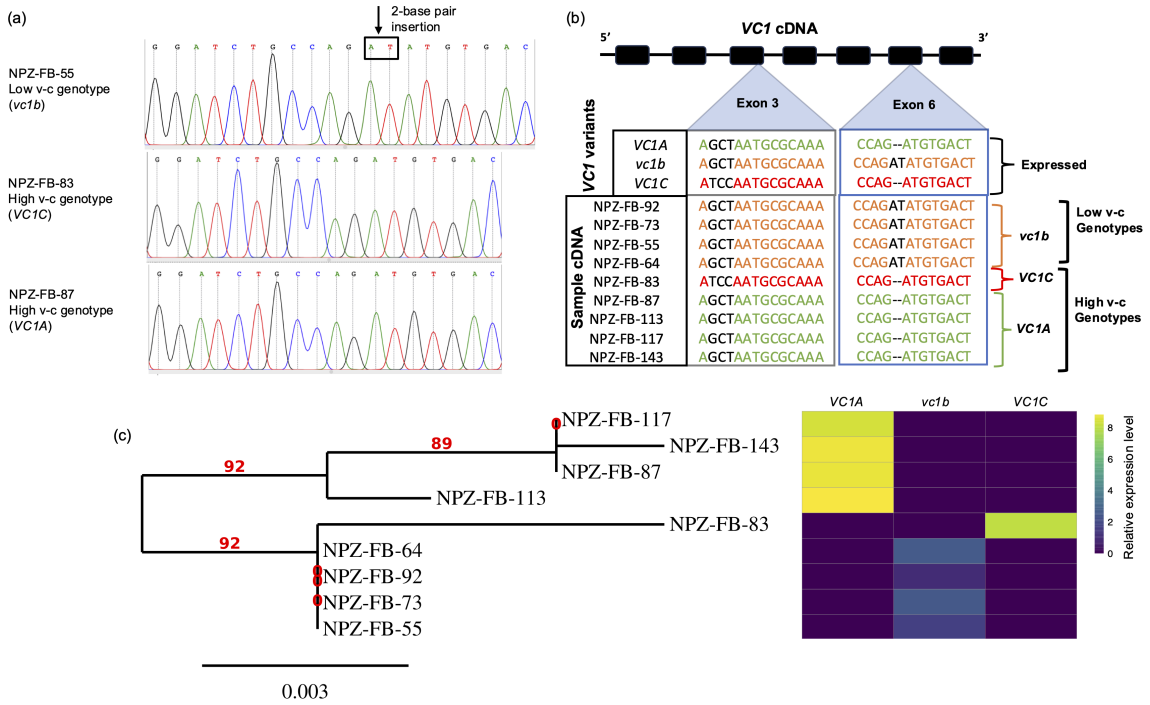


FIGURE 5
Identification of functional *VC1* variants. **(a)** Chromatograms obtained by Sanger sequencing of PCR fragments of VC1 cDNA show single peaks corresponding to only one variant of the gene. The position of the 2-base pair insertion is indicated by the arrow, and the nucleotides are enclosed in a box. **(b)** *VC1* cDNA analysis showing the expressed *VC1* variants. For each genotype, cDNA samples from two developmental stages were sequenced. In addition to the mutations shown here, single nucleotide polymorphisms in exons 4 and 5 were used to distinguish the variants (see Supplementary Table S6). **(c)** Phylogenetic tree showing the genetic relationships among the faba bean lines using maximum likelihood method, alongside a heatmap illustrating expression profiles for *VC1* variants.

key amino acid residues among these proteins indicates that their roles as RIBA proteins are well conserved (see Figure 2a), suggesting that *RIBA2* is another functional RIBA locus in faba bean. Notable differential expressions were evident among individual faba bean lines for this gene, with observed variations tending to correlate with differences in v-c contents (Figure 6a).

Furthermore, when comparing the proportion of *RIBA2* expression relative to the combined expression of both RIBA genes in high v-c cultivars, we observed an average transcript level at least 9.5-fold lower, ranging up to 20-fold (Figure 6b). This would represent about 5-10% of v-c content relative to high v-c cultivars and corresponds to the observed phenotypic differences between low and high v-c cultivars (Figure 6c). These results suggest that this neighboring homolog of *VC1* is a candidate locus with minor effect in v-c regulation.

Additionally, two neighboring single nucleotide polymorphisms within exon 6 of the *RIBA2* gene segregate with v-c phenotypes. We evaluated the segregation pattern of a SNP in exon 6 of this gene in a diversity panel of 97 genotypes. The results showed that this polymorphism can differentiate low v-c from high v-c genotypes, where allele C segregated with low v-c phenotype while allele G segregated with high v-c phenotype (Figure 6d; Supplementary Table S5), demonstrating that the SNPs are tightly linked to v-c phenotypes and can be utilized for marker-assisted selection in v-c breeding.

3.4 Genetic variations in v-c contents involve polymorphisms associated with differential expression of *VC1* and *RIBA2* genes

It is commonly observed that high v-c faba beans exhibit substantial variations in v-c content, whereas low v-c faba beans show much narrower variation, significantly below this threshold (Khamassi et al., 2013; Puspitasari et al., 2022). This pattern suggests that the 2 bp inactivating insertion in *VC1*, although critical, may not fully explain the observed phenotypic diversity. Based on our findings, we propose that polymorphisms influencing the expression levels of *VC1* and *RIBA2*, in addition to the 2 bp insertion, contribute to the variations in v-c content (Supplementary Figure S3). Specifically, *VC1* expression levels can vary up to five-fold among faba bean genotypes, potentially accounting for much of the variability within high v-c group. In contrast, the 2 bp mutation in *VC1* appears to be the primary determinant of the sharp phenotypic difference between high and low v-c genotypes. Although *RIBA2* generally exhibits low expression and may play a minor role, it could be the sole contributor to v-c biosynthesis in low v-c lines lacking functional *VC1*. This may explain the narrow v-c variation observed in these genotypes, which remains substantially lower than in high v-c lines.

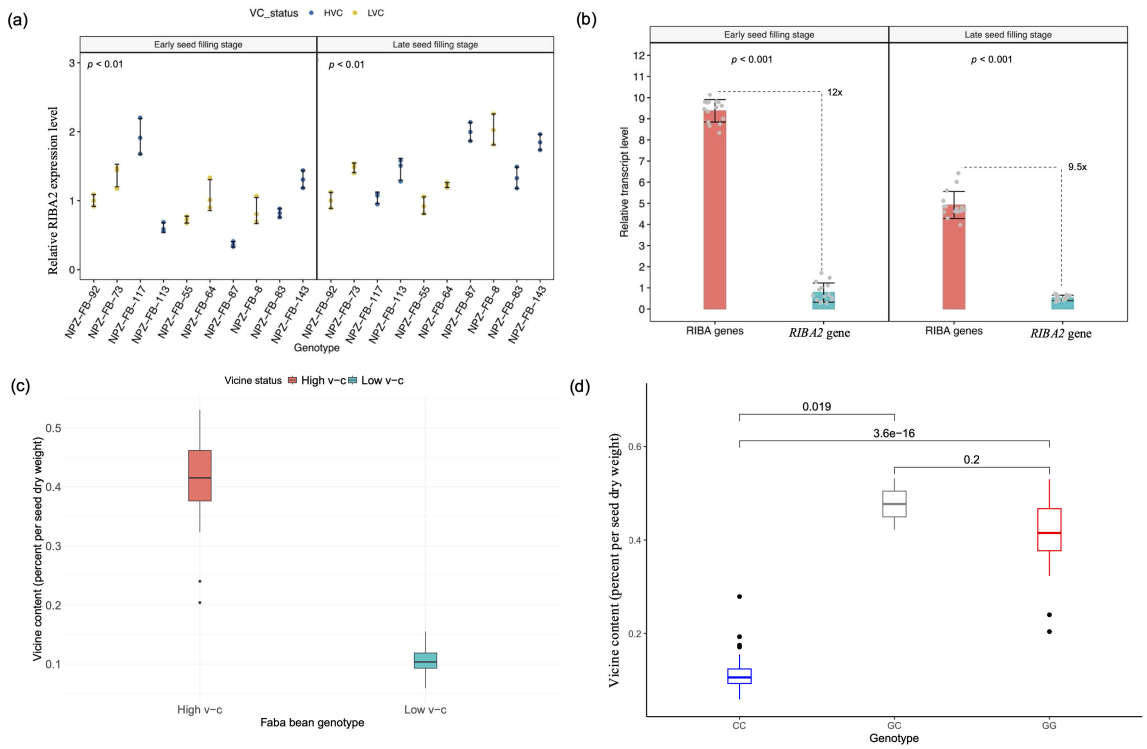


FIGURE 6
Analysis of *RIBA2* activity on vicine and convicine biosynthesis. **(a)** Relative expression levels of *RIBA2* across faba bean lines (means ± SD, n = 3). **(b)** Proportion of *RIBA2* expression relative to the combined RIBA gene (*VC1* and *RIBA2*) expression in high v-c lines at early and late seed filling stages (means ± SD, n = 15). Significance for comparisons between groups were determined using t-test. **(c)** Phenotypic differences in vicine content between high and low v-c faba bean genotypes **(d)** Allelic effects of a single nucleotide polymorphism within exon 6 of *RIBA2* gene in a panel of 97 diverse faba bean genotypes.

3.5 Implications of multiple *VC1* gene variants in molecular breeding for low vicine-convicine faba bean

It was observed that three different variants of *VC1* can be expressed. High v-c genotypes can express either *VC1A* or *VC1C*, while low v-c genotypes consistently express *vc1b*. Notably, *VC1C* shared close similarity with *vc1b*, having similar alleles at most SNP positions. This similarity caused genotypes expressing these variants to cluster under a major branch as observed in Figure 5c. Consequently, the existence of multiple *VC1* variants may present a significant challenge for marker-assisted selection of v-c content in faba bean breeding. Issues such as false heterozygous calls or incomplete segregation may arise due to the presence of multiple gene variants. To investigate this, we conducted genotyping assays using well-characterized low and high v-c genotypes. These assays involved single nucleotide polymorphic markers located within *VC1* gene, including some previously developed for v-c breeding. Since genomic DNA (gDNA) contains all gene variants, while complementary DNA (cDNA) only contains the expressed variant, we performed KASP genotyping assays with gDNA and cDNA as templates.

Our results revealed two key findings. Firstly, the presence of multiple variants could lead to bias in allele calls (Figures 7a, b; Table 1). In most cases, low v-c genotypes consistently exhibited false heterozygous signals when gDNA was used as a template. A similar trend can also occur in a few high v-c genotypes. However, when cDNA was used as the template, these genotypes were correctly identified as homozygous individuals. Likewise, in the assay targeting the 2 bp insertion, all low v-c genotypes were initially called as heterozygous when gDNA was used as the template, due to the presence of multiple copies, including both wild-type and mutant variants. However, these genotypes were correctly called as homozygous when cDNA was used as the template. Additionally, this observation further confirms that only one copy of *VC1* is expressed in each genotype analyzed.

Secondly, we observed that SNP alleles do not always segregate completely with the v-c phenotype. Low v-c alleles often segregate with some high v-c genotypes expressing *VC1C* variant. This results in the false clustering of these high v-c genotypes with low v-c genotypes, leading to inaccurate predictions of the v-c phenotype. (Figure 7b). This pattern was consistent with other *VC1*-based SNPs, except for the 2 bp insertion, which can efficiently distinguish the cultivars based on v-c contents. Beside this polymorphism, other SNPs within *VC1* were ineffective in predicting v-c contents in a diverse genetic background.

4 Discussion

4.1 *VC1* is multiallelic but exhibits single-copy expression

Until recently, the elucidation of genetic mechanisms underlying important traits in faba bean has been hindered by the unavailability of genomic tools for this crop, which has an enormous genome in which the largest chromosome is bigger than the entire human genome. This constraint equally impeded early endeavors to identify genes responsible for v-c biosynthesis in faba bean. In the absence of reference genomes, previous investigations relied on molecular markers generated by mapping mRNA contigs from faba bean to the genomes of closely related crops, such as *Medicago truncatula*, to identify regions controlling v-c content (Khazaei et al., 2015; Tacke et al., 2022). However, the involvement of *VC1* gene in v-c biosynthesis in faba bean was established recently and a frameshift mutation in this gene results in a loss-of-function, leading to a low v-c phenotype (Björnsdotter et al., 2021).

The availability of recently assembled *V. faba* reference genomes with functional gene annotations (Jayakodi et al., 2023) enabled us to map RIBA1 protein to these assemblies and identify all relevant RIBA genes associated with v-c biosynthesis. Our study

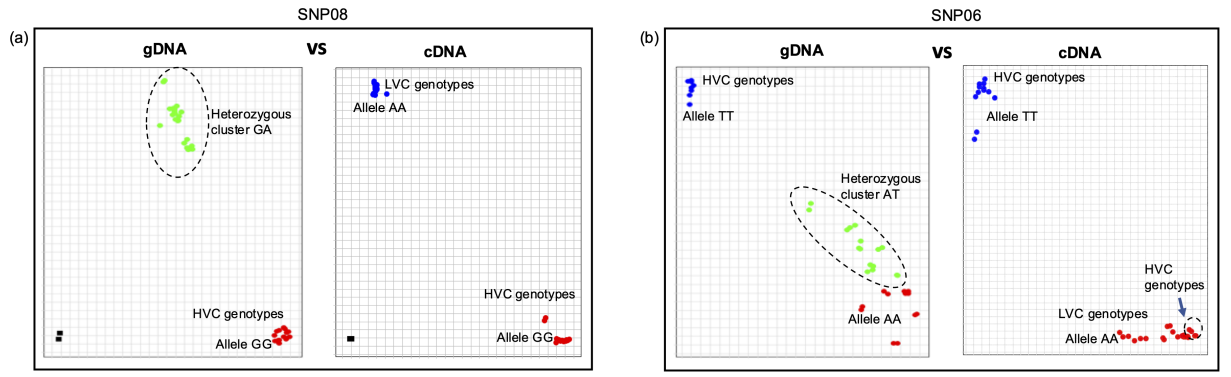


FIGURE 7
Multiple *VC1* copies can cause bias during marker analysis. Allelic discrimination plots for **(a)** 2 bp insertion in exon 6 and **(b)** an SNP within exon 5 of *VC1*. In both cases, the plots exhibited false heterozygous calls for all or most low vicine-convicine (v-c) (LVC) genotypes and some high v-c (HVC) genotypes when gDNA was employed as the template. In contrast, clear homozygous calls were observed with cDNA. Additionally, some high v-c genotypes expressing *VC1C* variant can cluster with low v-c genotypes, as observed in b (see arrow).

TABLE 1 Comparison of allele calls between gDNA and cDNA for SNPs detected by KASP Assays showed that multiple copies of *VC1* often lead to inaccurate allele calls.

Genotype	SNP08		SNP05		SNP06	
	gDNA	cDNA	gDNA	cDNA	gDNA	cDNA
NPZ-FB-117	GG	GG	CC	CC	TT	TT
NPZ-FB-113	GG	GG	CC	CC	TT	TT
NPZ-FB-87	GG	GG	CC	CC	TT	TT
NPZ-FB-83	GG	GG	AA	AA	AT	AA
NPZ-FB-143	GA	GG	AC	CC	AT	TT
NPZ-FB-92	GA	AA	AA	AA	AA	AA
NPZ-FB-73	GA	AA	AC	AA	AT	AA
NPZ-FB-55	GA	AA	AA	AA	AA	AA
NPZ-FB-64	GA	AA	AA	AA	AA	AA
NPZ-FB-8	GA	AA	AA	AA	AA	AA

Allele calls were also confirmed using cDNA sequences (see Supplementary Table S5).
High and low v-c genotypes are highlighted in red and yellow, respectively.

revealed multiple *VC1* variants, primarily characterized by mutations in intron 3, exon 3 and 6. We found that *VC1* is affected by copy number variations, and that all low v-c genotypes carry both functional and non-functional variants, although only the non-functional variant is expressed. Previously, only two allelic forms of *VC1* were known, where low v-c faba beans carried only mutant *vc1*, while high v-c faba beans carried the wild-type variant (Björnsdotter et al., 2021). These wild-type and mutant alleles correspond to *VC1A* and *vc1b* variants, respectively. Our identification of an additional variant, *VC1C*, expressed by some high v-c lines, underscores the diversity of *VC1* genes and highlights the significance of utilizing multiple genotypes in our study. However, a major structural variation in intron 3 across the three *VC1* variants did not appear to have functional relevance as it did not associate with the expression pattern.

We observed a differential expression of VC1, consistent with previous reports between high and low v-c genotypes (Björnsdotter et al., 2021). Surprisingly, copy number did not correlate with phenotypic expression. Although we observed a negative correlation between copy number and gene expression, this was due to the influence of low v-c genotypes sharing a single genetic source. These genotypes exhibited high copy numbers but lower expression levels. This observation was supported by cDNA sequencing which showed that each genotype expressed only one *VC1* variant despite having multiple variants in the genome. Previous reports on the transformation of hairy roots of Hedin with an additional copy of functional *VC1* did not result in increased vicine accumulation (Björnsdotter et al., 2021). This confirms that despite copy number variation, *VC1* is not sensitive to dosage.

The mechanisms involved in *VC1* dosage compensation are unclear; however, certain mechanisms have been proposed for genes with multiple copies. These mechanisms could involve complex processes at different transcription stages (Vaquerizas et al., 2009; Woodwark and Bateman, 2011), including microRNAs (miRNAs) that can activate or repress the transcription of duplicated or copy number variable (CNV) genes (Vasudevan et al., 2007; Li et al., 2008; Bartel, 2009; Woodwark and Bateman, 2011; Chang and Liao, 2012), imprinted or monoallelically expressed genes (Pauler et al., 2007), and DNA methylation (Suzuki and Bird, 2008; Law and Jacobsen, 2010).

The regulation of *VC1* differential expression between low and high v-c genotypes and the impact of the 2 bp insertion carried by low v-c genotypes on expression levels remain unclear. Differences in the expression of *VC1* genes can also stem from variations within the regulatory elements. Elements like enhancers and silencers have the potential to amplify or suppress the gene expression of the target gene (Bulger and Groudine, 2011; Kolovos et al., 2012). Upstream and downstream of *VC1* are various variants, such as short tandem repeats, InDels and SNPs. Structural differences within this region could potentially affect one or more regulatory elements. The interaction among these diverse regulatory components, their interplay with target promoters, and the involvement of epigenetic modifications can intricately regulate the expression of this gene (Bulger and Groudine, 2011; Kolovos et al., 2012; Cremer and Cremer, 2001; Maeda and Karch, 2011; Yang and Corces, 2011).

4.2 *RIBA2* is a functional RIBA locus in faba bean and a candidate for a minor effect v-c locus

The variation in *VC1* expression alone does not correlate with observed phenotypic differences among faba bean genotypes. Subsequent analysis of *VC1* cDNA in this study aligns with findings reported by Björnsdotter et al. (2021), that all low v-c cultivars express *vc1b* variant carrying 2 bp insertion in exon 6 which results in a non-functional protein. This implies that *VC1* is

not active in low v-c genotypes. However, as previously mentioned, this mutation only causes a significant reduction in v-c content but does not entirely eliminate it. For instance, the seed v-c content of the first low v-c genotype, initially reported by Duc et al. (1989) is approximately 0.04%, which is about 1/10 to 1/20 of high v-c contents. This substantial reduction in v-c content due to a single gene mutation highlights the major effect of *VC1* locus in v-c phenotype regulation and equally suggests the involvement of another gene with a minor effect.

In addition to the already known *VC1*, we identified a homologous gene, *RIBA2*, encoding putative bi-functional riboflavin proteins in faba bean genome. The distinct structure of the *RIBA2* gene compared to *VC1* suggests that it represents a second RIBA locus in faba bean. *RIBA2* is a bi-functional riboflavin gene with two catalytic domains, RibB and RibA, encoding DHBPS and GCHII enzymes, respectively. RIBA enzymes are known for their involvement in the riboflavin biosynthetic pathway in plants (Hiltunen et al., 2012). However, Björnsdotter et al. (2021) demonstrated that RIBA enzymes are also involved in v-c production in faba bean where v-c are synthesized in a three-step pathway that starts from the GTP cyclohydrolase II function of RIBA proteins. While *VC1* and *RIBA2* share nearly identical functional domains, *RIBA2* lacks the inactivating insertion present in *vc1b* mutant. Detailed analysis of *VC1* and *RIBA2* amino acid sequences showed conservation in all key amino acid residues necessary for catalytic activities, including those required for binding zinc ions essential for GCHII activity as demonstrated by Kaiser et al. (2002). These key amino acid residues are well conserved in *VC1* and *RIBA2* and other RIBA proteins within Fabaceae family but lacking in *vc1b*.

Subsequent analysis shows that *RIBA2* is expressed at much lower levels than *VC1* in seeds. Similarly, there was a significantly lower expression level relative to the combined riboflavin gene transcripts in high v-c genotypes, aligning with the observed phenotypic differences between low and high v-c genotypes. Previous studies consistently emphasize the *VC1* locus as a major determinant in v-c regulation (Ramsay et al., 1995; Gutierrez et al., 2006; Khazaei et al., 2015; Tacke et al., 2022). Our study substantiates this hypothesis, indicating several-fold higher expression of *VC1* relative to the homolog, *RIBA2*. These findings suggest that this homologous gene is a candidate for a minor effect locus for v-c content. The expression of this minor effect gene may explain why mutation in *vc1* does not eliminate v-c completely in affected genotypes. Tacke et al. (2022) used a transcript-based annotation method to identify a differentially expressed contig associated with v-c contents, as well as another, non-differentially expressed contig which mapped to the *VC1* locus. However, they could not fully decipher the locus due to the lack of genomic data.

It is essential to validate the involvement of *RIBA2* in v-c biosynthesis through mutant analysis, particularly in a *vc1* background. However, the lack of reliable transformation methods and mutant resources in faba bean significantly hinders functional gene validation through knockouts (O'Sullivan and Angra, 2016). Moreover, given that riboflavin biosynthesis is essential for fundamental cellular processes and overall plant development, it remains uncertain whether simultaneous loss-of-function mutations in both *VC1* and *RIBA2* would be viable.

4.3 The complexity of *VC1* locus limits the efficiency of localized SNPs for marker-assisted selection in v-c breeding

Various molecular markers have been developed to facilitate breeding of low v-c faba beans (Tacke et al., 2022; Khazaei et al., 2017). However, most existing v-c markers target polymorphisms within *VC1* genes. While these markers have proven valuable in some contexts, their efficiencies may be limited due to complex nature of *VC1*. Our research revealed that multiple *VC1* copies can lead to inaccurate allele calls, resulting in false clusters during marker analysis. This often happens because of the existence of closely homologous sequences in the genome (Makhoul et al., 2020). Consequently, conventional KASP assays may require optimization to improve accuracy (Makhoul et al., 2020). Nevertheless, the applicability of these SNPs might be confined to specific genetic backgrounds. In a diverse genetic background, comprising all expressed *VC1* variants, SNPs often do not segregate fully with the phenotype. Hence, these polymorphisms may not be efficient for MAS of v-c in faba bean breeding. Therefore, we strongly advise caution when utilizing *VC1*-based molecular markers for v-c content selection in breeding programs. Overreliance on only these markers for selection may inadvertently lead to incorrect prediction of v-c contents.

However, our study highlights the functional significance of the 2 bp insertion in exon 6 as a reliable polymorphism in *VC1* that effectively distinguishes genotypes based on the v-c phenotype. Optimizing the KASP assay targeting this polymorphism could significantly enhance its efficiency and specificity, particularly given the potential for inaccurate clustering due to multiple gene copies. Moreover, we have shown that SNPs within *RIBA2* show consistent segregation with v-c contents. These SNPs can accurately predict v-c content without bias and will be valuable for faba bean breeding.

5 Conclusion

We have demonstrated that *VC1* exists in multiple copies and shows CNV. However, copy number does not correlate with gene expression and suggests a tight regulation of the gene. Multiple *VC1* variants were expressed among low and high v-c genotypes which complicates molecular marker development for breeding. We also identified a diverging homolog of *VC1*, *RIBA2*, which shares nearly identical RIBA domains with *VC1*. Our results show that *RIBA2* has two functional domains similar to the active *VC1* genes, with highly conserved key amino acid residues, indicating well-conserved roles as RIBA proteins. Our findings suggest that this homologous gene is a candidate minor effect v-c locus, and its involvement can be validated using gene editing knockouts.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

SU: Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. MM: Methodology, Writing – review & editing, Formal analysis. AG: Funding acquisition, Resources, Writing – review & editing. CO: Methodology, Writing – review & editing, Formal analysis. RS: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The work was supported by DFG grant 469336000 to RJS from the German Research Society (DFG) for the International Research Training Group IRTG 2843 Accelerating Crop Genetic Gain. DFG grant 497667402 to AAG.

Acknowledgments

The authors gratefully acknowledge Stavros Tzigos and Annette Plank for their valuable technical assistance with nucleic acid extraction and plant management in the greenhouse, respectively. We thank Erick Owuor Mikwa for his insightful comments and suggestions. We also express our gratitude to Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ, Hohenlieth, Germany) for providing the faba bean seeds used in the study.

Conflict of interest

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

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1565210/full#supplementary-material

References

Adhikari, K., Khazaei, H., Ghaouti, L., Maalouf, F., Vandenberg, A., Link, W., et al. (2021). Conventional and molecular breeding tools for accelerating genetic gain in faba bean (Vicia faba L.). *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.744259

Badjona, A., Bradshaw, R., Millman, C., Howarth, M., and Dubey, B. (2023). Faba bean flavor effects from processing to consumer acceptability. *Foods* 12, 2237. doi: 10.3390/foods12112237

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002

Björnsdotter, E., Nadzieja, M., Chang, W., Escobar-Herrera, L., Mancinotti, D., Deepti Angra, D., et al. (2021). VC1 catalyses a key step in the biosynthesis of vicine from GTP in faba bean. *Nat. Plants* 7, 923–931. doi: 10.1038/s41477-021-00950-w

Borisjuk, L., Weber, H., Panitz, R., Manteuffel, R., and Wobus, U. (1995). Embryogenesis of V. faba L.: Histodifferentiation in relation to starch and storage protein synthesis. *Plant Physiol.* 147, 203–218. doi: 10.1016/S0176-1617(11)81507-5

Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. *Cell* 144, 327–339. doi: 10.1016/j.cell.2011.01.024

Caracuta, V., Weinstein-Evron, M., Kaufman, D., Yeshurun, R., Silvent, J., and Boaretto, E.. (2016). 14,000-year-old seeds indicate the Levantine origin of the lost progenitor of faba bean. *Sci. Rep.* 6, 37399. doi: 10.1038/srep37399

Carrillo-Perdomo, E., Vidal, A., Kreplak, J., Duborjal, H., Leveugle, M., Duarte, J., et al. (2020). Development of new genetic resources for faba bean (Vicia faba L.) breeding through the discovery of gene-based SNP markers and the construction of a high-density consensus map. *Sci. Rep.* 10, 6790. doi: 10.1038/s41598-020-63664-7

Chang, A. Y., and Liao, B. Y. (2012). DNA methylation rebalances gene dosage after mammalian gene duplications. *Mol. Biol. Evol.* 29, 133–144. doi: 10.1093/molbev/msr174

Choudhary, D. K., and Mishra, A. (2019). *In vitro* and in silico interaction of faba bean (Vicia faba L.) seed extract with xanthine oxidase and evaluation of antioxidant activity as a therapeutic potential. *Nat. Prod. Res.* 33, 2689–2693. doi: 10.1080/14786419.2018.1460831

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* 2, 292–301. doi: 10.1038/35066075

Doyle, J. J., and Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus* 12, 13–15.

Duc, G., Bao, S., Baum, M., Redden, B., Sadiki, M., Suso, M. J., et al. (2010). Diversity maintenance and use of Vicia faba L. genetic resources. *Field Crops Res.* 115, 270–278. doi: 10.1016/j.fcr.2008.10.003

Duc, G., Marget, P., Esnault, R., Le Guen, J., and Bastianelli, D. (1999). Genetic variability for feeding value of faba bean seeds (Vicia faba): comparative chemical composition of isogenics involving zero-tannin and zero-vicine genes. *J. Agric. Sci.* 133, 185–196. doi: 10.1017/S0021859699006905

Duc, G., Sixdenier, G., Lila, M., and Furstoss, V. (1989). "Search of genetic variability for vicine and convicine content in V. faba L.: A first report of a gene which codes for nearly zero-vicine and zero-convicine contents," in *International Workshop on 'Antinutritional Factors (ANF) in Legume Seeds'* (Pudoc, Wageningen).

FAOSTAT (2022). *Food and Agriculture Organization of the United Nations*. Available online at: http://faostat.fao.org (Accessed November 22, 2022).

Gutierrez, N., Avila, C. M., Duc, G., Marget, P., Suso, M. J., Moreno, M. T., et al. (2006). CAPs markers to assist selection for low vicine and convicine contents in faba bean (Vicia faba L.). *Theor. Appl. Genet.* 114, 59–66. doi: 10.1007/s00122-006-0410-3

Gutierrez, N., Giménez, M. J., Palomino, C., and Avila, C. M. (2011). Assessment of candidate reference genes for expression studies in Vicia faba L. by real-time quantitative PCR. *Mol. Breed.* 28, 13–24. doi: 10.1007/s11032-010-9456-7

Hiltunen, H.-M., Illarionov, B., Hedtke, B., Fischer, M., and Grimm, B. (2012). Arabidopsis RIBA proteins: two out of three isoforms have lost their bifunctional activity in riboflavin biosynthesis. *Int. J. Mol. Sci.* 13, 14086–14105. doi: 10.3390/ijms131114086

Jayakodi, M., Golicz, A. A., Kreplak, J., Fechete, L. I., Angra, D., Bednář, P, et al. (2023). The giant diploid faba genome unlocks variation in a global protein crop. *Nature* 615, 652–659. doi: 10.1038/s41586-023-05791-5

Kaiser, J., Schramek, N., Eberhardt, S., Puttmer, S., Schuster, M., and Bacher, A. (2002). Biosynthesis of vitamin B2. *Eur. J. Biochem.* 269, 5264–5270. doi: 10.1046/j.1432-1033.2002.03239.x

Kambal, A. E., Bond, D. A., and Toynbee-Clarke, G. (1976). A study on the pollination mechanism in field beans (Vicia faba L.). *J. Agric. Sci.* 87, 519–526. doi: 10.1017/S0021859600033128

Karkanis, A., Ntatsi, G., Lepse, L., Fernández, J. A., Vågen, I. M., Rewald, B., et al. (2018). Faba bean cultivation–revealing novel managing practices for more sustainable and competitive European cropping systems. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01115

Khamassi, K., Ben Jeddi, F., Hobbs, D., Irigoyen, J., Stoddard, F., O'Sullivan, D. M, et al. (2013). A baseline study of vicine–convicine levels in faba bean (Vicia faba L.) germplasm. *Plant Genet. Resour.* 11, 250–257. doi: 10.1017/S1479262113000105

Khazaei, H., O'Sullivan, D. M., Jones, H., Pitts, N., Sillanpää, M. J., Pärssinen, P., et al. (2015). Flanking SNP markers for vicine-convicine concentration in faba bean (Vicia faba L.). *Mol. Breed.* 35, 38. doi: 10.1007/s11032-015-0214-8

Khazaei, H., Purves, R. W., Song, M., Stonehouse, R., Bett, K. E., Stoddard, F. L., et al. (2017). Development and validation of a robust, breeder-friendly molecular marker for the vc - locus in faba bean. *Mol. Breed.* 37, 140. doi: 10.1007/s11032-017-0742-5

Kolovos, P., Knoch, T. A., Grosveld, F. G., Cook, P. R., and Papantonis, A. (2012). Enhancers and silencers: an integrated and simple model for their function. *Epigenet. Chromatin* 5, 1. doi: 10.1186/1756-8935-5-1

Law, J. A., and Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220. doi: 10.1038/nrg2719

Li, J., Musso, G., and Zhang, Z. (2008). Preferential regulation of duplicated genes by microRNAs in mammals. *Genome Biol.* 9, R132. doi: 10.1186/gb-2008-9-8-r132

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using realtime quantitative PCR and the 2ΔΔC(T) Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Luzzatto, L., and Arese, P. (2018). Favism and glucose-6-phosphate dehydrogenase deficiency. *N Engl. J. Med.* 378, 60–71. doi: 10.1056/nejmra1708111

Maeda, R. K., and Karch, F. (2011). Gene expression in time and space: additive vs hierarchical organization of cis-regulatory regions. *Curr. Opin. Genet. Dev.* 21, 187–193. doi: 10.1016/j.gde.2011.01.021

Makhoul, M., and Obermeier, C. (2022). Development of breeder-friendly KASP markers from genome-wide association studies results. *Methods Mol. Biol.* 2481, 287–310. doi: 10.1007/978-1-0716-2237-7_17

Makhoul, M., Rambla, C., Voss-Fels, K. P., Hickey, L. T., Snowdon, R. J., and Obermeier, C. (2020). Overcoming polyploidy pitfalls: a user guide for effective SNP conversion into KASP markers in wheat. *Theor. Appl. Genet.* 133, 2413–2430. doi: 10.1007/s00122-020-03608-x

O'Sullivan, D. M., and Angra, D. (2016). Advances in faba bean genetics and genomics. *Front. Genet.* 22. doi: 10.3389/fgene.2016.00150

Pauler, F. M., Koerner, M. V., and Barlow, D. P. (2007). Silencing by imprinted noncoding RNAs: is transcription the answer? *Trends Genet.* 23, 284–292.

Puspitasari, W., Allemann, B., Angra, D., Appleyard, H., Ecke, W., Möllers, C., et al. (2022). NIRS for vicine and convicine content in faba bean seeds enabled a GWAS in preparation for a marker-assisted adjustment of seed quality of German winter faba beans. *J. für Kulturpflanzen* 74, 1–10. doi: 10.5073/JfK.2022.01-02.01

Ramsay, G., van de Ven, W., Waugh, R., Griffi ths, D. W., and Powel, W. (1995). "Mapping quantitativetrait loci in faba beans," in *AEP (ed) 2nd Eur Conf on Grain Legume* (Copenhagen, AEP, Paris, France), 444—445.

Sixdenier, G., Cassecuelle, F., Guillaumin, L., and Duc, G. (1996). Rapid spectrophotometric method for reduction of vicine and convicine in faba bean seed. *FABIS Newslett.* 38, 42–44.

Stagnari, F., Maggio, A., Galieni, A., and Pisante, M. (2017). Multiple benefits of legumes for agriculture sustainability: an overview. *Chem. Biol. Technol. Agric.* 4, 2. doi: 10.1186/s40538-016-0085-1

Suzuki, M. M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* 9, 465–476. doi: 10.1038/nrg2341

Tacke, R., Ecke, W., Höfer, M., Sass, O., and Link, W. (2022). Fine-mapping of the major locus for vicine and convicine in faba bean (Vicia faba) and marker-assisted breeding of a novel, low vicine and convicine winter faba bean population. *Plant Breed* 141, 644–657. doi: 10.1111/pbr.13039

Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., and Luscombe, N. M. (2009). A census of human transcription factors: function, expression and evolution. *Nat. Rev. Genet.* 10, 252–263. doi: 10.1038/nrg2538

Vasudevan, S., Tong, Y., and Steitz, J. A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. *Science* 318, 1931–1934. doi: 10.1126/science.1149460

Warsame, A. O., Michael, N., O'Sullivan, D. M., and Tosi, P. (2020). Identification and quantification of major faba bean seed proteins. *J. Agri Food Chem.* 68, 8535–8544. doi: 10.1021/acs.jafc.0c02927

Woodwark, C., and Bateman, A. (2011). The characterisation of three types of genes that overlie copy number variable regions. *PLoSONE* 6, e14814. doi: 10.1371/journal.pone.0014814

Yang, J., and Corces, V. G. (2011). Chromatin insulators: a role in nuclear organization and gene expression. *Adv. Cancer Res.* 110, 43–76. doi: 10.1016/B978-0-12-386469-7.00003-7

3 Novel gene-copy-specific KASP markers for efficient marker-assisted development of low vicine faba bean (*Vicia faba* L.)

Ugwuanyi, S., Tietgen, H., Makhoul, M., Obermeier, C. and Snowdon, R. (2025). *Plant Breeding*, 144(4): 573–579. https://doi.org/10.1111/pbr.13276

Plant Breeding

WILEY

Plant Breeding

ORIGINAL ARTICLE OPEN ACCESS

Novel Gene-Copy-Specific KASP Markers for Efficient Marker-Assisted Development of Low Vicine Faba Bean (*Vicia faba* L.)

Samson Ugwuanyi[1,2] | Hanna Tietgen[3] | Manar Makhoul[1] | Christian Obermeier[1] | Rod J. Snowdon[1]

[1]Department of Plant Breeding, Justus Liebig University, Giessen, Germany | [2]Department of Plant Science and Biotechnology, University of Nigeria, Nsukka, Nigeria | [3]NPZ Innovation GmbH, Holtsee, Germany

Correspondence: Samson Ugwuanyi (samson.ugwuanyi@agrar.uni-giessen.de)

Received: 19 November 2024 | **Revised:** 23 January 2025 | **Accepted:** 3 March 2025

Funding: This work was supported by Deutsche Forschungsgemeinschaft, 469336000

Keywords: convicine | marker-assisted selection | molecular breeding | *RIBA2* | *VC1* | *Vicia faba*

ABSTRACT

Vicine and convicine (v-c) pose significant challenges to the global use of faba beans, causing favism in susceptible individuals. Developing low v-c faba beans is crucial to enhance their utilization and safety, but progress has been hindered by insufficient genomic tools for marker-assisted selection. This challenge is linked to the complex nature of the responsible locus, which makes marker development difficult. While a two-base pair insertion/deletion in *VC1* correlates with v-c phenotype, converting this variant into a reliable marker remains challenging due to multiple *VC1* copies. Here, we developed a set of five robust, breeder-friendly selection markers targeting *VC1* and a diverging homologue, *RIBA2*, for v-c breeding. The marker SNP08 was designed based on the 2 bp inactivating insertion in *VC1*. To overcome the bias caused by unexpressed homologous *VC1* copies, we employed a target-specific preamplification strategy, incorporating a preliminary PCR step to amplify specifically the *VC1* copy harbouring the mutation associated with v-c. This was followed by genotyping using a Kompetitive allele-specific PCR (KASP) assay, resulting in a robust *VC1*-copy-specific KASP marker capable of predicting v-c phenotypes with 100% accuracy. Additionally, we developed four simple, easy-to-use KASP markers based on single nucleotide polymorphisms (SNPs) within the single-copy neighbouring *VC1* homologue, *RIBA2*. Subsequent validation in a diverse panel showed consistent co-segregation with v-c phenotypes. The results demonstrated high accuracy of the markers in predicting v-c phenotypes, highlighting their suitability for marker-assisted development of low v-c faba beans. These genomic tools are valuable for faba bean breeding and will facilitate the development of high-quality faba bean protein.

1 | Introduction

Marker-assisted selection (MAS) is an important breeding tool that leverages trait-linked variants, such as single nucleotide polymorphisms (SNPs) and small insertions/deletions (InDels), to rapidly and cost-effectively develop cultivars. Several marker-trait association studies have significantly contributed to the facilitation of MAS by identifying significant associations between markers and traits (Raatz et al. 2019; Ugwuanyi et al. 2022). In modern breeding for simply inherited traits, trait-linked variants are frequently converted into reliable and breeder-friendly Kompetitive allele-specific PCR (KASP) assays for use in MAS. KASP, a fluorescence-based genotyping platform, is cost-effective and convenient for genotyping a low number of SNPs across a large set of genotypes. It has become a preferred choice due to its high throughput and convenience (Semagn et al. 2014). Converting SNPs to KASP markers is straightforward for polymorphisms within single-copy genes

or genomic regions unaffected by duplications. However, this conversion becomes very challenging in complex genomes, such as polyploids with homeologous or repetitive sequences, due to polymorphisms between homeologous or highly similar sequences from subgenomes. Molecular markers targeting such variations often exhibit heterozygous genotype clusters in homozygous individuals, rendering them unsuitable for selection (Makhoul et al. 2020). This limitation is also prevalent in diploid crops with complex and highly repetitive genomes, such as faba bean (*V. faba* L.), which has an enormous genome comprising more than 85% repetitive sequences (Novák et al. 2020; Jayakodi et al. 2023).

In faba bean, implementing MAS for low vicine and convicine (v-c) genotypes remains challenging despite known polymorphisms associated with variations in v-c phenotypes. Breeding for v-c content is an essential objective in faba bean because these antinutrients can induce the disease favism in individuals with glucose-6-phosphate dehydrogenase (G6PD) deficiency (Cappellini and Fiorelli 2008; Choudhary and Mishra 2019; Badjona et al. 2023). Additionally, v-c-rich faba beans can negatively impact feed conversion efficiency and productivity when included in livestock diets (Vilariño et al. 2009; Koivunen et al. 2014). Since the discovery of a low v-c (lvc) allele (Duc et al. 1989), which is associated with a more than 95% reduction in seed v-c levels and is safe for susceptible individuals, breeding efforts have focused on developing cultivars with low v-c content. However, a major challenge in breeding lies in developing reliable selection markers capable of efficiently distinguishing the low v-c phenotype, due to the complex nature of the v-c locus. *VC1* is a major locus on chromosome 1 of faba bean, identified to initiate the biosynthetic pathway (Björnsdotter et al. 2021). However, exploitation of SNP variants within this locus as selection markers (Khazaei et al. 2017; Tacke et al. 2022) has not been reliable for commercial breeding due to multiple tandem gene duplications involving this locus. We previously demonstrated the existence of several homologous *VC1* gene copies that remain unexpressed (Ugwuanyi et al. 2024). Consequently, SNPs within these pseudogenes do not segregate with v-c phenotypes and often bias marker analysis. Furthermore, a 2-bp insertion in *VC1*, known to be causal for the low v-c phenotype, was identified as a functional polymorphism suitable for selection in breeding programmes (Björnsdotter et al. 2021; Ugwuanyi et al. 2024). However, converting this genetic variant into an affordable, rapid and breeder-friendly marker is challenging, as some genotypes carry both mutant and wild type alleles, although only one allele is expressed and determines the v-c phenotype.

Similar challenges exist in other crops with complex genomes, where highly similar sequences or homologous sequences may impede reliable marker development (Wang et al. 2014; Makhoul et al. 2020). As a result, developing KASP markers for a trait based on such SNPs remains difficult, posing a significant challenge to the exploitation of important variants for MAS. Additionally, the *RIBA2* gene, a neighbouring homologue of *VC1* on chromosome 1, harbours SNPs that could be exploited for selection (Ugwuanyi et al. 2024). Currently, methods for v-c estimation in breeding programmes are either unreliable, as with available selection markers, or expensive and laborious, as with biochemical assays. Therefore, rapid, reliable and cost-effective molecular markers are highly desirable for an affordable and efficient development of low v-c faba beans. As a result, our study aimed to facilitate efficient and robust MAS in v-c breeding by (1) developing a robust *VC1*-copy-specific KASP marker using a target-specific preamplification (TPA) KASP strategy to overcome the challenge posed by homologous sequences and (2) developing four additional cost-effective and breeder-friendly KASP markers within the *RIBA2* gene and detailing their application across different faba bean populations.

2 | Materials and Methods

2.1 | Plant Materials and DNA Isolation

A subset of 14 diverse faba bean lines showing variations in seed v-c content were initially used to perform SNP-phenotype analysis to find co-segregating SNP variants (Table S1). Subsequent marker validation was conducted using a diverse set of 97 faba bean genotypes, including breeding lines and commercial cultivars, obtained from the commercial breeding company, Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ, Hohenlieth, Germany) (Table S2). These genotypes are known to show varying levels of seed v-c content. Seeds from each genotype were grown in 4-L plastic pots in the greenhouse. The plants were self-pollinated at flowering, and mature seeds originating from these plants were evaluated for vicine phenotypes. Vicine content was determined by spectrophotometry at a wavelength of 274 nm on a solution of two seeds soaked in water and diluted with HCl, according to the method of Sixdenier et al. (1996).

For DNA sampling, fresh leaves were collected from the first fully opened leaves of the plants. The fresh leaves were ground to a fine powder using TissueLyser II (Qiagen, Hilden Germany). Genomic DNA from each sample was isolated from the leaf powder following the Doyle and Doyle (1990) method and diluted to a working concentration of 20–30 ng/μL.

2.2 | Development of a TPA KASP Assay for V-C Selection

To improve the accuracy of allele calls for trait-linked variants within complex regions, we developed a TPA KASP assay. TPA-KASP is an optimized two-step KASP assay that integrates conventional PCR with standard allele-specific PCR within a single assay plate. In the initial step, we preamplified the target using a pair of PCR primers designed to selectively amplify only the target gene or region harbouring SNPs associated with the phenotype, followed by genotyping the preamplified target using a KASP assay in the second step (see Figure S1).

Specifically, in the first step, we designed a pair of PCR primers that specifically bind to only the expressed *VC1* variants as well as KASP assay primers (Figure 1a; Table 1). The PCR reaction was optimized in a final volume of 2.5 μL, comprising 1.25 μL of master mix (Colourless Master Mix, Promega), 0.125 μL each of forward and reverse primers and 1 μL of DNA pre-diluted to about 5 ng/μL. The reaction mix was pipetted into a 384-well plate and sealed. The PCR amplification was performed

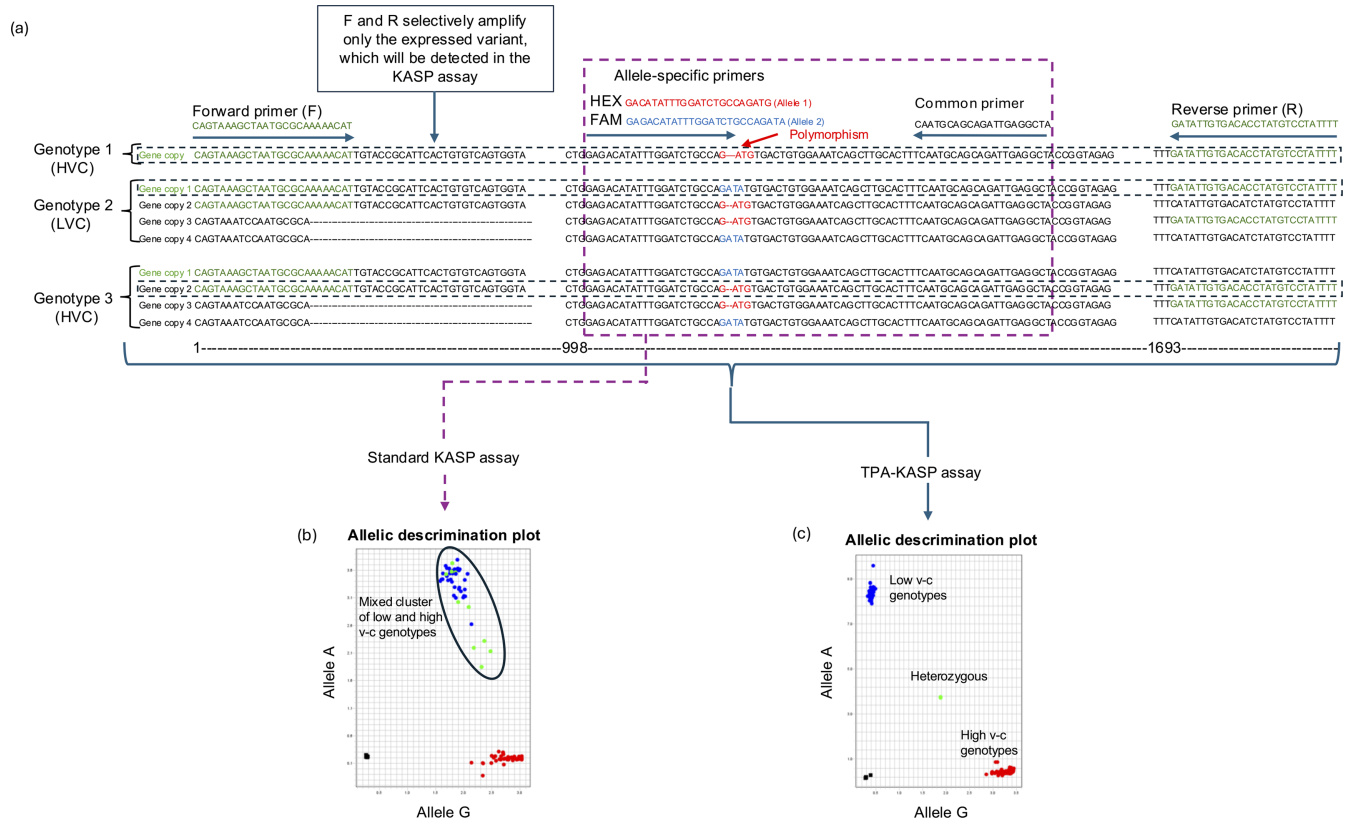


FIGURE 1 | Utilizing the TPA-KASP method to develop a reliable *VC1*-copy-specific marker for v-c breeding. (a) An alignment showing the *VC1* gene target region, which covers the polymorphism targeted for KASP genotyping. Forward (F) and reverse (R) primers are designed to span approximately 1.7 kb, including flanking regions for the insertion (see red arrow). F and R selectively amplify only the expressed gene variant, as indicated by the dotted boxes. (b) Allelic discrimination plot for the marker showing result from the standard KASP assay. The assay shows false heterozygous cluster comprising low and high v-c genotypes. (c) Allelic discrimination plot for the marker showing result from the TPA-KASP assay, with clear homozygous clusters corresponding to low (A allele) and high v-c genotypes (G allele).

TABLE 1 | *VC1*-copy-specific KASP primers for v-c phenotype prediction.

Marker ID	Primer ID	Primer sequences	Pre-PCR primers
SNP08	SNP08_Hex	gaaggtcggagtcaacggattGACATATTTGGATCTGCCAGATG	CAGTAAAGCTAATGCGCAAAAACAT[a] AAAATAGGACATAGGTGTCACAATATC[b]
	SNP08_Fam	gaaggtgaccaagttcatgctgaGACATATTTGGATCTGCCAGATA	
	SNP08_Com	TAGCCTCAATCTGCTGCATTG	

[a]forward primer.
[b]Reverse primer; lowercase letters represent the sequence of HEX and FAM tail.

following the thermal conditions: initial denaturation at 94°C for 2 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 61°C for 30 s, and elongation at 74°C for 2 min; followed by final elongation at 74°C for 5 min. After the PCR reaction was complete, the plate was removed and carefully unsealed.

Next, in the genotyping step, the KASP assay primer was prepared by mixing 46 μL of Milli-Q water, 30 μL of a common primer (100 μM) and 12 μL of each allele-specific primer (100 μM). For each KASP reaction, 0.08 μL of the primer mixture was combined with 2.5 μL of PACE-IR Genotyping Master Mix (low ROX, 3CR Bioscience, Harlow, UK) and added to 2.5 μL of PCR-amplified product on the 384-well plate, resulting in a final volume of 5 μL. Amplification was performed using the ViiA7 real-time PCR system (Life Technologies, USA) under the following conditions: 95°C for 15 min, followed by up to 40 cycles of 94°C for 20 s and 57°C for 1 min, according to Makhoul and Obermeier (2022).

2.3 | Development and Validation of *RIBA2* KASP Markers

Complete gene sequences of *RIBA2* from nine genotypes, including genotypes with high and low v-c content, were aligned using UGENE software to identify SNPs or small Indels. Each of these variants was converted into KASP markers by

designing two allele-specific primers and one common primer (Table 2). The KASP reaction for the *RIBA2* markers was conducted as previously described, but using genomic DNA instead of a PCR-amplified product. Markers that showed significant co-segregation with the v-c phenotype were subsequently validated using a diversity panel of 97 faba bean genotypes. For each marker, triplicates of water samples were used as a no-template control, and two replicates per sample were included. All marker assays were repeated at least twice. Detailed information on validated *RIBA2* KASP markers can be found in Table 2.

2.4 | Statistical Analysis

Statistical tests for significance were performed using R software 4.1 (R Core Team, 2021) to infer marker haplotype effects on v-c phenotypes. Differences were considered statistically significant at $p < 0.05$.

3 | Results

3.1 | A Robust *VC1*-Copy-Specific KASP Marker Using TPA Method for Efficient Marker-Assisted Development of Low V-C Faba Beans

Using a TPA strategy, we developed a reliable gene-copy-specific KASP marker, SNP08, for v-c selection based on the 2-bp nonsense mutation within the *VC1* gene. We observed that direct application of KASP assay for genotyping this mutation results in a mixed cluster of false heterozygous genotype calls in homozygous individuals comprising both low v-c and high v-c genotypes as shown in Figure 1b. We overcame this challenge by leveraging unique sequence variants within Intron 3 and a region downstream of the gene and selectively amplified only the expressed gene variant. The resulting amplicons were then picked up by KASP assay primers in the next step, producing accurate allele calls for the marker (TPA-KASP) (Figure 1c). The association of the allelic and phenotypic data is presented in Table 3.

Analysis of the segregation pattern of this marker perfectly correlated with v-c content in a diverse faba bean population. Unlike the direct application of the KASP assay, which resulted in false heterozygous calls (Figure 2a), the TPA-KASP approach accurately clustered high and low v-c genotypes separately as homozygous genotypes as observed in Figure 2b. The G allele, representing the absence of insertion (wild type), co-segregated with high v-c, while the A allele, representing the presence of insertion (mutant), co-segregated with the low v-c phenotype. This confirmed that the marker can distinguish genotypes based on v-c phenotypes with 100% accuracy in diverse faba bean populations, making it a robust tool for MAS in faba bean breeding programmes.

3.2 | MAS of Vicine and Convicine Content in Different Faba Bean Breeding Populations Using *RIBA2* KASP Markers

To provide more tools for v-c selection, we developed four additional cost-effective and breeder-friendly KASP markers, SNP_b1, SNP_b2, SNP_pr1 and SNP_pr2, within the *RIBA2* gene. Validation in a diverse panel showed that the markers segregated into two distinct haplotypes: Haplotype 1 and Haplotype 2. The

TABLE 3 | Association of SNP08 marker alleles with v-c phenotypes in a diverse set of 96 faba bean genotypes.

Genotype	Number of individuals	Mean vicine content	v-c status
GG	52	0.414	high v-c
AA	44	0.115	low v-c

TABLE 2 | List of *RIBA2* KASP markers for selection in v-c breeding.

Marker ID	Primer ID	Primer sequences
SNP_b1	SNP_b1_hex	gaaggtcggagtcaacggattGAACATCTTCTCCATCTCCAATG
	SNP_b1_fam	gaaggtgaccaagttcatgctGAACATCTTCTCCATCTCCAATA
	SNP_b1_com	TTTGCAACACTTATGGATTTTGA
SNP_b2	SNP_b2_hex	gaaggtcggagtcaacggattCGGCAGCCTCAATCTGTTC
	SNP_b2_fam	gaaggtgaccaagttcatgctCGGCAGCCTCAATCTGTTG
	SNP_b2_com	CTGCCAGATGTGACTGTGGA
SNP_pr1	SNP_pr1_hex	gaaggtcggagtcaacggattAACACAACTGAGAGGTGAGTGTAAT
	SNP_pr1_fam	gaaggtgaccaagttcatgctAACACAACTGAGAGGTGAGTGTAAA
	SNP_pr1_com	TGAGCTTCGCTTCCATTTG
SNP_pr2	SNP_pr2_hex	gaaggtcggagtcaacggattGCTTCCATTTGTGTGGAAATAGT
	SNP_pr2_fam	gaaggtgaccaagttcatgctGCTTCCATTTGTGTGGAAATAGA
	SNP_pr2_com	AAAAACACAACTGAGAGGTGAGTG

Note: Lowercase letters represent the sequence of HEX and FAM tail.

haplotypes correlated significantly with v-c phenotypes, with Haplotype 1 co-segregating with high v-c and Haplotype 2 with low v-c (Figure 3a), indicating high prediction accuracy. The association of the allelic and phenotypic data is presented in Table 4.

However, we found a few outliers with high v-c phenotypes within the low v-c cluster for these markers. These outlying genotypes correspond to *VC1C* genotypes (Haplotype 3) and differ from other high v-c genotypes corresponding to *VC1A* genotypes, as described in Ugwuanyi et al. (2024) (results not shown). While both groups of high v-c variants differ in all SNP loci within *VC1*, they share a common polymorphism—the absence of the 2-bp insertion in the expressed *VC1* gene. In contrast, the *VC1C* variant is similar to the low v-c variant, except

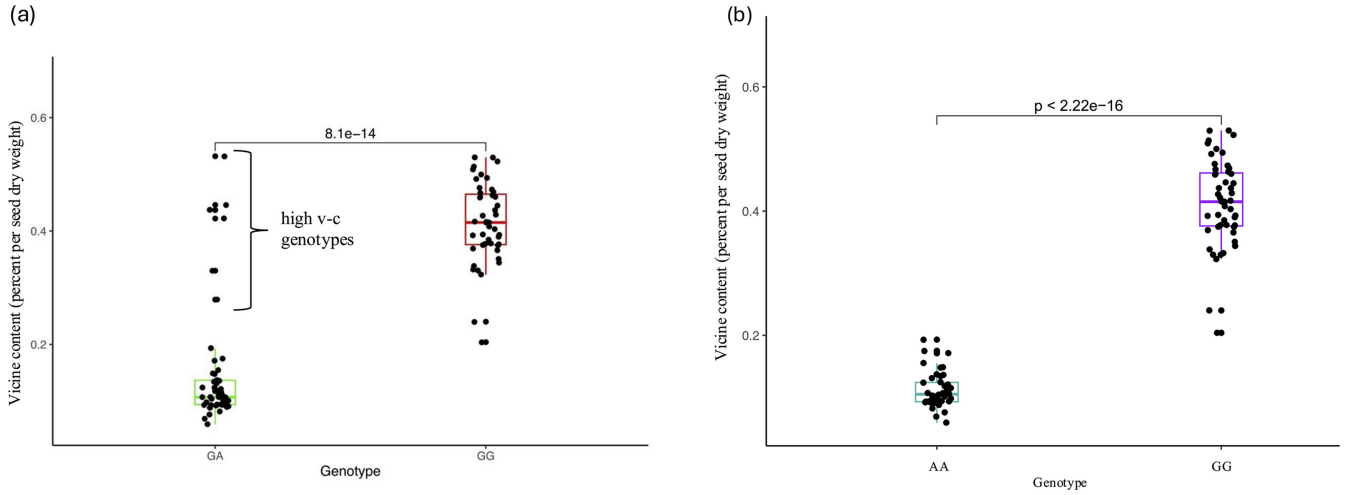


FIGURE 2 | Validation of a novel *VC1*-copy-specific KASP marker, SNP08, for marker-assisted development of low v-c faba beans. (a) A boxplot showing the allelic effect of the variant using standard KASP assay, where both high and low v-c genotypes cluster as heterozygotes due to non-specific signals from all copies with and without the insertion (b) A boxplot illustrating the allelic effect of the SNP08 marker on v-c content using target-specific preamplification method (TPA-KASP), which is specific to the expressed gene copy.

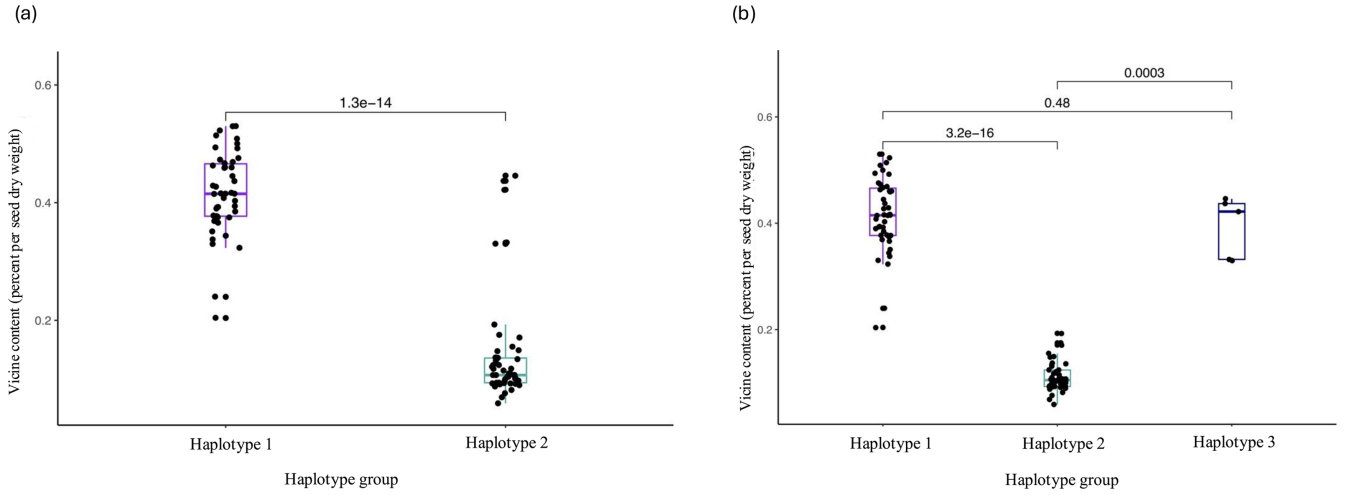


FIGURE 3 | Evaluation of the prediction accuracy of five v-c KASP markers, one amplifying from the *VC1* gene and four amplifying from the neighbouring homologous *RIBA2* gene, for efficient marker-assisted selection in different faba bean populations. (a) A boxplot illustrating the haplotype effect of four *RIBA2* markers in a diverse population of faba beans, where Haplotype 1 corresponds to high v-c, while Haplotype 2 corresponds to low v-c. Some high v-c genotypes corresponding to *VC1C* genotypes cluster together with low v-c genotypes as outliers in Haplotype 2. (b) A boxplot showing the haplotype effect of all five v-c markers, where Haplotype 1 and 3 represent *VC1A* and *VC1C* high v-c genotypes, respectively, while Haplotype 2 represents all low v-c genotypes.

TABLE 4 | Genotype calls and haplotypes for four *RIBA2* markers, along with corresponding phenotypes in a diverse faba bean panel.

SNP_b1	SNP_b2	SNP_pr1	SNP_pr2	Haplotype	Number of individuals	Mean vicine content	v-c status
AA	GG	AA	AA	Haplotype 1	46	0.416	high v-c
GG	CC	TT	TT	Haplotype 2	50	0.143	low v-c

for a few sequence variations, such as the absence of this 2-bp insertion in the active *VC1* gene. *VC1A* genotypes are the predominant high v-c genotypes, constituting about 90% of the high v-c genotypes in our population, including most commercial cultivars. In contrast, the outlying *VC1C* genotypes are rare, with only six genotypes representing this variant in our faba bean set.

We observed that *RIBA2* markers can effectively distinguish A-type high v-c genotypes from low v-c genotypes, but often not C-type. Hence, in populations with only the A-type allele, *RIBA2* markers are sufficient for selection and show 100% selection accuracy. Furthermore, the five v-c markers from the two neighbouring *VC1* and *RIBA2* genes segregated into three haplotypes, Haplotype 1, Haplotype 2 and Haplotype 3 (Table 5). Haplotypes 1 and 3 correlate with high v-c, while Haplotype 2 correlates with low v-c, where Haplotype 3 corresponds to *VC1C* genotypes (Figure 3b). Refer to Table S2 for a comprehensive list of the validation set with corresponding marker alleles and v-c content.

4 | Discussion

The presence of significant seed v-c levels in faba bean affects its global acceptance. Low v-c content is an essential trait required to be incorporated into modern cultivars to drive wider adoption of faba bean globally (Zanotto et al. 2020). However, progress in the development of low v-c cultivars has been slow due to a lack of reliable genomic tools for MAS. Previously developed molecular markers (Gutierrez et al. 2006; Khazaei et al. 2017; Tacke et al. 2022) do not allow for high prediction accuracy; hence, current breeding efforts rely on costly and time-consuming phenotyping. In this study, we developed and validated a set of five robust v-c markers within *VC1* and its diverging homologue, *RIBA2,* on chromosome 1, which are segregating with v-c content (Björnsdotter et al. 2021; Ugwuanyi et al. 2024). These markers showed high prediction accuracy, successfully distinguishing low v-c from high v-c faba beans.

The SNP08 marker was based on the 2-bp insertion in this gene. However, development of a marker based on this mutation proved challenging due to the presence of multiple *VC1* copies. We previously reported that faba bean lines carrying this mutation in the expressed gene copy always showed low v-c phenotype, but the presence of unexpressed wild type copies in the genome biased marker application (Ugwuanyi et al. 2024). We developed a reliable TPA-KASP method that enhances specificity and enables the easy conversion of SNPs in repetitive genomic regions. Since the inability to convert useful variants within complex or repetitive regions into KASP assays can be attributed to the small amplicon size limitation of KASP genotyping, which cannot take advantage of unique variations further away from the polymorphic site, we leveraged long amplicons to distinguish homologous sequences by exploiting unique variations localized further away from the SNP site by incorporating a preliminary PCR step, resulting in a cost-efficient and reliable selection marker. The ability of this marker to accurately predict v-c phenotypes is expected since the mutation causes a reading frameshift, which inactivates the gene function, resulting in a significant reduction of seed v-c content in mutants (Björnsdotter et al. 2021). Therefore, this two-step approach improves genotyping accuracy within repetitive regions, making it suitable for marker development in breeding programmes, especially in faba bean with its complex genome, as it is reliable, reproducible, cost-effective and ideal for high-throughput applications.

Furthermore, the markers based on the *RIBA2* gene showed high prediction accuracy and, being a single-copy gene, are straightforward to use. However, *RIBA2* markers often do not achieve 100% prediction accuracy in a mixed population. While the SNP08 marker is efficient and can be used exclusively for selection in any population, *RIBA2* markers remain valuable due to their ease of use, high selection accuracy and ability to differentiate most high v-c from low v-c genotypes. In a typical breeding programme with hundreds to thousands of samples, *RIBA2* markers are advantageous because they are convenient and straightforward, making them suitable for preselection. This preselection can eliminate over 90% of high v-c genotypes, depending on the breeding population, with the remaining pool subsequently screened with the SNP08 marker to ensure accurate selection. This strategy makes the selection process both convenient and specific.

Our study demonstrated the usefulness and precision of the quick and easy-to-use markers developed in this study for predicting v-c phenotypes, eliminating the need for expensive, time-consuming and often error-prone phenotyping of seed v-c content in breeding programmes. These KASP markers are reliable, cost-effective and amenable to high-throughput screening. Their application will facilitate rapid and efficient marker-assisted development of low v-c faba beans, contributing to the broader adoption and utilization of this important global protein crop. These genomic tools are valuable for faba bean breeding and will facilitate the development of high-quality faba bean protein.

TABLE 5 | Marker alleles and haplotype grouping associated with v-c phenotypes resulting from five robust v-c markers.

RIBA2 gene				*VC1* gene				
SNP_b1	SNP_b2	SNP_pr1	SNP_pr2	SNP08	Haplotype	Number of individuals	Mean vicine content	v-c status
AA	GG	AA	AA	GG	Haplotype 1	46	0.416	high v-c
GG	CC	TT	TT	AA	Haplotype 2	44	0.115	low v-c
GG	CC	TT	TT	GG	Haplotype 3	6	0.417	high v-c

Author Contributions

Samson Ugwuanyi: investigation, writing – original draft, methodology, visualization, writing – review and editing, formal analysis, software. **Hanna Tietgen:** writing – review and editing, methodology. **Manar Makhoul:** methodology, writing – review and editing. **Christian Obermeier:** methodology, writing – review and editing. **Rod J. Snowdon:** conceptualization, funding acquisition, writing – review and editing, supervision.

Acknowledgements

The authors gratefully acknowledge Stavros Tzigos and Annette Plank for their valuable technical assistance with nucleic acid extraction and plant management in the greenhouse, respectively. We are thankful to Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ, Hohenlieth, Germany) for providing the faba bean seeds and phenotypic data used in the study. Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors declare that all data reported in the paper are contained within the paper.

References

Badjona, A., R. Bradshaw, C. Millman, M. Howarth, and B. Dubey. 2023. "Faba Bean Flavor Effects From Processing to Consumer Acceptability." *Food* 12, no. 11: 2237. https://doi.org/10.3390/foods12112237.

Björnsdotter, E., M. Nadzieja, W. Chang, et al. 2021. "VC1 Catalyses a Key Step in the Biosynthesis of Vicine in Faba Bean." *Nature Plants* 7, no. 7: 923–931. https://doi.org/10.1038/s41477-021-00950-w.

Cappellini, M. D., and G. Fiorelli. 2008. "Glucose-6-Phosphate Dehydrogenase Deficiency." *Lancet* 371: 64–74. https://doi.org/10.1016/S0140-6736(08)60073–2.

Choudhary, D. K., and A. Mishra. 2019. "In Vitro and in Silico Interaction of Faba Bean (*Vicia faba* L.) Seed Extract With Xanthine Oxidase and Evaluation of Antioxidant Activity as a Therapeutic Potential." *Natural Product Research* 33: 2689–2693.

Doyle, J. J., and J. L. Doyle. 1990. "Isolation of Plant DNA From Fresh Tissue." *Focus* 12: 13–15.

Duc, G., G. Sixdenier, M. Lila, and V. Furstoss. 1989. "Search of Genetic Variability for Vicine and Convicine Content in *V. faba* L.: A First Report of a Gene Which Codes for Nearly Zero-Vicine and Zero-Convicine Contents." In 1. International Workshop on 'Antinutritional Factors (ANF) in Legume Seeds'. Wageningen. 23–25 Nov 1988: Pudoc.

Gutierrez, N., C. M. Avila, G. Duc, et al. 2006. "CAPs Markers to Assist Selection for Low Vicine and Convicine Contents in Faba Bean (*Vicia faba* L.)." *Theoretical and Applied Genetics* 114: 59–66.

Jayakodi, M., A. A. Golicz, J. Kreplak, et al. 2023. "The giant diploid faba genome unlocks variation in a global protein crop." *Nature* 615, no. 7953: 652–659. https://doi.org/10.1038/s41586-023-05791-5.

Khazaei, H., R. W. Purves, M. Song, et al. 2017. "Development and Validation of a Robust, Breeder-Friendly Molecular Marker for the VC—Locus in Faba Bean." *Molecular Breeding* 37: 140. https://doi.org/10.1007/s11032-017-0742-5.

Koivunen, E., P. Tuunainen, E. Valkonen, L. Rossow, and J. Valaja. 2014. "Use of Faba Beans (*Vicia faba* L.) in Diets of Laying Hens." *Agricultural and Food Science* 23: 165–172.

Makhoul, M., and C. Obermeier. 2022. "Development of Breeder-Friendly KASP Markers From Genome-Wide Association Studies Results." In *Genome-Wide Association Studies. Methods in Molecular Biology*, edited by D. Torkamaneh and F. Belzile, 287–310. New York: Springer. https://doi.org/10.1007/978-1-0716-2237-7_17.

Makhoul, M., C. Rambla, K. P. Voss-Fels, L. T. Hickey, R. J. Snowdon, and C. Obermeier. 2020. "Overcoming Polyploidy Pitfalls: A User Guide for Effective SNP Conversion Into KASP Markers in Wheat." *Theoretical and Applied Genetics* 133, no. 8: 2413–2430. https://doi.org/10.1007/s00122-020-03608-x.

Novák, P., M. S. Guignard, P. Neumann, et al. 2020. "Repeat-Sequence Turnover Shifts Fundamentally in Species With Large Genomes." *Nature Plants* 6, no. 11: 1325–1329. https://doi.org/10.1038/s41477-020-00785-x.

Raatz, B., C. Mukankusi, J. D. Lobaton, et al. 2019. "Analyses of African Common Bean (*Phaseolus vulgaris* L.) Germplasm Using a SNP Fingerprinting Platform: Diversity, Quality Control and Molecular Breeding." *Genetic Resources and Crop Evolution* 66: 707–722. https://doi.org/10.1007/s10722-019-00746-0.

Semagn, K., R. Babu, S. Hearne, and N. Olsen. 2014. "Single Nucleotide Polymorphism Genotyping Using Kompetitive Allele Specific PCR (KASP): Overview of the Technology and Its Application in Crop Improvement." *Molecular Breeding* 33: 1–14. https://doi.org/10.1007/s11032-013-9917-x.

Sixdenier, G., F. Cassecuelle, L. Guillaumin, and G. Duc. 1996. "Rapid Spectrophotometric Method for Reduction of Vicine and Convicine in Faba Bean Seed." *FABIS Newsletter* 38, no. 39: 42–44.

Tacke, R., W. Ecke, M. Höfer, O. Sass, and W. Link. 2022. "Fine-Mapping of the Major Locus for Vicine and Convicine in Faba Bean (*Vicia faba*) and Marker-Assisted Breeding of a Novel, Low Vicine and Convicine Winter Faba Bean Population." *Plant Breeding* 141: 644–657. https://doi.org/10.1111/pbr.13039.

Ugwuanyi, S., M. Makhoul, A. A. Golicz, C. Obermeier, and R. J. Snowdon. 2024. "*RIBA2* Regulates Baseline Vicine Content in Faba Bean." https://doi.org/10.1101/2024.08.06.606773.

Ugwuanyi, S., O. S. Udengwu, R. J. Snowdon, and C. Obermeier. 2022. "Novel Candidate Loci for Morpho-Agronomic and Seed Quality Traits Detected by Targeted Genotyping-By-Sequencing in Common Bean." *Frontiers in Plant Science* 13: 1014282. https://doi.org/10.3389/fpls.2022.1014282.

Vilariño, M., J. P. Métayer, K. Crépon, and G. Duc. 2009. "Effects of Varying Vicine, Convicine and Tannin Contents of Faba Bean Seeds (*Vicia faba* L.) on Nutritional Values for Broiler Chicken." *Animal Feed Science and Technology* 150: 114–121. https://doi.org/10.1016/j.anifeedsci.2008.08.001.

Wang, S., D. Wong, K. Forrest, et al. 2014. "Characterization of Polyploid Wheat Genomic Diversity Using a High-Density 90,000 Single Nucleotide Polymorphism Array." *Plant Biotechnology Journal* 12: 787–796. https://doi.org/10.1111/pbi.12183.

Zanotto, S., A. Vandenberg, and H. Khazaei. 2020. "Development and Validation of a Robust KASP Marker for zt2 Locus in faba Bean (*Vicia faba*)." *Plant Breeding* 139: 375–380. https://doi.org/10.1111/pbr.12772.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

4 Developmental stage-dependent gene expression modulates maternal control of seed vicine biosynthesis in faba bean

Ugwuanyi, S., Gerke-Fabian, T., Tietgen, H., Dreyer, F., Abbadi, A. and Snowdon, R. (2025). *Legume Science*, 7(4): e70057. https://doi.org/10.1002/leg3.70057.

Legume Science

WILEY

Legume Science Open Access

ORIGINAL ARTICLE OPEN ACCESS

Developmental Stage-Dependent Gene Expression Modulates Maternal Control of Seed Vicine Biosynthesis in Faba Bean

Samson Ugwuanyi[1,2] | Gerke-Fabian Thomas[1] | Hanna Tietgen[3] | Felix Dreyer[3] | Amine Abbadi[3] | Rod J. Snowdon[1]

[1]Department of Plant Breeding, Justus Liebig University, Giessen, Germany | [2]Department of Plant Science and Biotechnology, University of Nigeria, Nsukka, Nigeria | [3]NPZ Innovation GmbH, Holtsee, Germany

Correspondence: Samson Ugwuanyi (samson.ugwuanyi@agrar.uni-giessen.de)

Received: 2 September 2025 | **Revised:** 10 October 2025 | **Accepted:** 15 October 2025

Funding: The work was supported by DFG grant 469336000 to RJS from the German Research Society (DFG) for the International Research Training Group IRTG 2843 Accelerating Crop Genetic Gain.

Keywords: antinutritional factors | maternal inheritance | *RIBA2* | tissue-specific expression | *VC1* | *Vicia faba* | Vicine-convicine

ABSTRACT

Faba bean (*Vicia faba* L.) is a multipurpose legume valued for its high seed protein content and contribution to sustainable agriculture. However, its broader utilization is constrained by the presence of vicine and convicine (v-c), antinutritional compounds distributed throughout the plant. Although key genes involved in v-c biosynthesis have been identified, enabling genomic tools for breeding low v-c cultivars, the biosynthesis and accumulation of v-c within the plant remain incompletely understood. While previous studies reported maternal determination of seed v-c content, the underlying molecular mechanism remains unclear. In this study, we used multiple reciprocal crosses between five high and two low v-c parental lines to show that seed v-c content is determined by the maternal genotype, with F1 seeds consistently reflecting the maternal v-c phenotype. Expression analysis of *VICINE-CONVICINE 1* and *RIBOFLAVIN BIOSYNTHESIS PROTEIN 2* revealed maternal influence, with transcripts predominantly expressed in the seed coat, while the embryo remained transcriptionally inactive. Tissue-specific expression profiling confirmed that the activity of v-c-related genes is restricted to maternal tissues during seed development, suggesting that maternal determination is regulated by developmental stage-specific transcription. This absence of gene activity in the embryo explains the observed phenotypic delay in which the seed v-c content reflects the genotype of the previous generation. Despite maternal control of v-c expression, we found that segregation of the v-c trait follows Mendelian ratios regardless of cross direction, indicating that maternal phenotype does not affect breeding outcomes. These findings provide a molecular basis for maternal inheritance of v-c content and offer practical insights for breeding strategies, including seed-based genotyping, marker-assisted selection, and hybrid or synthetic breeding, supporting effective development of low v-c varieties.

1 | Introduction

Faba bean (*V. faba* L.) is a multipurpose grain legume globally recognized for its nutritional and ecological benefits. It is a rich source of dietary protein, playing a vital role in human and animal nutrition (Skovbjerg et al. 2020), and contributes to sustainable agriculture through symbiotic nitrogen fixation, diversification of cropping systems, and decreased fossil energy inputs in plant production (Jensen et al. 2010). With a yield potential of up to 8 t ha^{-1} (O'Donovan and Dunne 2021), faba bean holds significant promise for enhancing food and feed security. However, this potential remains largely unrealized, as current productivity is still well below the attainable yield, constrained by both biotic and abiotic stresses, as well as limited breeding attention to high and stable yield traits.

One of the most critical limitations to the wider use of faba bean is the presence of the antinutritional pyrimidine glucosides, vicine and convicine (v-c), which cause hemolytic anemia in human individuals with glucose-6-phosphate dehydrogenase (G6PD) deficiency. This disease is the most common human enzymopathy, affecting around 4% of the global population (Luzzatto and Arese 2018). Reduction or elimination of v-c content is therefore a major breeding target in faba bean (Augustin and Cole 2022).

The causal antinutritional compounds are synthesized through the riboflavin biosynthetic pathway by the activity of bifunctional riboflavin biosynthesis protein (RIBA) genes (Björnsdotter et al. 2021; Ugwuanyi, Makhoul, et al. 2025). The discovery of a naturally occurring low v-c accession by Duc et al. (1989) enabled the development of low v-c cultivars (Khazaei et al. 2019). Breeding efforts are increasingly supported by genomic tools that facilitate selection for the low v-c trait (Khazaei et al. 2017; Tacke et al. 2022; Ugwuanyi, Tietgen, et al. 2025).

Despite these advances, the biosynthesis and transport of v-c within the plant remain incompletely understood. Vicine and convicine are not limited to seeds; they have also been detected in various vegetative and reproductive tissues, including roots, shoots, leaves, flowers, and pods (Duc et al. 1989; Ramsay and Griffiths 1996; Goyoaga et al. 2008; Tacke 2023). Earlier studies suggest that v-c accumulation in seeds is determined maternally, with the seed coat proposed as the biosynthetic site and the embryo acting as a passive recipient (Brown and Roberts 1972; Duc et al. 1989). While this maternal determination of seed v-c content is widely reported, the underlying molecular mechanism remains unclear, particularly how the maternal genotype influences gene expression dynamics and segregation patterns in subsequent generations. Furthermore, multiple *VC1* gene copies are present in the faba bean genome, with selective expression of specific copies among genotypes (Ugwuanyi, Makhoul, et al. 2025). This raises questions about potential parent-of-origin effects on gene expression and the developmental regulation of v-c biosynthesis.

Therefore, the current study aims to elucidate the molecular mechanisms modulating the maternal control of v-c biosynthesis in faba bean. Specifically, we: i) developed multiple reciprocal populations using two low v-c recurrent parents and five high v-c donor parental lines, and performed detailed v-c phenotyping and molecular analyses across generations to investigate the molecular basis of maternal inheritance; ii) genotyped F2 populations to evaluate the segregation of alleles of v-c-related genes and to determine whether the direction of the cross or maternal phenotype is critical in crossing programs, thereby informing more efficient breeding strategies for low v-c cultivars.

2 | Materials and Methods

2.1 | Plant Materials

Inbred seed samples from seven elite spring faba bean breeding lines were obtained from the seed company Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ, Hohenlieth, Germany). These lines displayed varying levels of seed v-c content across different genetic backgrounds. Two lines (NPZ-FB-55 and NPZ-FB-92) had low vicine content, while five lines (NPZ-FB-70, NPZ-FB-83, NPZ-FB-113, NPZ-FB-116, and NPZ-FB-117) had high vicine content.

2.2 | Crossing Scheme and Greenhouse Cultivation

To generate multiple populations segregating for v-c content across diverse genetic backgrounds, reciprocal crosses were performed between all possible combinations of the two low v-c lines and the five high v-c lines (Figure 1). Seeds from each parent were sown in 5-l plastic pots (18 × 18 cm), filled with soil (Fruhstorfer Erde, Type N; HAWITA Gruppe GmbH, Vechta, Germany), and grown in a controlled, insect-proof greenhouse chamber.

Plants were maintained under standard agronomic practices, including irrigation, pest control, and application of inorganic fertilizer (WUXAL Universaldünger; Hauert HBG Dünger AG, Grossaffoltern, Switzerland). Daytime temperatures ranged from 18°C–25°C with 80%–90% relative humidity, and nighttime conditions ranged from 10°C to 12°C with 65%–75% humidity, under a 16-h photoperiod.

During flowering, reciprocal crosses were carried out to produce reciprocal F1 seeds from the ten high v-c × low v-c parental combinations (see Figure 1). Homozygosity of the parental plants and hybridity of the causal loci in all crosses was confirmed using two previously published v-c SNP markers, SNP08 and SNP_b1 (Ugwuanyi, Tietgen, et al. 2025). Confirmed F1 plants were self-pollinated to produce F2 populations.

2.3 | Nucleic Acid Extraction

Young leaf tissue for DNA isolation was collected from the first fully opened leaves of the parental, F1, and F2 plants. Samples were ground into fine powder using a TissueLyser II (Qiagen, Hilden, Germany). Genomic DNA was isolated from the leaf powder using a Biosprint96 DNA extraction robot (Qiagen, Hilden, Germany) following the manufacturer's instructions. DNA concentrations were measured using the Qubit dsDNA BR assay kit (ThermoFisher Scientific, Germany) and subsequently diluted to 20–50 ng/μL for downstream applications.

For RNA extraction used for gene expression and transcript analyses, whole seeds were collected at the early seed-filling stage (13–21 days after crossing or tripping, manipulation of petals to extrude stigma and anthers to ensure selfing) from parental, F1, and F2 plants. Because of the low multiplication rate of faba bean, only a subset of the F1 and F2 plants was sampled. Additionally, for tissue-specific analyses, young leaves and flowers were harvested directly from F1 plants of the cross NPZ-FB-92 × NPZ-FB-117, whereas seed coats, embryos, and whole seeds were obtained from seeds (F2 seeds) that developed on the same plants. All samples were flash-frozen in liquid nitrogen and stored at −80°C until further processing. Frozen tissues were ground with a mortar and pestle, and total RNA was isolated using the Zymo RNA Miniprep Kit (Zymo Research, Freiburg, Germany), following the manufacturer's protocol. In-column DNase I treatment was performed to eliminate genomic

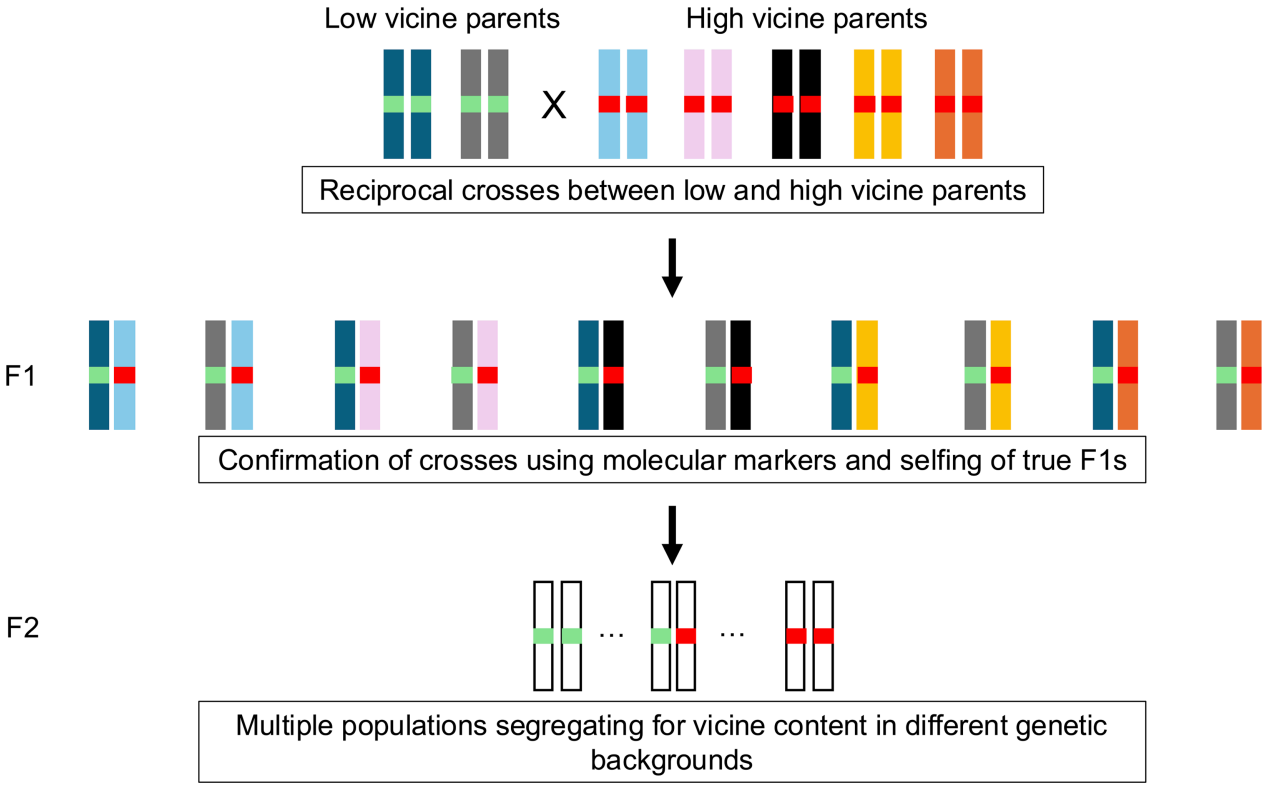


FIGURE 1 | **Population development.** Schematic representation of the crossing scheme used to develop multiple segregating populations for seed vicine content. Reciprocal crosses were made between combinations of the low v-c parental lines NPZ-FB-55 (dark blue) and NPZ-FB-92 (grey) with the high v-c parental lines NPZ-FB-70 (light blue), NPZ-FB-116 (pink), NPZ-FB-83 (black), NPZ-FB-117 (yellow), and NPZ-FB-113 (orange).

DNA contamination. RNA concentration and integrity were assessed using the Qubit RNA assay kit (ThermoFisher Scientific, Germany) and 1% agarose gel electrophoresis. First-strand cDNA synthesis was conducted using the RevertAid cDNA Synthesis Kit (ThermoFisher Scientific, Germany) as described in Ugwuanyi, Makhoul, et al. (2025).

2.4 | Quantification of Seed V-C Content

Mature seeds from the parental lines and F1/F2 populations were collected for vicine quantification. Each seed was milled separately into a homogenous powder using ball mill tubes in a TissueLyser II (Qiagen, Hilden, Germany). Extraction and quantification of v-c were performed using high-performance liquid chromatography (HPLC) according to the method in Zaar and Balko (2022). Each sample was measured in triplicate.

2.5 | KASP Genotyping

Genotyping of insertions/deletions (InDels) and single nucleotide polymorphisms (SNPs) within the coding sequences of the genes *VICINE-CONVICINE 1* (*VC1*) and *RIBOFLAVIN BIOSYNTHESIS PROTEIN 2* (*RIBA2*) was conducted using Kompetitive Allele-Specific PCR (KASP). Two robust KASP markers were used: SNP08, which detects the presence or absence of a 2-bp mutation in *VC1*, and SNP_b1, which targets a SNP within *RIBA2* (Ugwuanyi, Tietgen, et al. 2025). These two genes were selected because previous studies have linked them to variations in v-c content (Björnsdotter et al. 2021; Ugwuanyi, Makhoul, et al. 2025). Genotyping was performed using genomic DNA (gDNA) to determine the genotypes, and complementary DNA (cDNA) to assess expressed gene variants. For cDNA-based genotyping, the SNP_b1 marker was modified to target coding regions using the following primers:

Hex tail: gaaggtcggagtcaacggattCAAGAACATCTTCTCCATCTCCAATG

Fam tail: gaaggtgaccaagttcatgctCAAGAACATCTTCTCCATCTCCAATA

Common primer: GGGCCATTCACATCATACTGTTATAAA.

Detailed information on the KASP assay protocol is provided in Ugwuanyi, Tietgen, et al. (2025).

2.6 | Gene Expression Analysis

Relative transcript levels of *VC1* and *RIBA2* were determined using the previously synthesized cDNA. Each sample was analyzed in three biological and three technical replicates. For each gene target, three no-template controls (MilliQ water) were included. The *ELF1A* gene (Gutierrez et al. 2011) was used as an internal reference for normalizing cycle threshold (Ct) values. For *VC1* expression analysis, the following forward and reverse primers were used: TGGAGACATATTTGGATCTGCCAGATG and TAGCCTCAATCTGCTGCATTGA.

To assess cumulative expression of *VC1* and *RIBA2*, the following reverse primer was used with the same forward primer above: ATCCAATACCCCTACCTTCATGT. Each reaction consisted of 5 μL of 2×SYBR Green master mix (ThermoFisher Scientific, Germany), 1 μL each of 10-μM forward and reverse primers, 1 μL of cDNA, and 2 μL of MilliQ water. Quantitative PCR was performed on a ViiA7 real-time PCR system (Life Technologies, USA) under the following cycling conditions: initial denaturation at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Relative gene expression was calculated using the delta–delta Ct (ΔΔCt) method (Livak and Schmittgen 2001).

2.7 | Data Analysis

Experimental data were analyzed using R software version 4.1 (R Core Team 2021). Statistical significance was determined at $p < 0.05$.

3 | Results

3.1 | Maternal Genotype Determines Seed V-C Content in Faba Bean

The hybridity status of the F1 seeds from the 10 different reciprocal cross combinations between two low v-c and five high v-c parental lines was confirmed using SNP08 and SNP_b1, molecular markers linked to the v-c–related genes *VC1* and *RIBA2*, respectively (Ugwuanyi, Tietgen, et al. 2025). Phenotypic analysis of mature seeds confirmed that the two low v-c parental lines contained 0.232 and 0.183 mg/g seed dry weight of v-c, respectively, while v-c content in high v-c parents ranged from 2.509 to 3.957 mg/g. In all crosses, F1 seed v-c content corresponded to the maternal parent v-c phenotype. For example, F1 seeds derived from low v-c maternal plants averaged 0.323 mg/g, whereas those from high v-c maternal parents averaged 3.838 mg/g, demonstrating maternal control of seed v-c phenotype. F2 seeds derived from selfed F1 plants showed intermediate v-c levels (~1.803 mg/g), irrespective of cross direction, suggesting intermediate gene action (Figure 2a; Table 1).

3.2 | *VC1* and *RIBA2* Genes Are Transcriptionally Inactive in the Embryo During Seed Development

Expression analysis of the major v-c biosynthetic gene *VC1* (Björnsdotter et al. 2021) and the minor-effect candidate gene *RIBA2* (Ugwuanyi, Makhoul, et al. 2025) during early seed filling revealed 4-fold higher levels of *VC1* expression in high v-c parental lines compared to low v-c parents. We found that gene expression was clearly influenced by the maternal genotype. In F1 seeds, expression levels reflected those of the maternal parent from which the seed was harvested. For example, F1 seeds from high v-c maternal plants showed high *VC1* expression, while F1 seeds from low v-c maternal plants showed low expression. F2 seeds (harvested from F1 plants) displayed intermediate expression levels (Figure 2b). This pattern was also observed for the combined expression of *VC1* and *RIBA2*, and was consistent across all subpopulations (Figure 3). These results suggest that the genes are either inactive in the embryo during the seed-filling stage or that maternal imprinting is involved, such that only the maternal alleles are expressed while the paternal alleles are silenced.

To determine which alleles contributed to the observed gene expression, we first sequenced cDNA from representative parental lines (Table S1) and then assessed allele-specific transcripts in all samples using cDNA-based KASP genotyping. Our results revealed a monoallelic expression pattern in F1 seeds, with the expressed allele matching that of the maternal parent. In contrast,

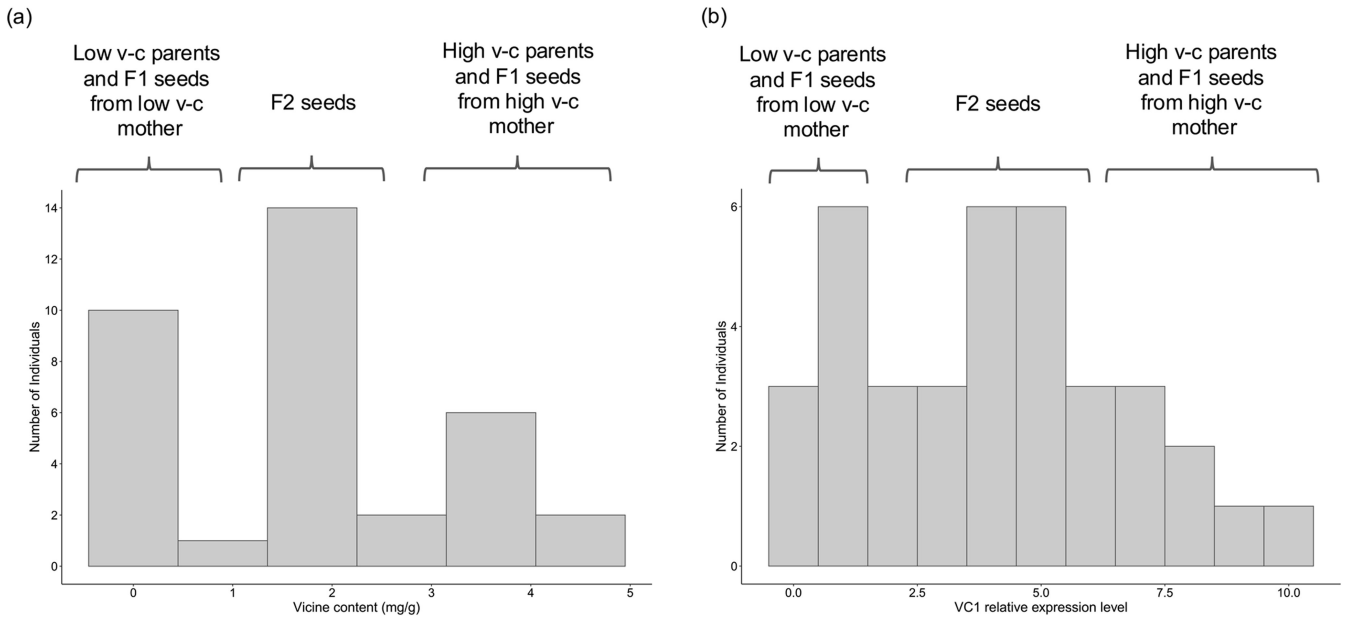


FIGURE 2 | **Inheritance pattern of vicine content and *VC1* expression.** Histograms showing the distribution of (a) vicine content and (b) *VC1* expression levels across parental lines and F1/F2 seeds.

TABLE 1 | Genotypic profiles of parental, F1, and F2 populations at SNP08 and SNP_b1 loci and their corresponding vicine contents. Genotyping was performed on both seed (complementary DNA) and leaf (genomic DNA) samples. Mend indicates that the population is segregating according to the Mendelian ratio (see Supplementary Table S2 for detailed information). P represents the parental generation while F1 and F2 represent the first and second filial generations. SNP08 and SNP_b1 are gene-specific molecular markers for *VC1* and *RIBA2*, respectively.

			Seed (complementary DNA)		Plant (genomic DNA)	
Genotype	**Generation**	**Vicine content (mg/g)**	**SNP08**	**SNP_b1**	**SNP08**	**SNP_b1**
FB-0092	P	0.183	AA	GG	AA	GG
FB-0055	P	0.232	AA	GG	AA	GG
FB-0083	P	3.761	GG	GG	GG	GG
FB-0116	P	2.509	GG	AA	GG	AA
FB-0117	P	3.957	GG	AA	GG	AA
FB-0070	P	2.792	GG	AA	GG	AA
FB-0113	P	3.313	GG	AA	GG	AA
FB-0055 × FB-0117	F1	0.433	AA	GG	GA	GA
FB-0117 × FB-0055	F1	4.085	GG	AA	GA	GA
FB-0055 × FB-0113	F1	0.463	AA	GG	GA	GA
FB-0113 × FB-0055	F1	3.315	GG	AA	GA	GA
FB-0092 × FB-0113	F1	0.268	AA	GG	GA	GA
FB-0113 × FB-0092	F1	3.873	GG	AA	GA	GA
FB-0055 × FB-0083	F1	0.270	AA	GG	GA	GG
FB-0083 × FB-0055	F1	3.415	GG	GG	GA	GG
FB-0092 × FB-0083	F1	0.279	AA	GG	GA	GG
FB-0083 × FB-0092	F1	4.501	GG	GG	GA	GG
FB-0092 × FB-0116	F1	0.291	AA	GG	GA	GA
FB-0116 × FB-0092	F1	3.122	GG	AA	GA	GA
FB-0116 × FB-0055	F1	3.944	GG	AA	GA	GA
FB-0092 × FB-0070	F1	0.227	AA	GG	GA	GA
FB-0055 × FB-0070	F2	1.926	GA	GA	Mend	Mend
FB-0070 × FB-0055	F2	1.657	GA	GA	Mend	Mend

(Continues)

TABLE 1 | (Continued)

Genotype	Generation	Vicine content (mg/g)	Seed (complementary DNA)		Plant (genomic DNA)	
			SNP08	SNP_b1	SNP08	SNP_b1
FB-0055 × FB-0083	F2	1.801	GA	GG	Mend	Mend
FB-0083 × FB-0055	F2	1.707	GA	GG	Mend	Mend
FB-0092 × FB-0117	F2	1.743	GA	GA	Mend	Mend
FB-0117 × FB-0092	F2	1.634	GA	GA	Mend	Mend
FB-0092 × FB-0083	F2	1.734	GA	GG	Mend	Mend
FB-0083 × FB-0092	F2	1.780	GA	GG	Mend	Mend
FB-0092 × FB-0113	F2	1.877	GA	GA	Mend	Mend
FB-0113 × FB-0092	F2	1.657	GA	GA	Mend	Mend
FB-0055 × FB-0113	F2	1.948	GA	GA	Mend	Mend
FB-0113 × FB-0055	F2	1.756	GA	GA	Mend	Mend
FB-0116 × FB-0092	F2	1.952	GA	GA	Mend	Mend
FB-0117 × FB-0055	F2	1.853	GA	GA	Mend	Mend
FB-0116 × FB-0055	F2	1.775	GA	GA	Mend	Mend
FB-0070 × FB-0092	F2	1.937	GA	GA	Mend	Mend
FB-0092 × FB-0070	F2	1.586	GA	GA	Mend	Mend

F2 seeds displayed biallelic expression (Table 1), consistent with their intermediate v-c phenotypes and indicative of intermediate gene action. Importantly, the transcripts corresponded to the genotype of the maternal plant or seed coat, which is entirely of maternal origin, rather than the embryo. This was confirmed by genotyping leaf DNA from parental and F1/F2 plants. Despite heterozygous or segregating genotypes at these generations, gene expression in seeds remained monoallelic in F1s and biallelic in F2s (Figure 4). These findings indicate that only maternal transcripts are expressed in seeds during development, suggesting that the embryo is transcriptionally inactive for v-c genes.

To confirm this, we performed qRT-PCR on RNA samples from leaf, flower, seed coat, embryo, and whole seeds from reciprocal F1 plants derived from the cross NPZ-FB-117 × NPZ-FB-92. These tissues were ideal for analysis because the maternal tissues were in the F1 generation, while the developing seeds were in the F2 generation. This allowed us to clearly distinguish transcripts originating from maternal versus embryonic tissues and to rule out the possibility of imprinting. We found that cumulative expression of *VC1* and *RIBA2* was highest in the seed coat and whole seed, with minimal expression in leaves, flowers, and embryos (Figure 5). We then compared cDNA and gDNA from the same tissues, revealing that the cDNA matched the gDNA in all tissues except the embryo, where cDNA matched the gDNA of the maternal plant. This result suggests that the transcripts detected in embryos likely originated from surrounding maternal tissues. Taken together, these results support the conclusion that v-c gene expression is primarily confined to maternal tissues, with the

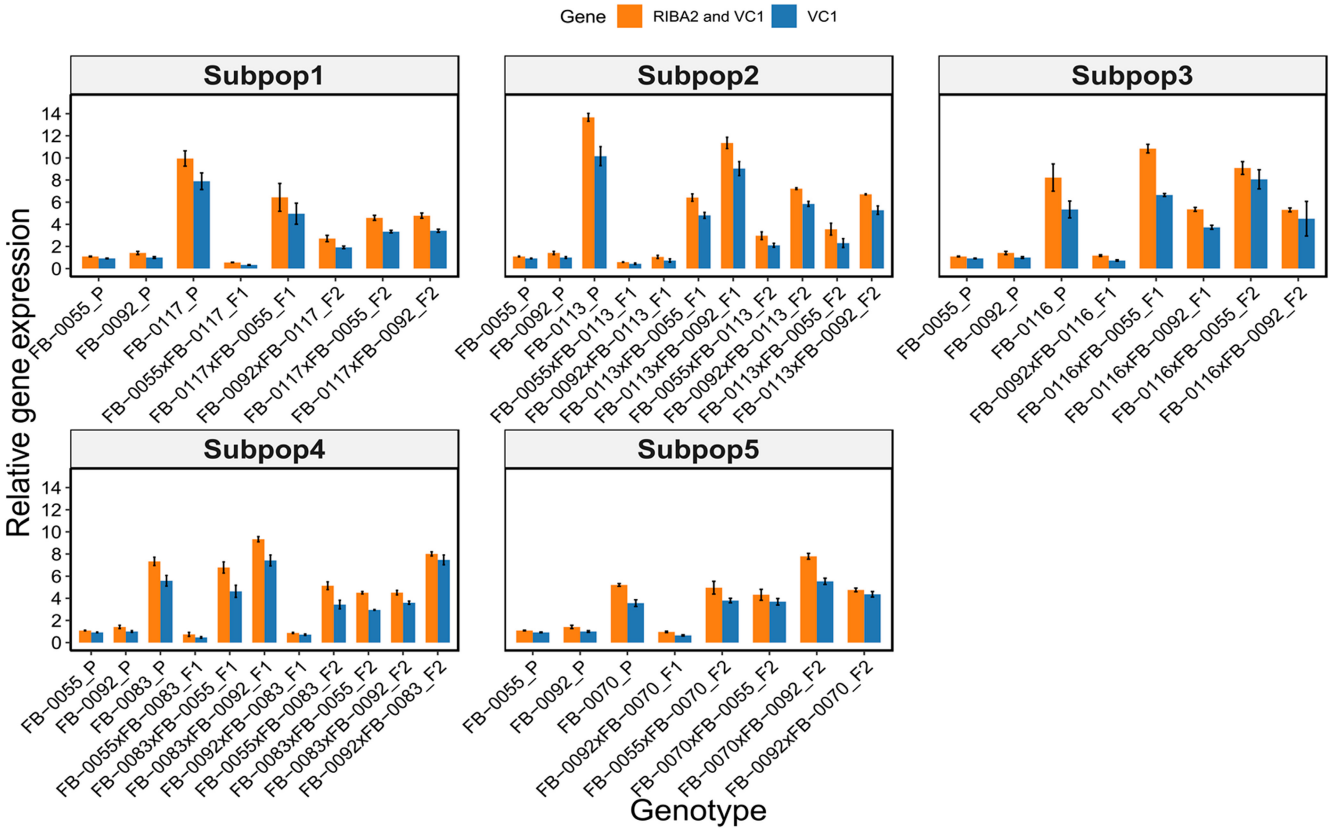


FIGURE 3 | **Relative transcript abundance of vicine-related genes.** Bar plot showing relative expression levels of *VC1* and *RIBA2* in immature seeds harvested from parental lines and F1/F2s.

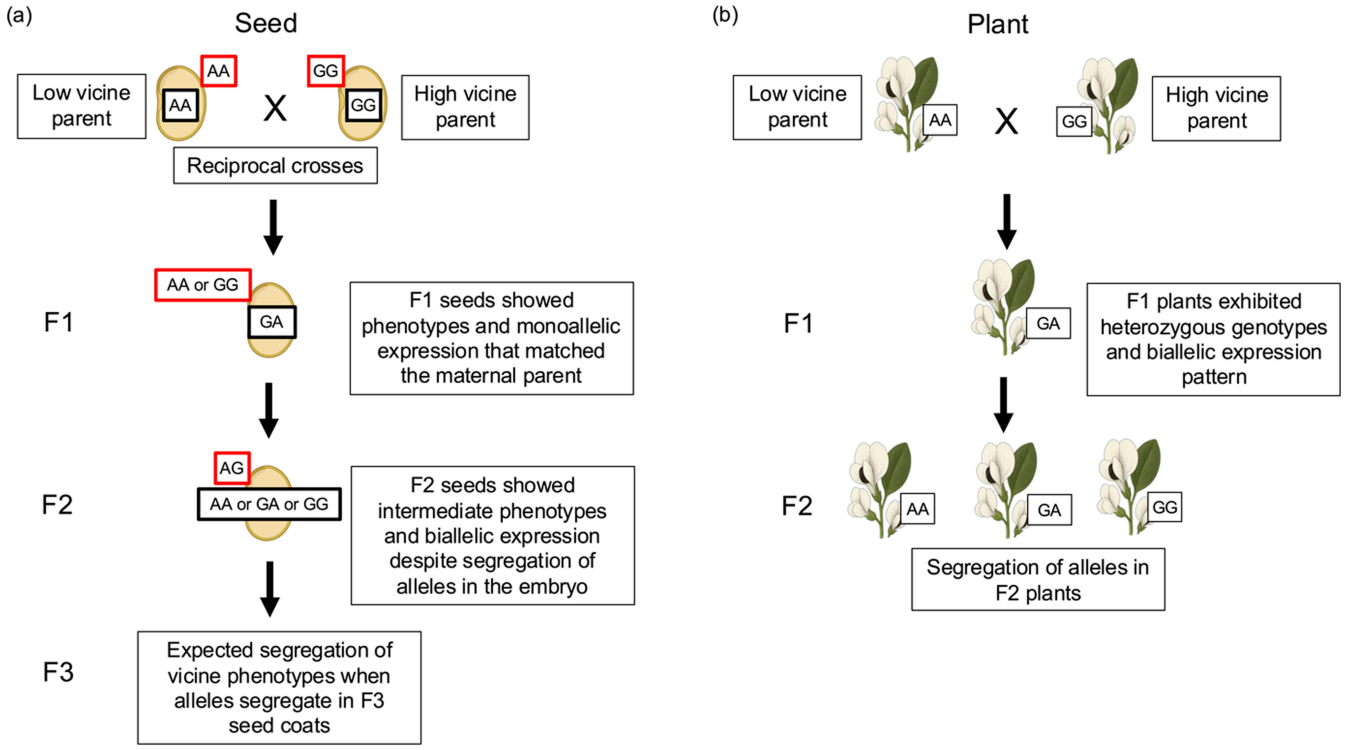
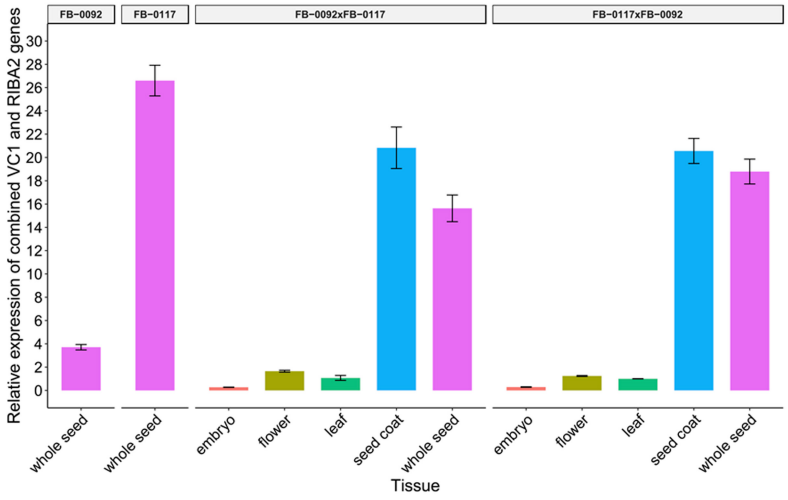


FIGURE 4 | **Schematic representation of the gene action and expression on vicine content.** (a) Seed-level and (b) plant-level segregation of vicine content and *VC1* genotypes across parental, and F1/F2s. Genotypes AA and GG represent the different SNP alleles associated with the *vc1* and *VC1* variants, respectively. Red and dark boxes represent the seed coat and embryo genotypes, respectively. *VC1* genotype combinations and vicine levels are traced from low and high vicine parental lines through hybrid and segregating generations.

(a)



(b)

		Complementary DNA		Genomic DNA	
Genotype	Tissue	SNP08	SNP_b1	SNP08	SNP_b1
NPZ-FB-92	whole seed	AA	GG	AA	GG
NPZ-FB-117	whole seed	GG	AA	GG	AA
NPZ-FB-92xNPZ-FB-117	embryo	GA	GA	GG	AA
	flower	GA	GA	GA	GA
	leaf	GA	GA	GA	GA
	seed coat	GA	GA	GA	GA
	whole seed	GA	GA	GA	GA
NPZ-FB-117xNPZ-FB-92	embryo	GA	GA	GG	AA
	flower	GA	GA	GA	GA
	leaf	GA	GA	GA	GA
	seed coat	GA	GA	GA	GA
	whole seed	GA	GA	GA	GA

FIGURE 5 | **Tissue-specific analyses of VC1 and RIBA2 transcripts.** (a) Relative expression of combined *VC1* and *RIBA2* genes across different faba bean tissues and (b) comparison of cDNA and gDNA detected in each tissue type. Leaves and flowers were harvested directly from F1 plants of the cross NPZ-FB-92 × NPZ-FB-117, whereas seed coats, embryos and whole seeds were obtained from seeds (F2 seeds) that developed on the same plants.

embryo remaining transcriptionally inactive during seed development. This reinforces the hypothesis that the seed coat is the primary site of v-c biosynthesis and the main determinant of seed v-c content.

3.3 | Segregation of V-C Genes Is Independent of Cross Direction

Given the maternal effect on v-c biosynthesis, it is important to determine the influence of the maternal phenotype on segregation of v-c alleles in breeding programs in order to determine the necessity or effectiveness of using low v-c lines as the female parent in crosses. To investigate maternal effects, we analyzed segregation patterns in F2 populations derived from both low v-c and high v-c maternal plants. Using SNP08 and SNP_b1, we genotyped 309 F2 individuals from all reciprocal populations (Table S2). The target genes *VC1* and *RIBA2* co-segregated into three distinct haplotypes, following normal Mendelian inheritance, with no recombination observed between the two loci.

At the single-gene level, *VC1* showed a typical 1:2:1 segregation ratio across the entire population (72 AA, 160 GA, 77 GG; $\chi^2 = 0.5534$, $p = 0.7583$). For *RIBA2*, a slight deviation from the expected ratio was observed (66 AA, 144 GA, 99 GG; $\chi^2 = 8.4757$, $p = 0.0144$), likely due to the lack of marker polymorphism in crosses involving the high-vc parental line NPZ-FB-083.

Further subpopulation analysis, grouped by maternal genotype (low v-c vs. high v-c), revealed consistent Mendelian segregation patterns in both cases (Figure 6; Table 2). These results indicate that the direction of the cross, and by extension, the maternal genotype and phenotype, do not affect inheritance of the v-c trait in the F2 generation.

4 | Discussion

4.1 | Maternal Control of V-C Content Is Mediated by Tissue-Specific, Stage-Dependent Gene Expression

Vicine and convicine are fungicidal glycoalkaloids that form part of the chemical defense against biotic stress, such as fungal infection (Sergeant et al. 2024). Reports have shown that v-c inhibits the in vitro growth of fungal pathogens (Bjerg et al. 1984; Pavlík et al. 2002). However, Tacke (2023) found no significant differences between the responses of low and high v-c faba beans to soil-borne pathogens, implying that the reduction or elimination of v-c from faba beans may not have any negative effects on the crop. This has paved the way for breeding programs aimed at developing faba bean cultivars with reduced v-c content.

Vicine and convicine are known to be synthesized during seed filling, predominantly in the seed coat (Brown and Roberts 1972; Pitz et al. 1981; Ramsay and Griffiths 1996), and are subsequently transported to the cotyledons (Pitz et al. 1981; Ray and Georges 2010). However, their presence in other tissues including roots, stems, flowers, and leaves suggests additional sites of biosynthesis (Griffiths and Ramsay 1996; Tacke 2023). Although maternal determination of seed v-c content has been recognized for decades (Duc et al. 1989), the underlying molecular mechanisms remain unknown.

Our findings demonstrate that v-c biosynthesis in faba bean is controlled by tissue-specific and developmentally regulated

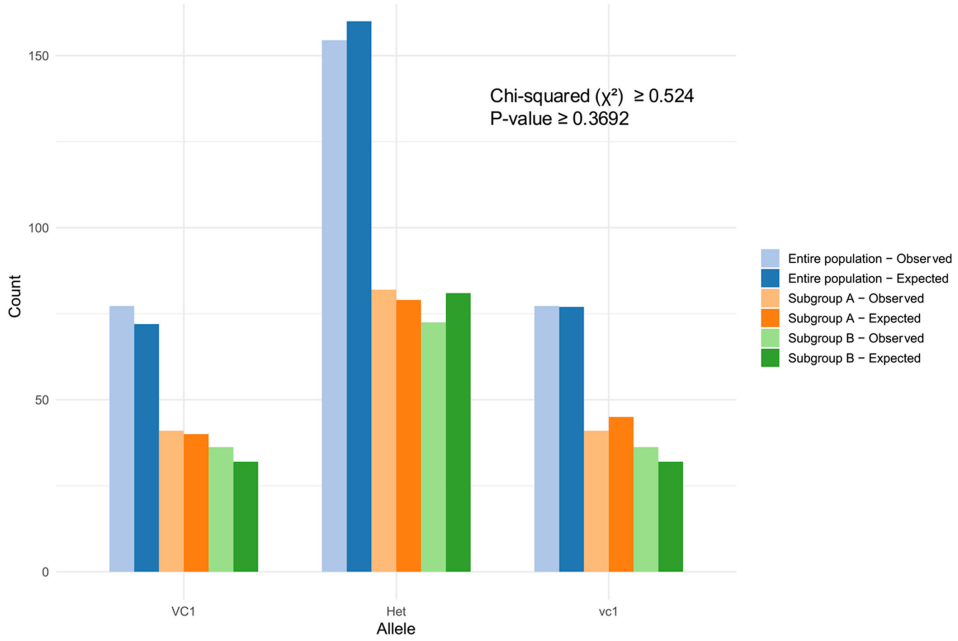


FIGURE 6 | **Segregation analysis of VC1 alleles.** Bar plot comparing observed *VC1* gene segregation pattern with expected Mendelian ratios. Subgroup A comprised F2 progenies originating from low v-c maternal lines while subpopulation B comprised progenies from high v-c maternal lines.

TABLE 2 | Segregation ratio of *VC1* and *RIBA2* genes in the F2 population.

Gene	Population	No. alleles				Expected value	Observed value	χ^2	*p*
		GG	GA	AA	Total				
VC1	Entire population	72	160	77	309	77.25:154.50:77.25	72:160:77	0.5534	0.7583
	Subpopulation A	40	79	45	164	41:82:41	40:79:45	0.52439	0.7694
	Subpopulation B	32	81	32	145	36.25:72.5:36.25	32:81:32	1.9931	0.3692
RIBA2	Entire population	99	144	66	309	77.25:154.50:77.25	99:144:66	8.4757	0.01444
	Subpopulation A	52	74	38	164	41:82:41	52:74:38	3.9512	0.1387
	Subpopulation B	47	70	28	145	36.25:72.5:36.25	47:70:28	5.1517	0.07609

gene expression. Using multiple reciprocal cross populations, we demonstrated that seed v-c content is determined by the maternal genotype, consistent with earlier observations (Brown and Roberts 1972; Duc et al. 1989; Tacke 2023). Seed development involves coordinated interaction between embryo and surrounding maternal tissues (Ohto et al. 2008), and maternal effects on traits can result from various mechanisms, including plastid inheritance, nutrient provisioning, or localized gene expression (Donohue 2009).

In the case of maternal control of seed v-c, localized gene expression in the seed coat is a likely mechanism. Previous studies have shown that v-c biosynthesis is developmentally regulated and occurs primarily during seed filling within the seed coat (Brown and Roberts 1972; Pitz et al. 1981; Ramsay and Griffiths 1996). Our findings provide molecular evidence supporting this model, demonstrating that v-c gene expression is confined to the seed coat during early seed development, while the embryo remains transcriptionally inactive. Additionally, we found that v-c genes are also expressed in other maternal tissues, such as leaves and flowers, where vicine and convicine are also detected at low levels (Tacke 2023), suggesting that biosynthesis may occur in multiple maternal tissues besides the seed coat.

Importantly, the absence of gene activity in the embryo, despite its high v-c accumulation, supports the hypothesis that v-c is synthesized in the maternal tissues and transported into the embryo for storage (Pitz et al. 1981; Ray and Georges, 2010; Björnsdotter et al. 2021). This pattern suggests a transcriptional regulation mechanism in which v-c gene expression is confined to maternal tissues by developmentally programmed gene silencing in the embryo, resulting in the complete maternal determination of v-c accumulation in developing seeds. Such maternal regulation is consistent with mechanisms observed in other crops such as maize (*Zea mays*) and in the model dicot *Arabidopsis thaliana*, where early development is driven by maternal transcripts while zygotic genome activation is delayed (Grini et al. 2002; Autran et al. 2011). Similar mechanisms of maternal tissue-specific gene expression have been described for other traits. For instance, seed oil content in Brassica crops is determined

by maternal photosynthetic tissues (Hua et al. 2012), while seed coat pigmentation in *A. thaliana* and maize is maternally regulated via localized gene expression (Debeaujon et al. 2000; Ron-Parra et al. 2016). Furthermore, maternal small RNAs and epigenetic regulators have been shown to delay zygotic genome activation during embryogenesis in *A. thaliana* and maize (Grini et al. 2002; Grimanelli et al. 2005; Autran et al. 2011). In faba bean, a similar transcriptional delay may underlie the observed inactivity of v-c biosynthesis genes in embryos, providing a plausible mechanistic explanation for the maternal control of seed v-c content.

4.2 | Considerations in V-C Breeding Programs

The maternal control of seed v-c biosynthesis aligns with the phenomenon of phenotypic delay, in which the seed phenotype reflects the maternal genotype rather than the zygotic genotype (Sun et al. 2020). This "generational lag" arises because the seed coat, which is responsible for seed v-c biosynthesis, is derived entirely from maternal tissue (Bewley et al. 2013), where the expression of v-c-related genes is confined. As a result, seed v-c content is determined by the genotype of the maternal parent, one generation prior to the generation of the embryo. This has important implications for various breeding strategies, particularly in marker-assisted selection (MAS), seed-based genotyping, and hybrid or synthetic cultivar development.

In MAS, the timing of phenotypic selection and a clear prediction of the desirable phenotype from the genotype of an appropriate generation is critical. For many seed traits, phenotyping is typically performed on F2 seeds when segregation occurs. However, this approach is unsuitable for v-c content. In early generations such as F2, phenotypic selection fails due to the absence of segregation in seeds caused by the masking effect of maternal gene expression and the resulting phenotypic lag. Therefore, selection for v-c content needs to be based on seeds produced by F2 plants (i.e., the F3 generation of seeds), as F3 seed phenotypes reflect the genotype of the F2 plant. This strategy ensures accurate selection and avoids misinterpretation of v-c content due to the maternal influence.

This maternal effect must also be considered in seed-based genotyping methods such as seed-chipping, a resource-efficient technique that extracts DNA directly from seeds, enabling rapid identification and planting of desirable genotypes from segregating populations (Gao et al. 2008; Zebosi et al. 2025). Although faba bean seeds do not yet segregate for v-c content at the F2 generation, the embryo genome is already segregating, allowing genotypes to be used for predicting the v-c phenotypes of the F3 seeds. However, the seed coat remains heterozygous at this stage because it is entirely derived from the F1 maternal plant. As a result, if the sampled tissue includes the seed coat, the extracted DNA will also be heterozygous and unsuitable for accurate genotypic prediction of v-c content. Therefore, to ensure reliable genotyping and accurate prediction of next generation seed phenotypes via seed-chipping, DNA must be isolated exclusively from the embryo or endosperm of F2 seeds, avoiding contamination from the seed coat. Alternatively, breeders need to germinate seeds and perform MAS using leaf or root tissue from juvenile plants.

In breeding programs aiming to incorporate the low v-c trait into synthetic cultivars composing populations from mixtures of different parental genotypes, strategies must ensure that progenies express the trait in the F2 and subsequent generations. Unlike many traits where hybrid performance is observable in the F1 generation, the v-c phenotype of the F1 seed reflects only the maternal genotype, masking any contribution from the paternal parent. As a result, only F2 seeds produced by F1 plants reflect the combined genetic contributions of both parents. If one parental component of a synthetic cultivar carries high v-c alleles, seeds of the population will exhibit intermediate v-c levels even if the F1 displays low v-c content, potentially failing to meet the criteria for low v-c varieties. Therefore, breeding programs need to prioritize the development of inbred lines that are homozygous for low v-c alleles to ensure consistent expression of the trait in synthetic combinations.

Finally, although v-c is maternally determined, our findings demonstrate that the maternal phenotype does not influence trait segregation in subsequent generations. In crosses between high and low v-c lines, the v-c trait segregated in a Mendelian ratio in F2 plants, regardless of maternal phenotype. Therefore, harvesting from high or low v-c maternal plants does not affect breeding outcomes.

5 | Conclusions

This study uncovers the molecular mechanism underlying the maternal control of seed v-c content in faba bean. We demonstrate that v-c gene expressions occur primarily in maternal tissues while embryos remain transcriptionally inactive during seed development. Consequently, the seed v-c phenotype is determined by the maternal genotype, whereas the zygotic genotype influences seed phenotype only in the next generation. This finding is critical for breeding programs, as it informs the adoption of appropriate strategies for the effective development of low v-c varieties.

Author Contributions

RJS and SU conceived the study; SU conducted the experiments, analyzed the data, interpreted the results, and drafted the manuscript; SU and GT performed HPLC analysis of vicine content; HT, FD, and AA provided the faba bean genotypes and preliminary vicine phenotypic data. All authors revised and approved the final manuscript.

Acknowledgments

The authors gratefully acknowledge Erika Schick and Liane Renno for their valuable technical assistance with the HPLC analysis. We are thankful to Ashwini Mudke and Lennart Scheer for their assistance with performing crosses. We appreciate the assistance of Stavros Tzigos and Annette Plank with plant management in the greenhouse. Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that supports the findings of this study are available in the supplementary material of this article.

References

Augustin, M. A., and M. B. Cole. 2022. "Towards a Sustainable Food System by Design Using Faba Bean Protein as an Example." *Trends in Food Science & Technology* 125: 1–11. https://doi.org/10.1016/j.tifs.2022.04.029.

Autran, D., C. Baroux, M. T. Raissig, et al. 2011. "Maternal Epigenetic Pathways Control Parental Contributions to Arabidopsis Early Embryogenesis." *Cell* 145, no. 5: 707–719. https://doi.org/10.1016/j.cell.2011.04.014.

Bewley, J. D., K. J. Bradford, H. W. M. Hilhorst, and H. Nonogaki. 2013. *Seeds: Physiology of Development, Germination and Dormancy*. Springer.

Bjerg, B., M. Heide, J. C. N. Knudsen, and H. Soerensen. 1984. "Inhibitory Effects of Convicine, Vicine and Dopa From *Vicia faba* on the *In Vitro* Growth Rates of Fungal Pathogens." *Zeitschrift für Pflanzenkrankheiten, Pflanzenpathologie und Pflanzenschutz* 91: 483–487.

Björnsdotter, E., M. Nadzieja, W. Chang, et al. 2021. "*VC1* Catalyses a key Step in the Biosynthesis of Vicine in Faba Bean." *Nature Plants* 7: 923–931. https://doi.org/10.1038/s41477-021-00950-w.

Brown, J. N., and E. H. Roberts. 1972. "The Inheritance of Vicine and Convicine Content in Faba Bean." *Heredity* 28, no. 2: 203–209. https://doi.org/10.1038/hdy.1972.62.

Debeaujon, I., K. M. Léon-Kloosterziel, and M. Koornneef. 2000. "Influence of the Testa on Seed Dormancy, Germination, and Longevity in Arabidopsis." *Plant Physiology* 122, no. 2: 403–414. https://doi.org/10.1104/pp.122.2.403.

Donohue, K. 2009. "Completing the Cycle: Maternal Effects as the Missing Link in Plant Life Histories." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 364: 1059–1074.

Duc, G., G. Sixdenier, M. Lila, and V. Furstoss. 1989. "Search of Genetic Variability for Vicine and Convicine Content in *Vicia faba* L. a First Report of a Gene Which Codes for Nearly Zero-Vicine and Zero-Convicine Contents." In *Recent Advances of Research in Antinutritional Factors in Legumes Seeds*, edited by J. Huisman, T. F. B. van der Poel, and I. E. Liener, 305–313. Pudoc.

Gao, S., C. Martinez, D. J. Skinner, A. F. Krivanek, J. H. Crouch, and Y. Xu. 2008. "Development of a Seed DNA-Based Genotyping System for Marker-Assisted Selection in Maize." *Molecular Breeding* 22: 477–494. https://doi.org/10.1007/s11032-008-9192-4.

Goyoaga, C., C. Burbano, C. Cuadrado, et al. 2008. "Content and Distribution of Vicine, Convicine and l-DOPA During Germination and Seedling Growth of Two *Vicia faba* L. Varieties." *European Food Research and Technology* 227: 1537–1542. https://doi.org/10.1007/s00217-008-0876-0.

Griffiths, D. W., and G. Ramsay. 1996. "The Distribution of Pyrimidinone Glucosides in Developing Seedlings of *Vicia faba* and *Vicia narbonensis*." *Journal of the Science of Food and Agriculture* 72: 469–475. https://doi.org/10.1002/(SICI)1097-0010(199612)72:4%3c469::AID-JSFA681%3e3.0.CO;2-P.

Grimanelli, D., E. Perotti, J. Ramirez, and O. Leblanc. 2005. "Timing of the Maternal-To-Zygotic Transition During Early Seed Development in Maize." *Plant Cell* 17, no. 4: 1061–1072. https://doi.org/10.1105/tpc.104.029819.

Grini, P. E., G. Jurgens, and M. Hulskamp. 2002. "Embryo and Endosperm Development Is Disrupted in the Female Gametophytic Capulet Mutants of Arabidopsis." *Genetics* 162: 1911–1925.

Gutierrez, N., M. J. Giménez, C. Palomino, and C. M. Avila. 2011. "Assessment of Candidate Reference Genes for Expression Studies in Vicia faba L. by Real-Time Quantitative PCR." *Molecular Breeding* 28: 13–24. https://doi.org/10.1007/s11032-010-9456-7.

Hua, W., R.-J. Li, G.-M. Zhan, et al. 2012. "Maternal Control of Seed Oil Content in *Brassica napus*: The Role of Silique Wall Photosynthesis." *Plant Journal* 69: 432–444. https://doi.org/10.1111/j.1365-313X.2011.04802.x.

Jensen, E. S., M. B. Peoples, and H. Hauggaard-Nielsen. 2010. "Faba Bean in Cropping Systems." *Field Crops Research* 115, no. 3: 203–216. https://doi.org/10.1016/j.fcr.2009.10.008.

Khazaei, H., R. W. Purves, J. Hughes, et al. 2019. "Eliminating Vicine and Convicine, the Main Anti-Nutritional Factors Restricting Faba Bean Usage." *Trends in Food Science & Technology* 91: 549–556. https://doi.org/10.1016/j.tifs.2019.07.051.

Khazaei, H., R. W. Purves, M. Song, et al. 2017. "Development and Validation of a Robust, Breeder-Friendly Molecular Marker for the VC—Locus in Faba Bean." *Molecular Breeding* 37: 140. https://doi.org/10.1007/s11032-017-0742-5.

Livak, K. J., and T. D. Schmittgen. 2001. "Analysis of Relative Gene Expression Data Using Realtime Quantitative PCR and the 2ΔΔC(T) Method." *Methods* 25: 402–408. https://doi.org/10.1006/meth.2001.1262.

Luzzatto, L., and P. Arese. 2018. "Favism and Glucose-6-Phosphate Dehydrogenase Deficiency." *New England Journal of Medicine* 378: 1068–1069.

O'Donovan, T. and Dunne, D. 2021. *Establishing High-Yielding Faba Bean*. Legume Hub. http://www.legumehub.eu.

Ohto, M., S. L. Stone, and J. J. Harada. 2008. "Genetic Control of Seed Development and Seed Mass." In *Seed Development, Dormancy and Germination*, edited by K. J. Bradford and H. Nonogaki, 1–24. Blackwell.

Pavlík, M., M. Váňová, V. Laudová, and J. Harmatha. 2002. "Fungitoxicity of Natural Heterocycle Glucoside Vicine Obtained From *Vicia faba* L. Against Selected Microscopic Filamentous Fungi." *Plant, Soil and Environment* 48: 543–547. https://doi.org/10.17221/4410-PSE.

Pitz, W. J., F. W. Sosulski, and C. C. Rowland. 1981. "Effect of Genotype and Environment on Vicine and Convicine Levels in Faba Beans (*Vicia faba* Minor)." *Journal of the Science of Food and Agriculture* 32: 1–8. https://doi.org/10.1002/jsfa.2740320102.

Ramsay, G., and D. W. Griffiths. 1996. "Accumulation of Vicine and Convicine in *Vicia faba* and *Vicia narbonensis*." *Phytochemistry* 42: 63–67. https://doi.org/10.1016/0031-9422(95)00870-5.

Ray, H. and F. Georges. 2010. "A Genomic Approach to Nutritional, Pharmacological and Genetic Issues of Faba Bean (*Vicia faba*): Prospects for Genetic Modifications." *GM Crops* 1, no. 2: 99–106. https://doi.org/10.4161/gmcr.1.2.11891.

R Core Team. 2021. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ron-Parra, J., M. M. M. Rivera, J. J. López, A. A. J. M. Cordero, L. D. L. C. Larios, and J. J. S. Gonzalez. 2016. "Maternal Genetic Inheritance of Red Pericarp in the Grain of Maize." *Maydica* 61: 1–5.

Sergeant, K., S. Goertz, S. Halime, et al. 2024. "Exploration of the Diversity of Vicine and Convicine Derivatives in Faba Bean (*Vicia faba* L.) Cultivars: Insights From LC-MS/MS Spectra." *Molecules* 29, no. 5: 1065. https://doi.org/10.3390/molecules29051065.

Skovbjerg, C. K., J. N. Knudsen, W. Füchtbauer, et al. 2020. "Evaluation of Yield, Yield Stability, and Yield–Protein Relationship in 17 Commercial Faba Bean Cultivars." *Legume Science* 2: e39. https://doi.org/10.1002/leg3.39.

Sun, L., C.-H. Yu, P. Y. Yung, and J. S. Weissman. 2020. "Effective Polyploidy Causes Phenotypic Delay in Antibiotic Resistance." *PLoS Computational Biology* 16, no. 2: e1007930. https://doi.org/10.1371/journal.pcbi.1007930.

Tacke, R. . 2023. "Genetic and Environmental Effects on Vicine and Convicine Levels in Faba Bean (Vicia faba L.)" [Doctoral Dissertation, Georg-August-Universität Göttingen]. Göttingen University Repository.

Tacke, R., W. Ecke, M. Höfer, O. Sass, and W. Link. 2022. "Fine-Mapping of the Major Locus for Vicine and Convicine in Faba Bean (*Vicia faba*) and Marker-Assisted Breeding of a Novel, Low Vicine and Convicine Winter Faba Bean Population." *Plant Breeding* 141: 644–657. https://doi.org/10.1111/pbr.13039.

Ugwuanyi, S., M. Makhoul, A. A. Golicz, C. Obermeier, and R. J. Snowdon. 2025. "Multiple Copy Number Variants of VC1 Gene Reveal Single-Copy Expression as a Key Determinant of Vicine Content." *Frontiers in Plant Science* 16: 1565210. https://doi.org/10.3389/fpls.2025.1565210.

Ugwuanyi, S., H. Tietgen, M. Makhoul, C. Obermeier, and R. Snowdon. 2025. "Novel Gene-Copy-Specific KASP Markers for Efficient Marker-Assisted Development of Low Vicine Faba Bean (*Vicia faba* L.)." *Plant Breeding* 144, no. 4: 573–579. https://doi.org/10.1111/pbr.13276.

Zaar, A., and C. Balko. 2022. "Klimaeinfluss auf die Qualität von Ackerbohnen (Vicia faba L.) bezüglich des Vicin- und Convicingehalts." In *55. Jahrestagung DGQ: Qualität 2030: Produktqualität in Zeiten des globalen Wandels*, vol. 469, 20. Julius-Kühn-Institut. https://doi.org/10.5073/20220404-103405.

Zebosi, B., J. Ssengo, L. F. Geadelmann, E. Unger-Wallace, and E. Vollbrecht. 2025. "An Effective and Safe Maize Seed Chipping Protocol Using Clipping Pliers With Applications in Small-Scale Genotyping and Marker-Assisted Breeding." *Bio-Protocol* 15, no. 3: e5200. https://doi.org/10.21769/BioProtoc.5200.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** cDNA sequences of VC1 and RIBA2 genes from selected parental lines. **Table S2:**: Genotypic data for F2 population based on the KASP assay analyses of SNP_b1 and SNP08.

5 General Discussion

5.1 Multiple tandem duplication drive copy number variation in the v-c major locus

Genomic studies in faba bean were long constrained by the absence of high-quality reference genomes. However, the recent release of complete reference genome assemblies with functional gene annotations for two important cultivars, Hedin2 (high v-c cultivar) and Tiffany (low v-c cultivar), has provided an invaluable resource for studying the genetic control of v-c biosynthesis (Jayakodi et al., 2023). These assemblies allowed for a more comprehensive investigation of the v-c trait, particularly given that the reference cultivars exhibit contrasting seed v-c levels.

A breakthrough in understanding v-c biosynthesis came when Björnsdotter et al. (2021) identified *VC1* as a key gene catalyzing the initial step of v-c biosynthesis. *VC1* encodes a bifunctional riboflavin protein with two enzymatic domains: 3,4-dihydroxy-2-butanone-4-phosphate synthase and GTP cyclohydrolase II, both involved in the riboflavin biosynthetic pathway. GCHII catalyzes the conversion of GTP into riboflavin precursors, including 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5′-phosphate (DARPP) and 5-amino-6-ribosylamino-2,3(1H,3H)-pyrimidinedione 5′-phosphate (ARPDP), which share structural similarity with vicine and convicine. Based on this, v-c biosynthesis is thought to occur as a byproduct of the riboflavin pathway. A 2-bp insertion within the GCHII domain of *VC1* disrupts its enzymatic function, leading to reduced v-c content (Björnsdotter et al., 2021). Despite these insights, several questions remain unanswered, such as why inactivation of *VC1* does not completely eliminate v-c biosynthesis, and what causes the residual v-c content in low v-c cultivars. Additionally, it is still unclear whether other riboflavin biosynthetic genes exist in the faba bean genome. To address these gaps, the *RIBA1* protein was mapped onto the newly available genome assemblies to identify all related genes.

The results described in Ugwuanyi et al. (2025a) revealed that multiple gene models encode bifunctional riboflavin proteins in faba bean, with most showing high similarity to *VC1*, indicating the presence of multiple *VC1* copies. Notably, three tandemly duplicated *VC1* copies were identified in the same genomic region in Hedin2, while four copies were found in Tiffany. This supports previous reports that many faba bean genes are affected by tandem duplication (Jayakodi et al., 2023). However, sequence variations were observed between these gene copies, particularly in exon 3, intron 3, and exon 6, leading to the classification of three major *VC1* variants. Interestingly, unlike earlier findings which suggested that low v-c genotypes exclusively carry only the mutant *VC1* variant while high v-c genotypes carry only the wild-type *VC1* (Björnsdotter et al., 2021; Tacke et al., 2022), the low v-c cultivar Tiffany was found

to carry both wild-type and mutant variants. This discrepancy may stem from previous studies relying on transcript-based approaches, which lack the resolution provided by reference genomes.

The analysis revealed substantial *VC1* copy number variation (CNV) among different faba bean genotypes, with gene copies ranging from two to five. However, expression levels did not correlate with copy number, suggesting the presence of dosage compensation mechanisms. This observation is similar to the findings on hilum colour, where an increased copy number of the underlying polyphenol oxidase gene does not correspond to trait expression (Jayakodi et al., 2023). It also aligns with previous evidence of a non-dosage effect at the *VC1* locus, as demonstrated by the transformation of Hedin2 with an additional *VC1* coding sequence (Björnsdotter et al., 2021). Instead, v-c content was determined by the expression patterns of individual gene variants. For example, low v-c genotypes exclusively expressed the variant carrying the 2-bp (AT) insertion, while high v-c genotypes expressed only the wild-type variant, even when both variants were present in the genome. This may explain why transcript-based studies reported only the wild-type or mutant variant in high and low v-c genotypes, respectively (Björnsdotter et al., 2021; Tacke et al., 2022).

The lack of dosage effect suggests the involvement of complex regulatory mechanisms, potentially including microRNAs (miRNAs), DNA methylation, or other epigenetic modifications (Vaquerizas et al., 2009; Woodwark and Bateman, 2011). Generally, miRNAs are known to fine-tune gene expression, and some studies suggest they preferentially regulate CNV genes (Bartel, 2009; Woodwark and Bateman, 2011), while some findings contradict this (Vasudevan et al., 2007; Li et al., 2008; Woodwark and Bateman, 2011; Chang and Liao, 2012).

Additionally, dosage compensation mechanisms such as imprinting or monoallelic expression are well documented (Pauler et al., 2007). DNA methylation is also a well-established mechanism for silencing duplicated genes (Suzuki and Bird, 2008; Law and Jacobsen, 2010; Chang and Liao, 2012). The faba bean genome is among the most highly methylated plant genomes (Jayakodi et al., 2023). It is therefore plausible that methylation may play a role in silencing additional *VC1* copies, thereby restricting expression to a single copy per genotype.

5.2 *RIBA2* is a potential minor effect v-c locus in faba bean

While *VC1* inactivation explains the low v-c phenotype, the persistence of residual v-c biosynthesis in *VC1* mutants suggests that additional loci may be contributing to the trait. In Ugwuanyi et al. (2025a), a second riboflavin biosynthetic gene, *RIBA2*, which shares high

sequence similarity with *VC1* was identified. Structurally, *RIBA2* differs from *VC1*, particularly in non-coding regions, although both genes show high similarity and clear conservation in their coding sequences and corresponding amino acid profiles. This indicates that *RIBA2* represents a distinct, second RIBA locus in faba bean. Bioinformatic analysis confirmed that *RIBA2* encodes a functional bifunctional riboflavin protein, with conserved catalytic domains for DHBPS and GCHII, enzymes known to participate in the riboflavin biosynthetic pathway in plants (Hiltunen et al., 2012). Key residues essential for enzymatic activity, including those involved in zinc ion binding, were highly conserved in both the wild-type *VC1* and this novel homolog (Kaiser et al., 2002). Additionally, transcript analysis revealed active *RIBA2* expression in seeds during early and late seed-filling stages, further supporting its functional role.

The discovery of another riboflavin locus is not unexpected, given that riboflavin (vitamin B2) is essential for plant growth and development. It serves as a precursor to flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), both of which are critical cofactors in numerous metabolic processes (Hiltunen et al., 2012). The presence of this second locus may explain why genotypes carrying mutant *VC1* gene still undergo normal growth and development.

Furthermore, gene expression results suggest that *RIBA2* may function as a minor-effect locus for riboflavin biosynthesis, with expression levels approximately eightfold lower than those of *RIBA1*. In high v-c genotypes, total riboflavin gene expression was at least ten times higher than *RIBA2* expression alone in low v-c genotypes, where *RIBA2* may represent the only actively expressed riboflavin gene. This supports the hypothesis that *RIBA2* could be responsible for the residual v-c content observed in these genotypes. These findings are consistent with previous studies identifying *VC1* as the major determinant of v-c biosynthesis, while also reinforcing earlier reports of additional minor-effect QTL (Ramsay et al., 1995; Gutiérrez et al., 2006; Khazaei et al., 2015; Tacke et al., 2022). Moreover, single nucleotide polymorphisms within *RIBA2* were found to co-segregate with v-c content. Functional validation of the role of *RIBA2* will require gene-editing experiments, such as CRISPR-Cas9 knockouts in a *vc1* background. However, complete inactivation of both *vc1* and *riba2* could be detrimental, as riboflavin biosynthesis is essential for plant viability. If this is the case, eliminating v-c content may require targeting downstream genes in the v-c biosynthetic pathway.

5.3 The presence of multiple *VC1* gene variants within genotypes impairs the accuracy of marker selection in low v-c breeding

The discovery of the low v-c trait by Duc et al. (1989) opened new opportunities for integrating this trait into elite faba bean cultivars. Furthermore, the identification of *VC1* as a major gene controlling v-c content laid a strong foundation for developing locus-specific molecular markers to facilitate efficient selection of low v-c genotypes in breeding programs. As expected, several gene-based markers have been developed (Khazaei et al., 2017; Tacke et al., 2022). However, these markers often do not performed as expected, as their predictive accuracy remains unreliable across diverse faba bean breeding populations. In some cases, the markers show false heterozygous clusters for homozygous genotypes.

Findings from Ugwuanyi et al. (2025a) suggest that this lack of accurate prediction stems from the presence of multiple *VC1* variants within faba bean genotypes. It was observed that genotypes carry multiple gene copies, and in some cases, these copies represent different variants. Despite this, only a single copy is actively expressed per genotype. However, existing markers were designed to target polymorphisms across the *VC1* locus, not specifically for individual variants. As a result, the markers fail to distinguish between expressed and unexpressed variants, leading to false heterozygous signals and incorrect phenotypic classification when multiple variants are present in a genotype.

This hypothesis was further confirmed by comparing marker genotypes in KASP assays between genomic DNA (gDNA) and complementary DNA (cDNA), considering that gDNA includes all gene copies present in the genome, while cDNA reflects only expressed genes. KASP genotyping using gDNA frequently resulted in inaccurate calls and incorrect v-c phenotype predictions, as expected due to interference from unexpressed gene copies. In contrast, cDNA-based KASP genotyping consistently produced accurate homozygous genotype calls that aligned with the actual v-c phenotypes.

These findings indicate that the structural complexity of the *VC1* locus poses a major challenge to the application of molecular markers in breeding for low v-c content. This challenge is not unique to faba bean; similar issues have been reported in other crops with complex genomes or regions, where molecular markers have shown reduced accuracy (Makhoul et al., 2020). A potential approach to overcome this limitation is to modify marker development strategies, for example, by designing gene-copy-specific markers that exclusively target the expressed gene variant rather than all *VC1* copies. This would enhance prediction accuracy and enable more reliable application in breeding programs.

5.4 Genomic resources to facilitate marker-assisted breeding of low vicine faba bean

The slow progress in developing low v-c cultivars has largely been due to the lack of reliable selection markers. The availability of efficient v-c markers would therefore accelerate the breeding of low v-c faba bean varieties, reducing reliance on labor-intensive and costly phenotyping methods (Gutierrez et al., 2006; Khazaei et al., 2017). Building on the improved understanding of the genetic control of v-c biosynthesis provided by the results in Ugwuanyi et al. (2025a), a set of five gene-copy-specific Kompetitive Allele-Specific PCR (KASP) markers was developed and validated for marker-assisted selection in Ugwuanyi et al. (2025b). One of these markers targets the 2-bp insertion mutation in *VC1*, while the remaining four are based on polymorphisms within the *RIBA2*, a single-copy riboflavin biosynthesis gene.

The 2-bp mutation was successfully converted into a selection marker by incorporating a target pre-amplification (TPA) step prior to KASP genotyping. This strategy helped overcome bias caused by multiple unexpressed gene copies. By using a pair of conventional PCR primers, the TPA step reduces genome complexity and selectively amplifies only the expressed *VC1* variant that influences v-c content. The resulting amplicon is then genotyped using the KASP assay, effectively avoiding interference from non-expressed gene variants. The TPA-KASP method was streamlined into a simple workflow that integrates the PCR and KASP steps on a single assay plate, making it efficient, convenient, and suitable for high-throughput applications. Validation in a diverse set of 97 faba bean genotypes demonstrated 100% prediction accuracy, successfully clustering genotypes into high and low v-c groups and confirming its utility in v-c breeding programs. In addition to the *VC1*-copy-specific marker, four KASP markers were developed targeting polymorphisms within *RIBA2*. Similarly, validation in a diverse set of faba bean genotypes confirmed their high selection accuracy, further demonstrating their suitability for breeding programs.

Until now, breeding programs for low v-c faba bean have relied on laborious phenotyping methods such as spectrophotometric/colorimetric assays (Sixdenier et al., 1996) or liquid chromatography (Marquardt and Frohlich, 1981; Khamassi et al., 2013; Pulkkinen et al., 2015). These methods are time-consuming, expensive and destructive, making them impractical for large-scale breeding applications. Moreover, they do not allow phenotyping at early developmental stages, thereby limiting their utility in selection. The v-c markers presented in Ugwuanyi et al. (2025b) offer a cost-effective, non-destructive, and high-throughput alternative for screening low v-c genotypes. These markers enable selection as early as the F2 and F3 generations, reducing breeding cycle time and resource expenditure. These markers will

enhance the efficiency of faba bean breeding programs and accelerate the development and adoption of low v-c cultivars.

5.5 Maternal tissue-localized gene expression controls the inheritance of seed v-c content

While many studies have reported that seed v-c content is maternally determined (Brown and Roberts, 1972; Duc et al., 1989; Tacke, 2023), the molecular mechanisms underlying this pattern of inheritance remain unclear. To investigate this, multiple reciprocal cross populations between high and low v-c parental lines were used to dissect the mechanisms of maternal inheritance at the level of gene expression (Ugwuanyi et al., 2025c). The findings demonstrate that the expression of v-c-related genes is restricted to only maternal tissues, including leaves, flowers, and the seed coat, with no detectable gene activity in the embryo. Specifically, gene expression localized in the seed coat determines v-c levels in seeds, indicating that seed v-c content is controlled by the maternal genotype. This supports the hypothesis that v-c is synthesized in maternal tissues and subsequently transported into the embryo for storage (Pitz et al., 1981; Ray and Georges, 2010; Björnsdotter et al., 2021).

This mechanism is consistent with maternal control of many seed traits observed in other crops. For example, in maize (*Zea mays*) and Arabidopsis (*Arabidopsis thaliana*), early seed development is driven by maternal transcripts, with delayed activation of the zygotic genome (Grimanelli et al., 2005; Autran et al., 2011). Similarly, seed oil content in Brassica is determined by maternal photosynthetic tissues (Hua et al., 2012) while seed coat pigmentation in Arabidopsis and maize is also maternally regulated (Debeaujon et al., 2000; Ron-Parra et al., 2016). In faba bean, the transcriptional silencing of v-c genes in the embryo may reflect a similar mechanism, representing a developmental-stage-specific, maternally regulated control of seed phenotype.

5.6 Conclusion

This thesis provides critical insights into the complex genetic regulation of v-c biosynthesis in faba bean, revealing copy number variation and single-copy expression of the *VC1* gene, and a contributing role of a second riboflavin biosynthetic gene, *RIBA2*, as key factors influencing v-c content in faba bean. This knowledge was successfully leveraged to develop highly robust genomic tools for MAS, overcoming the limitations of previous approaches and providing valuable tools for more efficient breeding of low v-c cultivars. These tools will accelerate breeding programs, reduce reliance on costly phenotyping and enable early-generation

selection, making the development of low v-c faba bean cultivars faster and more efficient. However, the role of *RIBA2* in v-c biosynthesis is yet to be validated by genome editing knockouts. As a result, future studies should therefore focus on functional validation of *RIBA2* and exploring the epigenetic mechanisms regulating *VC1* expression.

6 Summary

Summary

Vicine and convicine (v-c) are antinutritional compounds that significantly limit the global use of faba bean (*Vicia faba*) for food and feed. Consequently, reducing or eliminating these compounds through breeding is highly desirable. However, progress has been hindered by an incomplete understanding of the molecular mechanisms underlying their biosynthesis, which in turn has limited the effective application of genomic tools in breeding programs. Recently, the bifunctional riboflavin gene *VC1* was identified as a key enzyme in the v-c biosynthetic pathway. A frameshifting AT insertion in the coding sequence of this gene disrupts its function, leading to reduced, but not eliminated, v-c content. The availability of high-quality genome assemblies for two cultivars, Hedin2 and Tiffany, which differ in v-c content, has enabled deeper insights into the genetic basis of this trait.

The work described in this thesis revealed multiple tandem duplications of the *VC1* gene and consequent copy number variation, with different genotypes carrying between two and five copies. These copies, which carry structural variants, including insertions and deletions in exons 3 and 6 and intron 3, were classified into three major variants: *VC1*A, *VC1*B, and *VC1*C. High v-c genotypes were found to predominantly carry *VC1*A, while low v-c genotypes were characterized by the presence of *VC1*B and *VC1*C. Interestingly, copy number does not correspond to gene expression level. Instead, expression of a single variant appears to determine v-c content, suggesting a regulatory dosage compensation mechanism, possibly involving DNA methylation or microRNA activity. Importantly, low v-c genotypes still carry active *VC1* variants that remain unexpressed in seeds during both early and late-seed filling stages.

Furthermore, a neighbouring homolog of *VC1*, denominated as *RIBA2*, was identified. It shares high sequence similarity with *VC1* but differs structurally. Like *VC1* and other bifunctional riboflavin genes, *RIBA2* contains two expression domains, RibB and RibA, encoding 3,4-dihydroxy-2-butanone-4-phosphate synthase and GTP cyclohydrolase II, respectively, suggesting that it may function as a secondary riboflavin biosynthetic gene in *V. faba*. *RIBA2* accounts for approximately 5% of riboflavin gene transcripts in immature seeds. While *VC1* remains the primary determinant of v-c content, *RIBA2* may contribute as a minor-effect locus, potentially explaining residual v-c production in *VC1* mutants. However, functional validation via gene editing is needed to confirm its role.

Additionally, multiple reciprocal populations derived from high and low v-c lines were used to investigate maternal inheritance of the trait at the gene expression level. The findings showed that v-c-related gene expression is confined to maternal tissues, with no detectable activity in

the embryo. Specifically, gene expression in the seed coat regulates v-c levels in seeds, indicating that maternal genotype determines seed v-c content through developmental stage-specific expression.

Building on these insights, a set of reliable molecular markers were developed to support v-c breeding programs. One marker targets a previously described AT insertion in *VC1*, while four others are based on novel polymorphisms that were detected within *RIBA2*. To mitigate marker bias from multiple *VC1* copies, a TPA-KASP approach was implemented, combining PCR and KASP genotyping in a single assay plate. The method facilitates simple implementation as a diagnostic tool for marker-assisted selection, significantly improving selection accuracy while considerably reducing the cost and effort of screening in breeding populations. Subsequent validation of the marker set in 97 faba bean genotypes revealed 100% accuracy of the assay, effectively distinguishing high and low v-c lines. These markers offer a high-throughput, cost-effective, and early-generation selection tool, critically enabling efficient screening for low v-c genotypes at the F2/F3 stage to enhance breeding efficiency.

In conclusion, the methods, tools and results in this thesis advance our understanding of the complex genetic regulation of v-c biosynthesis and provides valuable genomics-based breeding tools to support the development of high-quality faba bean protein.

7 Zusammenfassung

Zusammenfassung

Vicin und Convicin (v-c) sind antinutritive Verbindungen, die die weltweite Nutzung der Ackerbohne (*Vicia faba*) als Nahrungs- und Futtermittel erheblich einschränken. Daher ist eine Reduktion oder Eliminierung dieser Verbindungen durch Züchtung von großem Interesse. Der Fortschritt wurde jedoch durch ein unvollständiges Verständnis der molekularen Mechanismen ihrer Biosynthese behindert, was wiederum den effektiven Einsatz genomischer Werkzeuge in Züchtungsprogrammen begrenzte. Kürzlich wurde das bifunktionale Riboflavin-Gen *VC1* als Schlüsselenzym im v-c-Biosyntheseweg identifiziert. Eine frameshift-verursachende AT-Insertion in der kodierenden Sequenz dieses Gens unterbricht dessen Funktion, was zu einer Verringerung, jedoch nicht zu einer vollständigen Eliminierung des v-c-Gehalts führt. Die Verfügbarkeit hochwertiger Genomassemblies für zwei Sorten, Hedin2 und Tiffany, die sich im v-c-Gehalt unterscheiden, ermöglichte tiefere Einblicke in die genetische Grundlage dieses Merkmals.

Die in dieser Dissertation beschriebenen Arbeiten zeigten mehrfache tandemartige Duplikationen des *VC1*-Gens und daraus resultierende Kopienzahlvariationen, wobei verschiedene Genotypen zwischen zwei und fünf Kopien tragen. Diese Kopien, die strukturelle Varianten wie Insertionen und Deletionen in Exon 3 und 6 sowie Intron 3 enthalten, wurden in drei Hauptvarianten eingeteilt: *VC1A*, *VC1B* und *VC1C*. Genotypen mit hohem v-c-Gehalt trugen überwiegend *VC1A*, während Genotypen mit niedrigem v-c-Gehalt durch das Vorhandensein von *VC1B* und *VC1C* gekennzeichnet waren. Interessanterweise entsprach die Kopienzahl nicht dem Expressionsniveau des Gens. Stattdessen scheint die Expression einer einzelnen Variante den v-c-Gehalt zu bestimmen, was auf einen regulatorischen Dosis-Kompensationsmechanismus hindeutet, der möglicherweise DNA-Methylierung oder MikroRNA-Aktivität einbezieht. Bemerkenswert ist, dass Genotypen mit niedrigem v-c-Gehalt weiterhin aktive *VC1*-Varianten tragen, die jedoch nicht exprimiert wurden.

Darüber hinaus wurde ein benachbartes Homolog von *VC1* identifiziert, das als *RIBA2* bezeichnet wurde. Es weist eine hohe Sequenzähnlichkeit zu *VC1* auf, unterscheidet sich jedoch strukturell. Wie *VC1* und andere bifunktionale Riboflavin-Gene enthält *RIBA2* zwei Expressionsdomänen, RibB und RibA, die für 3,4-Dihydroxy-2-butanon-4-phosphat-Synthase bzw. GTP-Cyclohydrolase II kodieren, was darauf hindeutet, dass es als sekundäres Riboflavin-Biosynthesegen in V. *faba* fungieren könnte. *RIBA2* macht etwa 5 % der Riboflavin-Gen-Transkripte in unreifen Samen aus. Während *VC1* der primäre Bestimmungsfaktor für den v-c-Gehalt bleibt, könnte *RIBA2* als Lokus mit geringem Effekt beitragen und möglicherweise die

Restproduktion von v-c in *VC1*-Mutanten erklären. Eine funktionelle Validierung durch Genom-Editierung ist jedoch erforderlich, um seine Rolle zu bestätigen.

Zusätzlich wurden mehrere reziproke Populationen, die von Linien mit hohem und niedrigem v-c-Gehalt abgeleitet wurden, genutzt, um die maternale Vererbung des Merkmals auf Ebene der Genexpression zu untersuchen. Die Ergebnisse zeigten, dass die v-c-bezogene Genexpression auf maternale Gewebe beschränkt ist und im Embryo keine Aktivität nachweisbar ist. Insbesondere reguliert die Genexpression in der Samenschale den v-c-Gehalt der Samen, was darauf hinweist, dass der maternale Genotyp den Samen-v-c-Gehalt durch entwicklungsstadien-spezifische Expression bestimmt.

Aufbauend auf diesen Erkenntnissen wurde ein Set zuverlässiger molekularer Marker entwickelt, um v-c-Züchtungsprogramme zu unterstützen. Ein Marker richtet sich gegen eine bereits beschriebene AT-Insertion in *VC1*, während vier weitere auf neu entdeckten Polymorphismen innerhalb von *RIBA2* basieren. Um Marker-Bias aufgrund mehrerer *VC1*-Kopien zu vermeiden, wurde ein TPA-KASP-Ansatz implementiert, der PCR und KASP-Genotypisierung in einer einzigen Assay-Platte kombiniert. Die Methode ermöglicht eine einfache Implementierung als Diagnosetool für marker-gestützte Selektion, verbessert die Selektionsgenauigkeit erheblich und reduziert gleichzeitig die Kosten und den Aufwand für das Screening in Züchtungspopulationen deutlich. Die anschließende Validierung des Markersets an 97 Ackerbohnengenotypen zeigte eine 100%ige Genauigkeit des Assays und ermöglichte eine eindeutige Unterscheidung zwischen Linien mit hohem und niedrigem v-c-Gehalt. Diese Marker bieten ein hochdurchsatzfähiges, kosteneffizientes und früh einsetzbares Selektionswerkzeug, das eine effiziente Identifizierung von Linien mit niedrigem v-c-Gehalt bereits in der F2/F3-Generation ermöglicht und so die Züchtungseffizienz steigert.

Zusammenfassend vertiefen die in dieser Dissertation vorgestellten Methoden, Werkzeuge und Ergebnisse unser Verständnis der komplexen genetischen Regulation der v-c-Biosynthese und liefern wertvolle genomgestützte Züchtungsinstrumente zur Entwicklung hochwertiger Ackerbohnenproteine.

8 References

References

Adamidou, S., Nengas, I., Grigorakis, K., Nikolopoulou, D., and Jauncey, K. (2011). Chemical composition and antinutritional factors of field peas (Pisumsativum), chickpeas (Cicer arietinum), and faba beans (Viciafaba) as affected by extrusion preconditioning and drying temperatures. Cereal Chemistry, 88:80–86. https://doi.org/10.1094/CCHEM-05-10-0077

Adhikari, K., Khazaei, H., Ghaouti, L., Maalouf, F., Vandenberg, A., Link, W., and O'Sullivan, M. (2021). Conventional and molecular breeding tools for accelerating genetic gain in faba bean (*Vicia faba* L.). Frontiers in Plant Science, 12, 744259. doi:10.3389/fpls.2021.744259

Ali, M.B.M., Welna, G.C., Sallam, A., Martsch, R., Balko, C., Gebser, B., Sass, O. and Link, W. (2016). Association analyses to genetically improve drought and freezing tolerance of faba bean (*Vicia faba* L.). Crop Science, 56: 1036-1048. https://doi.org/10.2135/cropsci2015.08.0503

Alonso, R., Aguirre, A., and Marzo, F. (2000). Effects of extrusion and traditional processing methods on antinutrients and in vitro digestibility of protein and starch in faba and kidney beans. Food Chemistry, 68(2), 159–165. https://doi.org/10.1016/S0308-8146(99)00169-7

Anderson, J. C., Idowu, A. O., Singh, U., and Singh, B. (1994). Physicochemical characteristics of flours of faba bean as influenced by processing methods. Plant Foods for Human Nutrition, 45:371–379. https://doi.org/10.1007/BF01088087

Arese, P., Gallo, V., Pantaleo, A. and Turrini, F. (2012). Life and death of glucose-6-phosphate dehydrogenase (G6PD) deficient erythrocytes – role of redox stress and band 3 modifications. Transfusion Medicine and Hemotherapy, 39:328-334

Augustin, M. A. and Cole, M. B. (2022). Towards a sustainable food system by design using faba bean protein as an example. Trends in Food Science and Technology, 125:1-11. https://doi.org/10.1016/j.tifs.2022.04.029.

Autran, D., Baroux, C., Raissig, M.T., Lenormand, T., Wittig, M., Grob, S., Steimer, A., Barann, M., Klostermeier, U.C., Leblanc, O., Vielle-Calzada, J.P., Rosenstiel, P., Grimanelli, D. and Grossniklaus, U. (2011). Maternal epigenetic pathways control parental

contributions to Arabidopsis early embryogenesis. Cell, 145(5):707-19. doi: 10.1016/j.cell.2011.04.014.

Bai, T., Nosworthy, M. G., House, J. D., and Nickerson, M. T. (2018). Effect of tempering moisture and infrared heating temperature on the nutritional properties of desi chickpea and hull-less barley flours, and their blends. Food Research International, 108:430–439.

Bakr, A. A. (1996). Effect of Egyptian cooking methods of faba beans on its nutritive value, dietary protein utilization and iron deficiency anemia 1. The role of main technological pretreatments. Plant Foods for Human Nutrition, 49:83–92. https://doi.org/10.1007/BF01092525

Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136: 215–233

Beleia, A., Thu Thao, L. T., and Ida, E. I. (1993). Lowering phytic phosphorus by hydration of soybeans. Journal of Food Science, 58:375–378. https://doi.org/10.1111/j.1365-2621.1993.tb04278.x

Björnsdotter, E., M. Nadzieja, W. Chang, et al. (2021). *VC1* catalyses a key step in the biosynthesis of vicine in faba bean. Nature Plants, 7(7):923–931. https://doi.org/10.1038/s41477-021-00950-w

Brown, J. N., and Roberts, E. H. (1972). The inheritance of vicine and convicine content in faba bean. Heredity, 28(2):203–209. https://doi.org/10.1038/hdy.1972.62

Cappellini, M.D. and Fiorelli, G. (2008). Glucose-6-phosphate dehydrogenase deficiency. Lancet, 371(9606):64-74. doi: 10.1016/S0140-6736(08)60073-2

Caracuta, V., Weinstein-Evron, M., Kaufman, D., Yeshurun, R., Silvent, J., and Boaretto, E. (2016). 14,000-year-old seeds indicate the Levantine origin of the lost progenitor of faba bean. Scientific Reports, 6, 37399. https://doi.org/10.1038/srep37399

Cardador-Martínez, A., Maya-Ocaña, K., Ortiz-Moreno, A., Herrera-Cabrera, B. E., Dávila-Ortiz, G., Múzquiz, M., Martín-Pedrosa, M., Burbano, C., Cuadrado, C., and Jiménez-Martínez, C. (2012.) Effect of roasting and boiling on the content of vicine, convicine and L-3, 4-dihydroxyphenylalanine in *Vicia faba* L. Journal of Food Quality, 35(6):419–428.

Chaieb, N., González, J. L., López-Mesas, M., Bouslama, M., and Valiente, M. (2011). Polyphenols content and antioxidant capacity of thirteen faba bean (*Vicia faba* L.) genotypes cultivated in Tunisia. Food Research International, 44(4), 970–977.

Chang, A.Y. and Liao, B.Y. (2012). DNA methylation rebalances gene dosage after mammalian gene duplications. Mol Biol Evol. 29(1):133–44. doi: 10.1093/molbev/msr174. Epub 2011 Aug 6. PMID: 21821837

Coda, R., Melama, L., Rizzello, C. G., Curiel, J. A., Sibakov, J., Holopainen, U., Pulkkinen, M., and Sozer, N. (2015). Effect of air classification and fermentation by Lactobacillus plantarum VTT E-133328 on faba bean (*Vicia faba* L.) flour nutritional properties. International Journal of Food Microbiology, 193:34–42.

Crépon, K., Marget, P., Peyronnet, C., Marget, P., Peyronnet, C., Carrouée, B., et al. (2010). Nutritional value of faba bean (*Vicia faba* L.) seeds for feed and food. Field Crop Res. 115:329–339. doi: 10.1016/j.fcr.2009.09.016

Dai, J., and Mumper, R. J. (2010). Plant phenolics: Extraction, analysis and their antioxidant and anticancer properties. Molecules, 15(10), 7313–7352.

Debeaujon, I., Léon-Kloosterziel, K.M. and Koornneef, M. (2000). Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. Plant Physiology, 122(2):403-414. doi: 10.1104/pp.122.2.403.

Denton, M.D., Phillips, L.A., Peoples, M.B. et al. (2017). Legume inoculant application methods: effects on nodulation patterns, nitrogen fixation, crop growth and yield in narrow-leaf lupin and faba bean. Plant Soil, 419:25–39. https://doi.org/10.1007/s11104-017-3317-7

Desoky, E-S.M., Mansour, E., El-Sobky, E-S.E.A., Abdul-Hamid, M.I., Taha, T.F., Elakkad, H.A., Arnaout, S.M.A.I., Eid, R.S.M., El-Tarabily, K.A. and Yasin, M.A.T. (2021). Physio-biochemical and agronomic responses of faba beans to exogenously applied nano-silicon under drought stress conditions. Frontiers in Plant Science, 12:637783. doi: 10.3389/fpls.2021.637783

Dhull, S. B., Kidwai, M. K., Noor, R., Chawla, P., and Rose, P. K. (2022). A review of nutritional profile and processing of faba bean (*Vicia faba* L.). Legume Science, 4(3), e129. https://doi.org/10.1002/leg3.129

Duc, G. et al. (2015). Faba Bean. In: De Ron, A. (eds) Grain Legumes. Handbook of Plant Breeding, vol 10. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-2797-5_5

Duc, G., Bao, S., Baum, M., Redden, B., Sadiki, M., Suso, M. J., et al. (2010). Diversity maintenance and use of *Vicia faba* L. genetic resources. Field Crops Res, 115:270–278. doi: 10.1016/j.fcr.2008.10.003

Duc, G., G. Sixdenier, M. Lila, and V. Furstoss. (1989). Search of Genetic Variability for Vicine and Convicine Content in V. faba L.: A First Report of a Gene Which Codes for Nearly Zero-Vicine and Zero-Convicine Contents. In 1. International Workshop on 'Antinutritional Factors (ANF) in Legume Seeds'. Wageningen. 23–25 Nov 1988: Pudoc.

Duc, G., Marget, P., Page, D. and Domoney, C. (2004). Facile breeding markers to lower contents of vicine and convicine in faba bean seeds and trypsin inhibitors in pea seeds. Proceedings of the fourth international workshop on antinutritional factors in legume seeds and oilseeds, 2004/03/08-10, Toledo, Spain. In M. Musquiz, G.D. Hill, M.M. Cuadrado, C. Burbano (Eds.), Recent advances of Research in antinutritional factors in legume seeds and oilseeds, Wageningen Academic Publishers, The Netherlands (2004), pp. 281-285. https://doi.org/10.3920/978-90-8686-524-6

FAOSTAT (2025). Food and Agriculture Organization of the United Nations. Available online at: http://faostat.fao.org (accessed on April 8, 2025).

Fernando, S. (2021). Production of protein-rich pulse ingredients through dry fractionation: A review. LWT, 141, 110961.

Francis, G., Makkar, H. P. S., and Becker, K. (2001). Antinutritional factors present in plant-derived alternate fish feed ingredients and their effects in fish. Aquaculture, 199:197–227. https://doi.org/10.1016/S0044-8486(01)00526-9

Frelin, O. et al. (2015). A directed-overflow and damage-control N-glycosidase in riboflavin biosynthesis. Biochem. J. 466:137–145.

Gallo, V., Skorokhod, O.A., Simula, L.F., Marrocco, T., Tambini, E., Schwarzer, E., et al. (2018). No red blood cell damage and no hemolysis in G6PD-deficient subjects after ingestion of low vicine/convicine *Vicia faba* seeds. Blood, 131: 1621-1625

Gauttam, V. and Kalia, A.N. (2013). High performance thin layer chromatography method for simultaneous estimation of vicine, trigonelline and withaferin-a in a polyherbal antidiabetic formulation. International Journal of Pharmacy and Pharmaceutical Sciences. 5(2): 367-371

Gemede, H. F. (2014). Antinutritional factors in plant foods: Potential health benefits and adverse effects. International Journal of Nutrition and Food Sciences, 3, 284.

Goyoaga, C., Burbano, C., Cuadrado, C., Varela, A., Guillamón, E., Pedrosa, M.M., Muzquiz, M. (2008). Content and distribution of vicine, convicine and L-DOPA during germination and seedling growth of two *Vicia faba* L. varieties. European Food Research and Technology, 227 (5): 1537-1542

Graham, R.D. and Gregorio, G. (2001). Breeding for nutritional characteristics in cereals. Novartis Found Symp, 236:205-14; discussion 214-8. doi: 10.1002/9780470515778.ch15

Grimanelli D, Perotti E, Ramirez J, Leblanc O. (2005). Timing of the maternal-to-zygotic transition during early seed development in maize. Plant Cell, 17(4):1061-72. https://doi.org/10.1105/tpc.104.029819

Gu, B. J., Masli, M. D. P., and Ganjyal, G. M. (2020). Whole faba bean flour exhibits unique expansion characteristics relative to the whole flours of lima, pinto, and red kidney beans during extrusion. Journal of Food Science, 85, 404–413. https://doi.org/10.1111/1750-3841.14951

Gulewicz, P., Martinez-Villaluenga, C., Kasprowicz-Potocka, M., and Frias, J. (2014). Non-nutritive compounds in Fabaceae family seeds and the improvement of their nutritional quality by traditional processing – A review. Polish Journal of Food and Nutrition Sciences, 64(2), 75–89.

Gutierrez, N., Avila, C. M., Moreno, M. T., and Torres, A. M. (2008). Development of SCAR markers linked to zt-2, one of the genes controlling absence of tannins in faba bean. Aust. J. Agric. Res. 59, 62–68. doi: 10.1071/AR07019

Gutierrez, N., Avila, C.M., Duc, G., Marget, P., Suso, M.J., Moreno, M.T. and Torres, A.M. (2006). CAPs markers to assist selection for low vicine and convicine contents in faba bean (*Vicia faba* L.). Theoretical and Applied Genetics. 114: 59-66.

Hall, A. E., and Moraru, C. I. (2021). Effect of high pressure processing and heat treatment on in vitro digestibility and trypsin inhibitor activity in lentil and faba bean protein concentrates. LWT, 152, 112342.

Hamza, M. A., Shehata, A. M. E. T., and Stegemann, H. (1987). Effect of traditional methods of processing on the electrophoretic patterns of faba bean water soluble proteins. Plant Foods for Human Nutrition, 36, 253–262. https://doi.org/10.1007/BF01892347

Hefni, M. E., Shalaby, M. T., and Witthöft, C. M. (2015). Folate content in faba beans (Viciafaba L.)—Effects of cultivar, maturity stage, industrial processing, and bioprocessing. Food Science and Nutrition, 3, 65–73. https://doi.org/10.1002/fsn3.192

Hendawey, M.H. and Younes, A.M.A. (2013). Biochemical evaluation of some faba bean cultivars under rainfed conditions at El-Sheikh Zuwayid. Annals of Agricultural Sciences, 58(2): 183-193. https://doi.org/10.1016/j.aoas.2013.07.010

Herridge, D.F., Peoples, M.B., Boddey, R.M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil 311:1–18

Hiltunen, H.-M.; Illarionov, B.; Hedtke, B.; Fischer, M.; Grimm, B. (2012). Arabidopsis RIBA Proteins: Two out of Three Isoforms Have Lost Their Bifunctional Activity in Riboflavin Biosynthesis. Int. J. Mol. Sci, 13, 14086–14105. doi:10.3390/ijms131114086

Hua, W., Li, R.-J., Zhan, G.-M., Liu, J., Li, J., Wang, X.-F., Liu, G.-H. and Wang, H.-Z. (2012), Maternal control of seed oil content in Brassica napus: the role of silique wall photosynthesis. The Plant Journal, 69: 432-444. https://doi.org/10.1111/j.1365-313X.2011.04802.x

IPCC Summary for Policymakers. In Special Report on Climate Change and Land (eds Shukla, P. R. et al.) (IPCC, 2019).

Jamalian, J. (1999). Removal of favism-inducing factors vicine and convicine and the associated effects on the protein content and digestibility of faba beans (*Vicia faba* L.). Journal of the Science of Food and Agriculture, 79, 1909–1914.

Jamalian, J. and Ghorbani, M. (2005), Extraction of favism-inducing agents from whole seeds of faba bean (*Vicia faba* L var major). J. Sci. Food Agric., 85: 1055-1060. https://doi.org/10.1002/jsfa.2075

Jayakodi, M., A. A. Golicz, J. Kreplak, et al. (2023). The giant diploid faba genome unlocks variation in a global protein crop. Nature 615(7953) 652–659. https://doi.org/10.1038/s41586-023-05791-5

Jensen, E.S., Peoples, M.B. and Hauggaard-Nielsen, H. (2010). Faba bean in cropping systems. Field Crops Research, 115(3): 203-216. https://doi.org/10.1016/j.fcr.2009.10.008

Kader, Z. M. A. (1995). Study of some factors affecting water absorption by faba beans during soaking. Food Chemistry, 53, 235–238. https://doi.org/10.1016/0308-8146(95)93926-I

Kaiser, J.; Schramek, N.; Eberhardt, S.; Puttmer, S.; Schuster, M.; Bacher, A. (2002). Biosynthesis of vitamin B2. Eur. J. Biochem, 269, 5264–5270

Karkanis, A., Ntatsi, G., Lepse, L., Fernández, J.A., Vågen, I.M., Rewald, B., Alsina, I., Kronberga, A., Balliu, A., Olle, M., Bodner, G., Dubova, L., Rosa, E. and Savvas, D. (2018). Faba bean cultivation – revealing novel managing practices for more sustainable and competitive European cropping systems. Frontiers in Plant Science, 9:1115. doi: 10.3389/fpls.2018.01115

Khalil, A. (1995). The effect of cooking, autoclaving and germination on the nutritional quality of faba beans. Food Chemistry, 54, 177–182.

Khamassi K, Ben Jeddi F, Hobbs D, et al. (2013). A baseline study of vicine–convicine levels in faba bean (*Vicia faba* L.) germplasm. Plant Genetic Resources, 11(3):250-257. doi:10.1017/S1479262113000105

Khazaei, H., O'Sullivan, D.M., Jones, H., Pitts, N., Sillanpää, M.J., Pärssinen, P., Manninen, O. and Stoddard, F.L. (2015). Flanking SNP markers for vicine-convicine concentration in faba bean (*Vicia faba* L.). Molecular Breeding, 35: 38.

Khazaei, H., O'Sullivan, D. M., Stoddard, F. L., Adhikari, K. N., Paull, J. G., Schulman, A. H., Andersen, S. U., and Vandenberg, A. (2021). Recent advances in faba bean genetic and genomic tools for crop improvement. Legume Science, 3(3), e75. https://doi.org/10.1002/leg3.75

Khazaei, H., Purves, R. W., Hughes, J., Link, W., O'Sullivan, D. M., Schulman, A. H., Björnsdotter, E., Geu-Flores, F., Nadzieja, M., Andersen, S. U., Stougaard, J., Vandenberg, A. and Stoddard, F. L. (2019). Eliminating vicine and convicine, the main anti-nutritional factors restricting faba bean usage. Trends in Food Science and Technology, 91:549-556. https://doi.org/10.1016/j.tifs.2019.07.051

Khazaei, H., Purves, R.W., Song, M., Stonehouse R., Bett, K.E., Stoddard, F.L. and Vandenberg, A. (2017). Development and validation of a robust, breeder-friendly molecular marker for the vc - locus in faba bean. Mol Breeding 37, 140. doi:10.1007/s11032-017-0742-5

Köpke, U., and Nemecek, T. (2010). Ecological services of faba bean. Field Crops Res. 115, 217–233. doi: 10.1016/j.fcr.2009.10.012

Labba, I.M., Frøkiær, H. and Sandberg, A. (2021). Nutritional and antinutritional composition of fava bean (*Vicia faba* L., var. minor) cultivars. Food Research International, 140. https://doi.org/10.1016/j.foodres.2020.110038.

Law, JA, Jacobsen, SE. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet. 11:204–220

Li J, Musso G, Zhang Z (2008) Preferential regulation of duplicated genes by microRNAs in mammals. Genome Biol 9: R132.

Link, W., Balko, C., and Stoddard, F. L. (2010). Winter hardiness in faba bean: physiology and breeding. Field Crops Res. 115, 287–296. doi: 10.1016/j.fcr.2008.08.004

Liu, C., Pei, R., and Heinonen, M. (2022). Faba bean protein: A promising plant-based emulsifier for improving physical and oxidative stabilities of oil-in-water emulsions. Food Chemistry, 369, 130879. https://doi.org/10.1016/j.foodchem.2021.130879

Liu, L., Knight, J. D., Lemke, R. L., and Farrell, R. E. (2019). A side-by-side comparison of biological nitrogen fixation and yield of four legume crops. Plant Soil. 442, 169–182. doi: 10.1007/s11104-019-04167-x

Lizarazo, C. I., Lampi, A. M., Sontag-Strohm, T., Liu, J., Piironen, V., and Stoddard, F. L. (2015). Nutritive quality and protein production from grain legumes in a boreal climate. J. Sci. Food Agric. 95, 2053–2064. doi: 10.1002/jsfa.6920

Longobardi, F., Sacco, D., Casiello, G., Ventrella, A., and Sacco, A. (2015). Chemical profile of the carpino broad bean by conventional and innovative physicochemical analyses. J. Food Qual. 38, 273–284. doi: 10.1111/jfq.12143

Luo,Y-W., and Xie, W-H. (2013). Effect of different processing methods on certain antinutritional factors and protein digestibility in green and white faba bean (*Vicia faba* L.). CYTA- Food. 11, 43–49. doi: 10.1080/19476337.2012.681705

Marquardt, R.R. and Frohlich, A.A. (1981). Rapid reversed-phase high-performance liquid chromatography method for the quantitation of vicine, convicine and related compounds. Journal of Chromatography A, 208 (2): 373-379 https://doi.org/10.1016/S0021-9673(00)81950-0

Martineau-Côté, D., Achouri, A., Pitre, M., Wanasundara, J., Karboune, S., and L'Hocine, L. (2023). Investigation of the nutritional quality of raw and processed Canadian faba bean (*Vicia faba* L.) flours in comparison to pea and soy using a human in vitro gastrointestinal digestion model. Food Research International, 173, 113264.

Meletis, J. and Konstantopoulos, K. (2004). Favism - from the "avoid fava beans" of Pythagoras to the present. Haema, 7: 17-21

Mesfin T (2020). Impacts of Soil Acidity on Growth Performance of Faba bean (*Vicia faba* L.) and Management Options. Acad. Res. J. Agri. Sci. Res. 8(4): 423-431

Multari, S., Stewart, D. and Russell, W.R. (2015), Potential of Fava Bean as Future Protein Supply to Partially Replace Meat Intake in the Human Diet. COMPREHENSIVE REVIEWS IN FOOD SCIENCE AND FOOD SAFETY, 14: 511-522. https://doi.org/10.1111/1541-4337.12146

Nasar-Abbas, S. M., Siddique, K. H. M., Plummer, J. A., White, P. F., Harris, D., Dods, K., and D'Antuono, M. (2009). Faba bean (*Vicia faba* L.) seeds darken rapidly and phenolic content falls when stored at higher temperature, moisture and light intensity. Food Science and Technology, 42(10), 1703–1711.

Neme, K., Bultosa, G., and Bussa, N. (2015). Nutrient and functional properties of composite flours processed from pregelatinised barley, sprouted faba bean and carrot flours. Int. J. Food Sci. Technol. 50, 2375–2382. doi: 10.1111/ijfs.12903

Nkhoma, E.T., Poole, C., Vannappagari, V., Hall, S.A., Beutler, E. (2009). The global prevalence of glucose-6-phosphate dehydrogenase deficiency: a systematic review and meta-analysis. Blood Cells Mol Dis, 42(3):267-78. doi: 10.1016/j.bcmd.2008.12.005

Nosworthy, M. G., Medina, G., Franczyk, A. J., Neufeld, J., Appah, P., Utioh, A., Frohlich, P., and House, J. D. (2018). Effect of processing on the in vitro and in vivo protein quality of beans (Phaseolus vulgaris and *Vicia faba*). Nutrients, 10, 671. https://doi.org/10.3390/nu10060671

Novák, P., Guignard, M. S., Neumann, P., Kelly, L. J., Mlinarec, J., Koblížková, A., … Leitch, A. R. (2020). Repeat-sequence turnover shifts fundamentally in species with large genomes. Nature Plants, 6: 1325–1329. https://doi.org/10.1038/s41477-020-00785-x

Oomah, B. D., Luc, G., Leprelle, C., Drover, J. C. G., Harrison, J. E., and Olson, M. (2011). Phenolics, phytic acid, and phytase in Canadian-grown low-tannin faba bean (*Vicia faba* L.) genotypes. Journal of Agricultural and Food Chemistry, 59(8), 3763–3771.

Osman, A. M. A., Hassan, A. B., Osman, G. A., Mohammed, N., Rushdi, M. A., Diab, E. E., and Babiker, E. E. (2014). Effects of gamma irradiation and/or cooking on nutritional quality of faba bean (Viciafaba L.) cultivars seeds. Journal of Food Science and Technology, 51, 1554–1560. https://doi.org/10.1007/s13197-012-0662-7

Patterson, C. A., Curran, J., and Der, T. (2017). Effect of processing on antinutrient compounds in pulses. Cereal Chem. 94, 2–10. doi: 10.1094/CCHEM-05-16-0144-FI

Pauler FM, Koerner MV, Barlow DP (2007) Silencing by imprinted noncoding RNAs: is transcription the answer? Trends Genet 23: 284–292

Pereira, D. M., Valentão, P., Pereira, J. A., and Andrade, P. B. (2009). Phenolics: From chemistry to biology. Molecules, 14(6), 2202–2211.

Pitz, W.J., Sosulski, F.W. and C.C. Rowland. (1981). Effect of genotype and environment on vicine and convicine levels in faba beans (*Vicia faba* minor). Journal of the Science of Food and Agriculture, 32, 1-8. https://doi.org/10.1002/jsfa.2740320102

Polanowska, K., Grygier, A., Kuligowski, M., Rudzińska, M., and Nowak, J. (2020). Effect of tempe fermentation by three different strains of Rhizopus oligosporus on nutritional characteristics of faba beans. Food Science and Technology, 122, 109024.

Pulkkinen, M., Gautam, M., Lampi, A-M., Ollilainen, V., Stoddard, F., Sontag-Strohm, T., Salovaara, H. and Piironen, V. (2015). Determination of vicine and convicine from faba bean with an optimized high-performance liquid chromatographic method. Food Research International, 76(1): 168-177 https://doi.org/10.1016/j.foodres.2015.05.031

Puspitasari, W. (2017). Association analyses to genetically study reproduction and seed quality features of faba bean (*Vicia faba* L.). PhD thesis. Georg-August-University Göttingen, Germany.

Ramsay, G., van de Ven, W., Waugh, R., Griffiths, D.W. and Powel, W. (1995). Mapping Quantitative Trait Loci in Faba Beans. In: Improving Production and Utilization of Grain Legumes. [AEP (Ed.)], Copenhagen, Denmark. 2nd European conference on grain legumes. pp 444–445.

Ray, H. and F. Georges, F. (2010). A genomic approach to nutritional, pharmacological and genetic issues of faba bean (*Vicia faba*): Prospects for genetic modifications. GM Crops, 1 (2):99-106. https://doi.org/10.4161/gmcr.1.2.11891

Ray, H., Bock, C. and Georges, F. (2015). Faba bean: Transcriptome analysis from etiolated seedling and developing seed coat of key cultivars for synthesis of proanthocyanidins, phytate, raffinose family oligosaccharides, vicine, and convicine. The Plant Genome, 8:1 https://doi.org/10.3835/plantgenome2014.07.002

Ron-Parra, J., Rivera, M. M. M., López, J. J., Cordero, A. A. J. M., Larios, L. D. L. C. and Gonzalez, J. J. S. (2016). Maternal genetic inheritance of red pericarp in the grain of maize. Maydica, 61, 1–5.

Rosa-Sibakov, N., Heiniö, R. L., Cassan, D., Holopainen-Mantila, U., Micard, V., Lantto, R., and Sozer, N. (2016). Effect of bioprocessing and fractionation on the structural, textural and sensory properties of gluten-free faba bean pasta. LWT- Food Science and Technology, 67, 27–36.

Sandberg, A.S. (2002). Bioavailability of minerals in legumes. British Journal of Nutrition, 88 (S3) (2002), pp. 281-285. 10.1079/BJN/2002718

Sharan, S., Zanghelini, G., Zotzel, J., Bonerz, D., Aschoff, J., Saint-Eve, A., and Maillard, M. N. (2021). Fava bean (*Vicia faba* L.) for food applications: From seed to ingredient processing and its effect on functional properties, antinutritional factors, flavor, and color. Comprehensive Reviews in Food Science and Food Safety, 20(1), 401–428.

Sharma, A., and Sehgal, S. (1992). Effect of processing and cooking on the antinutritional factors of faba bean (*Vicia faba*). Food Chemistry, 43(5), 383–385.

Shi, D., House, J. D., Wanasundara, J. P. D., and Nickerson, M. T. (2022). Comparative evaluation of the nutritional value of faba bean flours and protein isolates with major legumes in the market. Cereal Chemistry, 99(5), 1013–1029.

Shi, D., Stone, A. K., Marinangeli, C. P. F., Carlin, J., and Nickerson, M. T. (2024). Faba bean nutrition: Macronutrients, antinutrients, and the effect of processing. Cereal Chemistry, 101, 1181–1197. https://doi.org/10.1002/cche.10830

Shi, L., Arntfield, S. D., and Nickerson, M. (2018). Changes in levels of phytic acid, lectins and oxalates during soaking and cooking of Canadian pulses. Food Research International, 107, 660–668.

Shi, L., Mu, K., Arntfield, S. D., and Nickerson, M. T. (2017). Changes in levels of enzyme inhibitors during soaking and cooking for pulses available in Canada. J. Food Sci. Technol. 54, 1014–1022. doi: 10.1007/s13197-017-2519-6

Siah, S., Wood, J. A., Agboola, S., Konczak, I., and Blanchard, C. L. (2014). Effects of soaking, boiling and autoclaving on the phenolic contents and antioxidant activities of faba beans (*Vicia faba* L.) differing in seed coat colours. Food Chemistry, 142, 461–468.

Sixdenier, G., F. Cassecuelle, L. Guillaumin, and G. Duc. (1996). Rapid Spectrophotometric Method for Reduction of Vicine and Convicine in Faba Bean Seed. FABIS Newsletter 38(39): 42–44

Soltis, D. E., Soltis, P. S., Bennett, M. D., and Leitch, I. J. (2003). Evolution of genome size in the angiosperms. American Journal of Botany, 90, 1596–1603. https://doi.org/10.3732/ajb.90.11.1596

Stone, A., Waelchli, K., Çabuk, B., McIntosh, T., Wanasundara, J., Arntfield, S., and Nickerson, M. (2021). The levels of bioactive compounds found in raw and cooked Canadian pulses. Food Science and Technology International, 27(6), 528–538.

Suzuki MM and Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 9:465–476

Tacke, R. (2023). Genetic and environmental effects on vicine and convicine levels in faba bean (*Vicia faba* L.) [Doctoral dissertation, Georg-August-Universität Göttingen]. Göttingen University Repository.

Tacke, R., W. Ecke, M. Höfer, O. Sass, and W. Link. (2022). Fine-mapping of the major locus for vicine and convicine in faba bean (*Vicia faba*) and marker-assisted breeding of a novel, low vicine and convicine winter faba bean population. Plant Breeding 141: 644–657. https://doi.org/10.1111/pbr.13039.

Tako, E., Beebe, S.E., Reed, S., Hart, J.J. and Glahn, R.P. (2014). Polyphenolic compounds appear to limit the nutritional benefit of biofortified higher iron black bean (Phaseolus vulgarisL.). Nutrition Journal, 13 (1), p. 28. 10.1186/1475-2891-13-28

Turco, I., Ferretti, G., and Bacchetti, T. (2016). Review of the health benefits of faba bean (*Vicia faba* L.) polyphenols. Journal of Food and Nutrition Research, 55(4), 283–293.

Ugwuanyi S, Makhoul M, Golicz AA, Obermeier C and Snowdon RJ (2025a) Multiple copy number variants of VC1 gene reveal single-copy expression as a key determinant of vicine content. Frontiers in Plant Science, 16:1565210. doi: 10.3389/fpls.2025.1565210

Ugwuanyi, S., Tietgen, H., Makhoul, M., Obermeier, C. and Snowdon, R. (2025b). Novel gene-copy-specific KASP markers for efficient marker-assisted development of low vicine faba bean (*Vicia faba* L.). Plant Breeding 144, no. 4: 573–579. https://doi.org/10.1111/pbr.13276.

Ugwuanyi, S., Thomas, G.-F., Tietgen, H., Dreyer, F., Abbadi, A., and Snowdon, R. (2025c). Developmental stage-dependent gene expression modulates maternal control of seed vicine biosynthesis in faba bean. Legume Science 7, no. 4: e70057. https://doi.org/10.1002/leg3.70057.

Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM (2009) A census of human transcription factors: function, expression and evolution. Nat Rev Genet 10: 252–263.

Vasudevan, S., Tong, Y. and Steitz, J.A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. Science 318: 1931–1934

Verni, M., De Mastro, G., De Cillis, F., Gobbetti, M., and Rizzello, C. G. (2019). Lactic acid bacteria fermentation to exploit the nutritional potential of Mediterranean faba bean local biotypes. Food Research International, 125, 108571.

Vidal-Valverde, C., Frias, J., Sotomayor, C., Diaz-Pollan, C., Fernandez, M., and Urbano, G. (1998). Nutrients and antinutritional factors in faba beans as affected by processing. European Food Research and Technology, 207, 140–145.

Vogelsang-O'Dwyer, M., Petersen, I. L., Joehnke, M. S., Sørensen, J. C., Bez, J., Detzel, A., Busch, M., Krueger, M., O'Mahony, J. A., Arendt, E. K., and Zannini, E. (2020). Comparison of faba bean protein ingredients produced using dry fractionation and isoelectric precipitation: Techno-functional, nutritional and environmental performance. Foods, 9(3), 322.

Wang, X., Warkentin, T. D., Briggs, C. J., Oomah, B. D., Campbell, C. G., and Woods, S. (1998). Total phenolics and condensed tannins in field pea (Pisum sativum L.) and grass pea (Lathyrus sativus L.). Euphytica, 101(1), 97–102.

Warsame, A. O., Michael, N., O'Sullivan, D. M., and Tosi, P. (2020). Identification and quantification of major faba bean seed proteins. J. Agri Food Chem. 68, 8535–8544. doi: 10.1021/acs.jafc.0c02927

Webb, A., Cottage, A., Wood, T., Khamassi, K., Hobbs, D., Gostkiewicz, K., et al. (2016). A SNP-based consensus genetic map for synteny-based trait targeting in faba bean (*Vicia faba* L.). Plant Biotechnology Journal, 14 (2016), pp. 177-185. https://doi.org/10.1111/pbi.12371

Winterbourn, C.C. (1989). Inhibition of autoxidation of divicine and isouramil by the combination of superoxide dismutase and reduced glutathione. Archives of Biochemistry and Biophysics, 271 (1989), pp. 447-455

Woodwark, C and Bateman, A (2011). The Characterisation of Three Types of Genes that Overlie Copy Number Variable Regions. PLoSONE, 6(5): e14814. doi:10.1371/journal.pone.0014814

Woyengo, T. A., and Nyachoti, C. M. (2012). IIeal digestibility of amino acids for zero-tannin faba bean (*Vicia faba* L.) fed to broiler chicks. Poult. Sci. 91, 439–443. doi: 10.3382/ps.2011-01678

Zhu, X.G., Long, S.P. and Ort, D.R. (2010). Improving photosynthetic efficiency for greater yield. Annu. Rev. Plant Biol., 61: 235-261

Appendix

Appendix I: Supplementary material from

Ugwuanyi, S., Makhoul, M., Golicz, A.A., Obermeier, C. and Snowdon, R.J. (2025). Multiple copy number variants of VC1 gene reveal single-copy expression as a key determinant of vicine content. Front. Plant Sci. 16:1565210. doi: 10.3389/fpls.2025.1565210

Multiple copy number variants of *VC1* gene reveal single-copy expression as a key determinant of vicine content

Samson Ugwuanyi*[1], Manar Makhoul[1], Agnieszka A. Golicz[1], Christian Obermeier[1] and Rod J. Snowdon[1]

[1]Department of Plant Breeding, Justus Liebig University, Giessen, Germany.
*Correspondence:

Samson Ugwuanyi

samson.ugwuanyi@agrar.uni-giessen.de

Supplementary Material

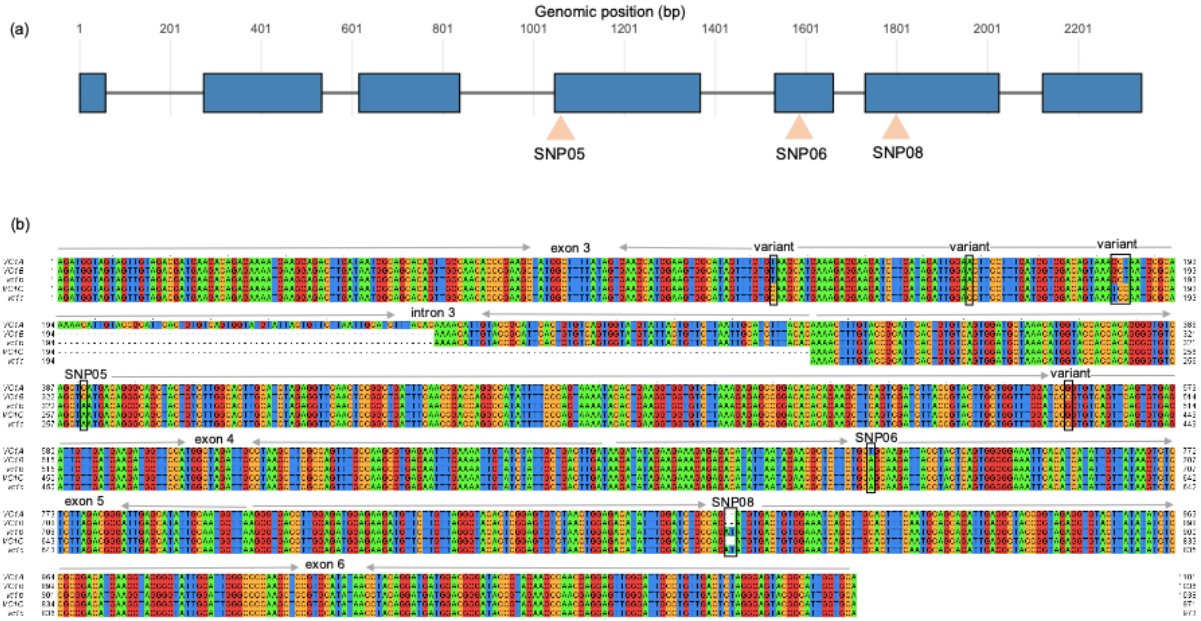


Figure S1: Structural variations in *VC1* genes. (a) Schematic representation of the *VC1* gene from the Hedin reference genome, showing the positions of KASP primers. (b) Alignment of reference *VC1* genes from Hedin and Tiffany showing polymorphisms in the coding regions that can distinguish different gene variants.

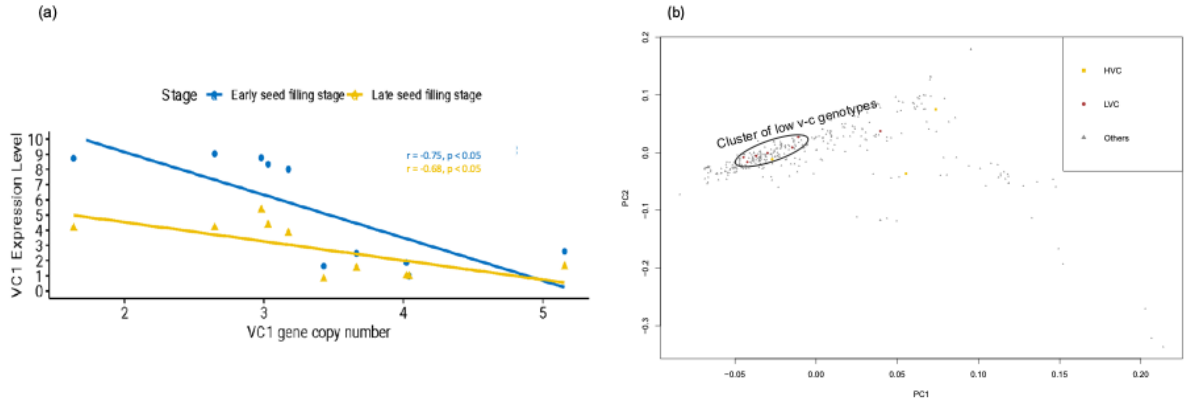


Figure S2: **Tracing the source of low vicine and convicine (v-c) trait**. (a) Correlation analysis showing the relationship between *VC1* copy number and *VC1* expression (means ± SD, n = 3). (b) Principal component analysis showed that most low v-c genotypes cluster together as they share a single genetic origin, as shown in dark circle.

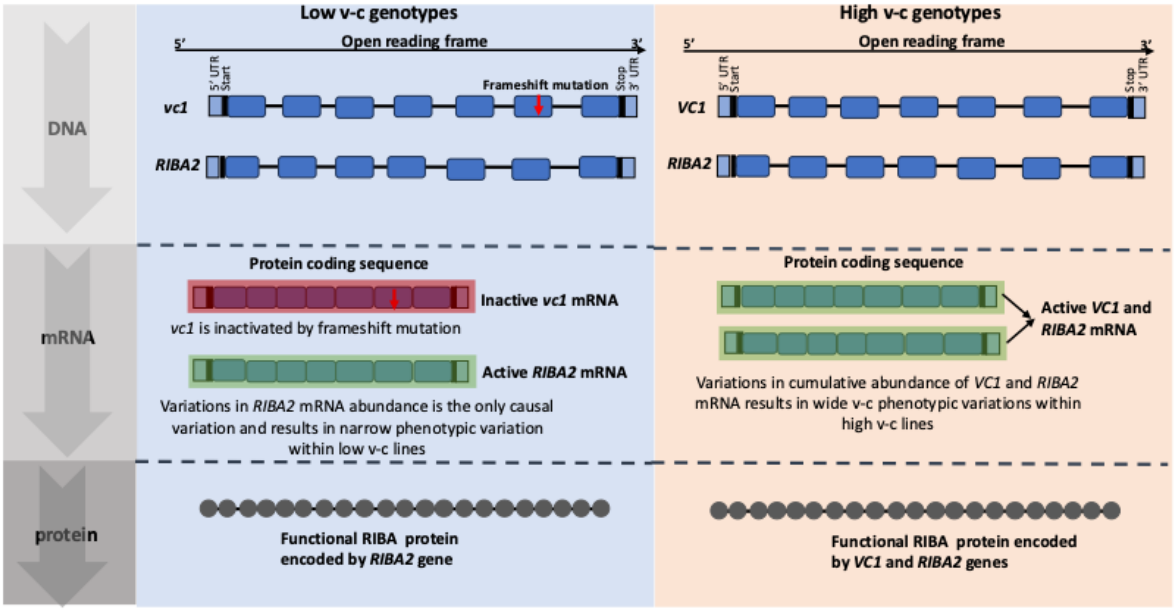


Figure S3: Schematic representation of genetic variations in vicine and convicine (v-c) contents involving *VC1* and *RIBA2* genes.

Table S1 List of faba bean lines and corresponding experiments

Genotype	VC status	Experiments
NPZ-FB-92	lvc	Gene validations, CNV analysis, Gene expression analysis, cDNA sequencing, KASP assay
NPZ-FB-73	lvc	Gene validations, CNV analysis, Gene expression analysis, cDNA sequencing, KASP assay
NPZ-FB-56	lvc	Gene validations, CNV analysis, KASP assay
NPZ-FB-117	HVC	Gene validations, CNV analysis, Gene expression analysis, cDNA sequencing, KASP assay
NPZ-FB-114	HVC	Gene validations, CNV analysis, KASP assay
NPZ-FB-110	HVC	Gene validations, CNV analysis, KASP assay
NPZ-FB-115	lvc	Gene validations, CNV analysis, KASP assay
NPZ-FB-113	HVC	Gene validations, CNV analysis, Gene expression analysis, cDNA sequencing, KASP assay
NPZ-FB-52	lvc	Gene validations, CNV analysis, KASP assay
NPZ-FB-55	lvc	Gene validations, CNV analysis, Gene expression analysis, cDNA sequencing, KASP assay
NPZ-FB-206	lvc	Gene validations, CNV analysis, KASP assay
NPZ-FB-64	lvc	Gene validations, CNV analysis, Gene expression analysis, cDNA sequencing, KASP assay
NPZ-FB-138	HVC	Gene validations, CNV analysis, KASP assay
NPZ-FB-87	HVC	Gene validations, CNV analysis, Gene expression analysis, cDNA sequencing, KASP assay
NPZ-FB-8	lvc	Gene validations, CNV analysis, Gene expression analysis, KASP assay
NPZ-FB-83	HVC	Gene validations, CNV analysis, Gene expression analysis, cDNA sequencing, KASP assay
NPZ-FB-88	HVC	Gene validations, CNV analysis, KASP assay
NPZ-FB-143	HVC	Gene validations, CNV analysis, Gene expression analysis, cDNA sequencing, KASP assay

VC = vicine and convicine
HVC = high vicine and convicine
lvc = low vicine and convicine

Table S2 List of primers for PCR and qPCR

Primer	Target	Sequence	Expected amplicon size (bp)
Forward	VC1A/VC1B	AAAGCTAATGCGCAAAAACA	230/165
Reverse	VC1A/VC1B	TCAGTCAAACTAACTCCAGATGTGT	
Forward	VC1C	GATGGTGGACAGTAAATCCAATG	1504
Reverse	VC1C	TTGACACGTTTTGTCTCCAAGTA	
Forward	RIBA2	AGAGGAAGACCTTGATAGATTGGAACT	1700
Reverse	RIBA2	CGTAGATATACAAGGACACCTCTTCC	
Forward	VC1 copy number	AGGGGTATTGGATTGGGCC	119
Reverse	VC1 copy number	CCAATGCCGTACTCCCTAGA	
Forward	VC1 expression	TCCCAGTAAAATACACTGAAGGTG	127
Reverse	VC1 expression	CCATCTTCATCAACAATCTCACA	
Forward	RIBA2 expression	ACAGGGGTGTCAGCTCATGATAG	120
Reverse	RIBA2 expression	ACCACCATCTCTGTATTTTAGTGG	
Forward	VC1 cDNA sequencing	AAGCATGAAAGAGGAAGATCTTGATAG	714
Reverse	VC1 cDNA sequencing	CCGCGGAGATATATAAGTACACC	
Forward	RIBA2 cDNA sequencing	AGAGGAAGACCTTGATAGATTGGAACT	702
Reverse	RIBA2 cDNA sequencing	CGTAGATATACAAGGACACCTCTTCC	

Table S3 List of KASP assay primers

Marker ID	Primer ID	Primer sequences
SNP05	SNP05_Hex	GAAGGTCGGAGTCAACGGATTCACCACAGGGGTGTCAGCTA
	SNP05_Fam	GAAGGTGACCAAGTTCATGCTCACCACAGGGGTGTCAGCTC
	SNP05_Com	ACCTCTAGATGCAAGTGCCAAG
SNP06	SNP06_Hex	GAAGGTCGGAGTCAACGGATTCATATTAATAGAACGCTCTTCTGCA
	SNP06_Fam	GAAGGTGACCAAGTTCATGCTCATATTAATAGAACGCTCTTCTGCT
	SNP06_Com	ATATGATGTGAATTTCCCCCACT
SNP08	SNP08_Hex	GAAGGTCGGAGTCAACGGATTGACATATTTGGATCTGCCAGATG
	SNP08_Fam	GAAGGTGACCAAGTTCATGCTGAGACATATTTGGATCTGCCAGATA
	SNP08_Com	TAGCCTCAATCTGCTGCATTG

Table S4 Vicia faba RIBA genes. Summary of RIBA genes in faba bean genome, including their names, chromosomal positions, nucleotide sequences, and predicted protein sequences

Sequences can be found online at:
https://www.frontiersin.org/articles/10.3389/fpls.2025.1565210/full#supplementary-material

Table S5 A diverse set of 97 faba bean genotypes used to validate a 2bp insertion in vc1 (SNP08) and the co-segregation of RIBA2 SNP with v-c phenotypes

Genotype name	SNP08	RIBA2 SNP	vicine content (percent per seed dry weight)
NPZ-FB-10.2	GG	GG	0.46
NPZ-FB-18.3	GG	GG	0.204
NPZ-FB-19.5	GA	CC	0.104
NPZ-FB-22.4	GG	GG	0.415
NPZ-FB-23.5	GG	GG	0.375
NPZ-FB-24.1	GG	GG	0.351
NPZ-FB-25.5	GG	GG	0.394
NPZ-FB-26.1	GG	GG	0.415
NPZ-FB-27.2	GG	GG	0.344
NPZ-FB-28.5	GG	GG	0.369
NPZ-FB-29.1	GG	GG	0.53
NPZ-FB-30.2	GG	GG	0.523
NPZ-FB-31.3	GA	CC	0.115
NPZ-FB-32.1	GG	GG	0.427
NPZ-FB-33.3	GG	GG	0.473
NPZ-FB-34.4	GG	GG	0.377
NPZ-FB-35.4	GG	GG	0.39
NPZ-FB-36.1	GG	GG	0.53
NPZ-FB-37.1	GG	GG	0.476
NPZ-FB-38.4	GG	GG	0.437
NPZ-FB-39.2	GG	GG	0.415
NPZ-FB-40.3	GG	GG	0.459
NPZ-FB-41.5	GG	GG	0.416
NPZ-FB-42.3	GA	CC	0.118
NPZ-FB-43.5	GA	CC	0.098
NPZ-FB-44.5	GA	CC	0.093
NPZ-FB-45.2	GA	CC	0.106
NPZ-FB-46.4	GA	CC	0.107
NPZ-FB-47.3	GA	CC	0.134
NPZ-FB-48.5	GA	CC	0.069
NPZ-FB-49.3	GA	CC	0.059
NPZ-FB-50.4	GA	CC	0.094
NPZ-FB-51.2	GG	GG	0.366
NPZ-FB-52.4	GA	CC	0.148
NPZ-FB-53.3	GA	CC	0.171
NPZ-FB-54.1	GG	GG	0.514
NPZ-FB-55.2	GA	CC	0.092
NPZ-FB-56.3	GA	CC	0.11

Genotype name	SNP08	RIBA2 SNP	vicine content (percent per seed dry weight)
NPZ-FB-57.5	GA	CC	0.155
NPZ-FB-58.1	GA	CC	0.103
NPZ-FB-59.4	GA	CC	0.104
NPZ-FB-60.4	GA	CC	0.093
NPZ-FB-61.2	GA	CC	0.107
NPZ-FB-62.1	GA	CC	0.118
NPZ-FB-63.2	GA	CC	0.136
NPZ-FB-64.2	GA	CC	0.137
NPZ-FB-65.3	GA	CC	0.149
NPZ-FB-66.4	GA	CC	0.446
NPZ-FB-67.2	GG	CC	0.332
NPZ-FB-68.2	GG	GG	0.463
NPZ-FB-69.2	GG	GG	0.385
NPZ-FB-70.4	GG	GG	0.33
NPZ-FB-71.2	GG	GG	0.467
NPZ-FB-72.4	GG	GG	0.378
NPZ-FB-73.2	GG	GG	0.24
NPZ-FB-74.4	GA	CC	0.108
NPZ-FB-75.1	GG	GG	0.408
NPZ-FB-76.4	GG	GG	0.509
NPZ-FB-77.1	GG	GG	0.46
NPZ-FB-78.3	GG	GG	0.377
NPZ-FB-79.3	GG	GG	0.429
NPZ-FB-80.1	GG	GG	0.393
NPZ-FB-82.2	GG	GG	0.494
NPZ-FB-83.3	GA	CC	0.33
NPZ-FB-84.4	GG	GG	0.323
NPZ-FB-85.3	GA	CC	0.422
NPZ-FB-86.3	GG	GG	0.469
NPZ-FB-87.2	GG	GG	0.338
NPZ-FB-88.4	GA	CC	0.437
NPZ-FB-89.1	GG	GG	0.392
NPZ-FB-90.1	GG	GG	0.403
NPZ-FB-91.5	GG	GG	0.492
NPZ-FB-92.3	GA	CC	0.098
NPZ-FB-93.1	GG	GG	0.5
NPZ-FB-94.3	GG	GG	0.375
NPZ-FB-95.4	GG	GG	0.445
NPZ-FB-96.3	GA	CC	0.094
NPZ-FB-97.3	GA	CC	0.094
NPZ-FB-98.4	GA	CC	0.09
NPZ-FB-99.1	GA	CC	0.124
NPZ-FB-100.1	GA	CC	0.124
NPZ-FB-101.1	GA	CC	0.093
NPZ-FB-102.1	GA	CC	0.091
NPZ-FB-103.1	GA	CC	0.107

Genotype name	SNP08	RIBA2 SNP	vicine content (percent per seed dry weight)
NPZ-FB-104.3	GA	CC	0.093
NPZ-FB-105.5	GA	CC	0.1
NPZ-FB-106.3	GA	CC	0.076
NPZ-FB-107.5	GA	CC	0.131
NPZ-FB-108.2	GA	CC	0.101
NPZ-FB-109.3	GA	CC	0.121
NPZ-FB-110.1	GA	GC	0.532
NPZ-FB-111.3	GA	CC	0.082
NPZ-FB-112.1	GA	CC	0.175
NPZ-FB-114.2	GA	CC	0.193
NPZ-FB-115.2	GA	CC	0.088
NPZ-FB-116.3	GA	CC	0.279
NPZ-FB-117.4	GG	GG	0.417

For SNP08, allele G represents wild type *VC1*, while allele A represent mutant *vc1*

Table S6 cDNA sequences of *VC1* and *RIBA2*

Sequences can be found online at:
https://www.frontiersin.org/articles/10.3389/fpls.2025.1565210/full#supplementary-material

Appendix II: Supplementary material from

Ugwuanyi, S., Tietgen, H., Makhoul, M., Obermeier, C. and Snowdon, R. (2025). Novel gene-copy-specific KASP markers for efficient marker-assisted development of low vicine faba bean (*Vicia faba* L.). Plant Breeding, 144(4): 573–579. https://doi.org/10.1111/pbr.13276

Novel gene-copy specific KASP markers for efficient marker-assisted development of low vicine faba bean (*Vicia faba* L.)

Samson Ugwuanyi*[1,2], Hanna Tietgen[3], Manar Makhoul[1], Christian Obermeier[1] and Rod J. Snowdon[1]

[1]Department of Plant Breeding, Justus Liebig University, Giessen, Germany.

[2]Department of Plant Science and Biotechnology, University of Nigeria, Nsukka, Nigeria.

[3]NPZ Innovation GmbH, Holtsee, Germany.

*Correspondence:

Samson Ugwuanyi

samson.ugwuanyi@agrar.uni-giessen.de

Supplementary Material

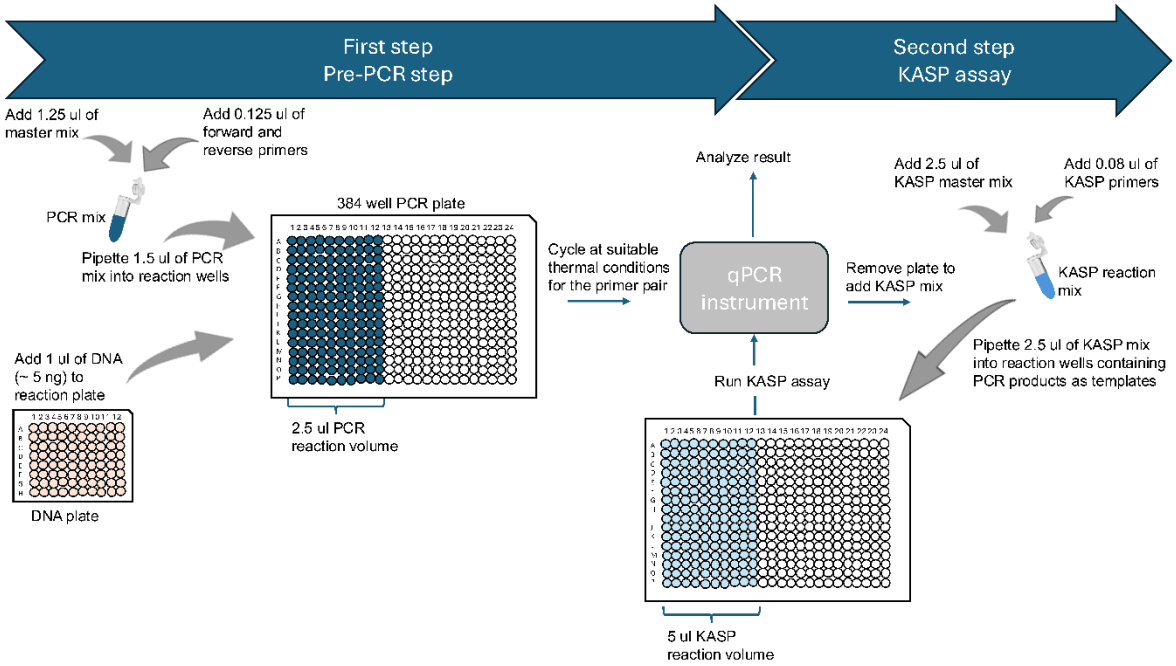


Figure S1: Workflow for the target-specific preamplification KASP assay. In the first step, the reaction mix for PCR was prepared. A pair of PCR primers was specifically designed to amplify only the target variant. The PCR reaction took place directly in a KASP assay plate in a final volume of 2.5 µL. The reaction was cycled for up to 30 cycles under appropriate thermal conditions. Following the PCR reaction, the KASP mix was prepared in the next step and added to the same reaction plate. Subsequently, KASP genotyping assay was performed and analysed.

Table S1 A subset of diverse faba bean lines for initial co-segregation analysis

Genotype	v-c status	vicine content
NPZ-FB-78	hvc	0.377
NPZ-FB-92	lvc	0.098
NPZ-FB-73	hvc	0.24
NPZ-FB-56	lvc	0.11
NPZ-FB-117	hvc	0.417
NPZ-FB-114	lvc	0.193
NPZ-FB-110	hvc	0.532
NPZ-FB-115	lvc	0.088
NPZ-FB-52	lvc	0.148
NPZ-FB-55	lvc	0.092
NPZ-FB-64	lvc	0.137
NPZ-FB-87	hvc	0.338
NPZ-FB-83	hvc	0.33
NPZ-FB-88	hvc	0.437

Table S2 Faba bean genotypes with corresponding vicine contents and genotype calls for all markers, including vicine status prediction by the markers

Genotype	SNP_b1	SNP_b2	SNP_pr1	SNP_pr2	SNP08	Haplotype group	Marker prediction	vicine content
NPZ-FB-10	AA	GG	AA	AA	GG	Hap1	hvc	0.46
NPZ-FB-18	AA	GG	AA	AA	GG	Hap1	hvc	0.204
NPZ-FB-19	GG	CC	TT	TT	AA	Hap2	lvc	0.104
NPZ-FB-22	AA	GG	AA	AA	GG	Hap1	hvc	0.415
NPZ-FB-23	AA	GG	AA	AA	GG	Hap1	hvc	0.375
NPZ-FB-24	AA	GG	AA	AA	GG	Hap1	hvc	0.351
NPZ-FB-25	AA	GG	AA	AA	GG	Hap1	hvc	0.394
NPZ-FB-26	AA	GG	AA	AA	GG	Hap1	hvc	0.415
NPZ-FB-27	AA	GG	AA	AA	GG	Hap1	hvc	0.344
NPZ-FB-28	AA	GG	AA	AA	GG	Hap1	hvc	0.369
NPZ-FB-29	AA	GG	AA	AA	GG	Hap1	hvc	0.53
NPZ-FB-30	AA	GG	AA	AA	GG	Hap1	hvc	0.523
NPZ-FB-31	GG	CC	TT	TT	AA	Hap2	lvc	0.115
NPZ-FB-32	AA	GG	AA	AA	GG	Hap1	hvc	0.427
NPZ-FB-33	AA	GG	AA	AA	GG	Hap1	hvc	0.473
NPZ-FB-34	AA	GG	AA	AA	GG	Hap1	hvc	0.377
NPZ-FB-35	AA	GG	AA	AA	GG	Hap1	hvc	0.39
NPZ-FB-36	AA	GG	AA	AA	GG	Hap1	hvc	0.53
NPZ-FB-37	AA	GG	AA	AA	GG	Hap1	hvc	0.476
NPZ-FB-38	AA	GG	AA	AA	GG	Hap1	hvc	0.437
NPZ-FB-39	AA	GG	AA	AA	GG	Hap1	hvc	0.415
NPZ-FB-40	AA	GG	AA	AA	GG	Hap1	hvc	0.459
NPZ-FB-41	AA	GG	AA	AA	GG	Hap1	hvc	0.416
NPZ-FB-42	GG	CC	TT	TT	AA	Hap2	lvc	0.118
NPZ-FB-43	GG	CC	TT	TT	AA	Hap2	lvc	0.098
NPZ-FB-44	GG	CC	TT	TT	AA	Hap2	lvc	0.093
NPZ-FB-45	GG	CC	TT	TT	AA	Hap2	lvc	0.106
NPZ-FB-46	GG	CC	TT	TT	AA	Hap2	lvc	0.107
NPZ-FB-47	GG	CC	TT	TT	AA	Hap2	lvc	0.134
NPZ-FB-48	GG	CC	TT	TT	AA	Hap2	lvc	0.069
NPZ-FB-49	GG	CC	TT	TT	AA	Hap2	lvc	0.059
NPZ-FB-50	GG	CC	TT	TT	AA	Hap2	lvc	0.094
NPZ-FB-51	AA	GG	AA	AA	GG	Hap1	hvc	0.366
NPZ-FB-52	GG	CC	TT	TT	AA	Hap2	lvc	0.148
NPZ-FB-53	GG	CC	TT	TT	AA	Hap2	lvc	0.171
NPZ-FB-54	AA	GG	AA	AA	GG	Hap1	hvc	0.514
NPZ-FB-55	GG	CC	TT	TT	AA	Hap2	lvc	0.092
NPZ-FB-56	GG	CC	TT	TT	AA	Hap2	lvc	0.11
NPZ-FB-57	GG	CC	TT	TT	AA	Hap2	lvc	0.155
NPZ-FB-58	GG	CC	TT	TT	AA	Hap2	lvc	0.103
NPZ-FB-59	GG	CC	TT	TT	AA	Hap2	lvc	0.104

Genotype	SNP_b1	SNP_b2	SNP _pr1	SNP _pr2	SNP08	Haplotype group	Marker prediction	vicine content
NPZ-FB-60	GG	CC	TT	TT	AA	Hap2	lvc	0.093
NPZ-FB-61	GG	CC	TT	TT	AA	Hap2	lvc	0.107
NPZ-FB-62	GG	CC	TT	TT	AA	Hap2	lvc	0.118
NPZ-FB-63	GG	CC	TT	TT	AA	Hap2	lvc	0.136
NPZ-FB-64	GG	CC	TT	TT	AA	Hap2	lvc	0.137
NPZ-FB-65	GG	CC	TT	TT	AA	Hap2	lvc	0.149
NPZ-FB-66	GG	CC	TT	TT	GG	Hap3	hvc	0.446
NPZ-FB-67	GG	CC	TT	TT	GG	Hap3	hvc	0.332
NPZ-FB-68	AA	GG	AA	AA	GG	Hap1	hvc	0.463
NPZ-FB-69	AA	GG	AA	AA	GG	Hap1	hvc	0.385
NPZ-FB-70	AA	GG	AA	AA	GG	Hap1	hvc	0.33
NPZ-FB-71	AA	GG	AA	AA	GG	Hap1	hvc	0.467
NPZ-FB-72	AA	GG	AA	AA	GG	Hap1	hvc	0.378
NPZ-FB-73	AA	GG	AA	AA	GG	Hap1	hvc	0.24
NPZ-FB-74	GG	CC	TT	TT	AA	Hap2	lvc	0.108
NPZ-FB-75	AA	GG	AA	AA	GG	Hap1	hvc	0.408
NPZ-FB-76	AA	GG	AA	AA	GG	Hap1	hvc	0.509
NPZ-FB-77	AA	GG	AA	AA	GG	Hap1	hvc	0.46
NPZ-FB-78	AA	GG	AA	AA	GG	Hap1	hvc	0.377
NPZ-FB-79	AA	GG	AA	AA	GG	Hap1	hvc	0.429
NPZ-FB-80	AA	GG	AA	AA	GG	Hap1	hvc	0.393
NPZ-FB-82	AA	GG	AA	AA	GG	Hap1	hvc	0.494
NPZ-FB-83	GG	CC	TT	TT	GG	Hap3	hvc	0.33
NPZ-FB-84	AA	GG	AA	AA	GG	Hap1	hvc	0.323
NPZ-FB-85	GG	CC	TT	TT	GG	Hap3	hvc	0.422
NPZ-FB-86	AA	GG	AA	AA	GG	Hap1	hvc	0.469
NPZ-FB-87	AA	GG	AA	AA	GG	Hap1	hvc	0.338
NPZ-FB-88	GG	CC	TT	TT	GG	Hap3	hvc	0.437
NPZ-FB-89	AA	GG	AA	AA	GG	Hap1	hvc	0.392
NPZ-FB-90	AA	GG	AA	AA	GG	Hap1	hvc	0.403
NPZ-FB-91	AA	GG	AA	AA	GG	Hap1	hvc	0.492
NPZ-FB-92	GG	CC	TT	TT	AA	Hap2	lvc	0.098
NPZ-FB-93	AA	GG	AA	AA	GG	Hap1	hvc	0.5
NPZ-FB-94	AA	GG	AA	AA	GG	Hap1	hvc	0.375
NPZ-FB-95	AA	GG	AA	AA	GG	Hap1	hvc	0.445
NPZ-FB-96	GG	CC	TT	TT	AA	Hap2	lvc	0.094
NPZ-FB-97	GG	CC	TT	TT	AA	Hap2	lvc	0.094
NPZ-FB-98	GG	CC	TT	TT	AA	Hap2	lvc	0.09
NPZ-FB-99	GG	CC	TT	TT	AA	Hap2	lvc	0.124
NPZ-FB-100	GG	CC	TT	TT	AA	Hap2	lvc	0.124
NPZ-FB-101	GG	CC	TT	TT	AA	Hap2	lvc	0.093
NPZ-FB-102	GG	CC	TT	TT	AA	Hap2	lvc	0.091
NPZ-FB-103	GG	CC	TT	TT	AA	Hap2	lvc	0.107

Genotype	SNP_b1	SNP_b2	SNP _pr1	SNP _pr2	SNP08	Haplotype group	Marker prediction	vicine content
NPZ-FB-104	GG	CC	TT	TT	AA	Hap2	lvc	0.093
NPZ-FB-105	GG	CC	TT	TT	AA	Hap2	lvc	0.1
NPZ-FB-106	GG	CC	TT	TT	AA	Hap2	lvc	0.076
NPZ-FB-107	GG	CC	TT	TT	AA	Hap2	lvc	0.131
NPZ-FB-108	GG	CC	TT	TT	AA	Hap2	lvc	0.101
NPZ-FB-109	GG	CC	TT	TT	AA	Hap2	lvc	0.121
NPZ-FB-110	GA	GC	TA	TA	GA	het	hvc	0.532
NPZ-FB-111	GG	CC	TT	TT	AA	Hap2	lvc	0.082
NPZ-FB-112	GG	CC	TT	TT	AA	Hap2	lvc	0.175
NPZ-FB-114	GG	CC	TT	TT	AA	Hap2	lvc	0.193
NPZ-FB-115	GG	CC	TT	TT	AA	Hap2	lvc	0.088
NPZ-FB-116	GG	CC	TT	TT	AA	Hap2	lvc	0.279
NPZ-FB-117	AA	GG	AA	AA	GG	Hap1	hvc	0.417

hvc = high vicine-convicine; lvc = low vicine-convicine

Appendix III: Supplementary material from

Ugwuanyi, S., Thomas, G.-F., Tietgen, H., Dreyer, F., Abbadi, A. and Snowdon, R. (2025). Developmental stage-dependent gene expression modulates maternal control of seed vicine biosynthesis in faba bean. *Legume Science,* 7(4): e70057. https://doi.org/10.1002/leg3.70057

Developmental stage-dependent gene expression modulates maternal control of seed vicine biosynthesis in faba bean

Samson Ugwuanyi*[1,2], Gerke-Fabian Thomas[1], Hanna Tietgen[3], Felix Dreyer[3], Amine Abbadi[3] and Rod J. Snowdon[1]

[1]Department of Plant Breeding, Justus Liebig University, Giessen, Germany.
[2]Department of Plant Science and Biotechnology, University of Nigeria, Nsukka, Nigeria.
[3]NPZ Innovation GmbH, Holtsee, Germany.

*Correspondence:

Samson Ugwuanyi

samson.ugwuanyi@agrar.uni-giessen.de

Supplementary Material

Table S1: cDNA sequences of *VC1* and *RIBA2* genes from selected parental lines

Sequences can be found online at:
https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Fleg3.70057&file=leg370057-sup-0001-Table_S1_S2.xlsx

Table S2 Genotypic data for F2 population based on the KASP assay analyses of SNP_b1 and SNP08

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0055xFB-0070	FB-0055xFB-0070-7.1	GA	GA	Het
FB-0055xFB-0070	FB-0055xFB-0070-7.2	GA	GA	Het
FB-0055xFB-0070	FB-0055xFB-0070-7.3	GA	GA	Het
FB-0055xFB-0070	FB-0055xFB-0070-7.4	GG	AA	Hap2
FB-0055xFB-0070	FB-0055xFB-0070-7.5	GA	GA	Het
FB-0055xFB-0083	FB-0055xFB-0083-40.1	GA	GA	Het
FB-0055xFB-0083	FB-0055xFB-0083-40.3	AA	GG	Hap1
FB-0055xFB-0083	FB-0055xFB-0083-40.4	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-36.1	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-36.2	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-36.3	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-36.4	GG	AA	Hap2
FB-0117xFB-0055	FB-0117xFB-0055-36.5	GA	GA	Het
FB-0055xFB-0117	FB-0055xFB-0117-15.1	AA	GG	Hap1
FB-0055xFB-0117	FB-0055xFB-0117-15.2	AA	GG	Hap1
FB-0055xFB-0117	FB-0055xFB-0117-15.3	GA	GA	Het
FB-0055xFB-0117	FB-0055xFB-0117-15.4	GG	AA	Hap2
FB-0055xFB-0117	FB-0055xFB-0117-15.5	AA	GG	Hap1
FB-0117xFB-0055	FB-0117xFB-0055-35.1	AA	GG	Hap1
FB-0117xFB-0055	FB-0117xFB-0055-35.2	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-35.3	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-35.4	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-35.5	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-6.1	AA	GG	Hap1

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0092xFB-0113	FB-0092xFB-0113-6.2	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-6.3	AA	GG	Hap1
FB-0092xFB-0113	FB-0092xFB-0113-6.4	AA	GG	Hap1
FB-0092xFB-0113	FB-0092xFB-0113-6.5	GA	GA	Het
FB-0055xFB-0113	FB-0055xFB-0113-47.1	GA	GA	Het
FB-0055xFB-0113	FB-0055xFB-0113-47.2	GA	GA	Het
FB-0055xFB-0113	FB-0055xFB-0113-47.3	AA	GG	Hap1
FB-0055xFB-0113	FB-0055xFB-0113-47.4	GA	GA	Het
FB-0055xFB-0113	FB-0055xFB-0113-47.5	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-73.1	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-73.2	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-73.3	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-73.4	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-73.5	GA	GA	Het
FB-0116xFB-0055	FB-0116xFB-0055-63.1	AA	GG	Hap1
FB-0116xFB-0055	FB-0116xFB-0055-63.2	GA	GA	Het
FB-0116xFB-0055	FB-0116xFB-0055-63.3	AA	GG	Hap1
FB-0116xFB-0055	FB-0116xFB-0055-63.4	GA	GA	Het
FB-0116xFB-0055	FB-0116xFB-0055-63.5	GA	GA	Het
FB-0083xFB-0055	FB-0083xFB-0055-12.1	GG	AA	Hap2
FB-0083xFB-0055	FB-0083xFB-0055-12.2	GG	GG	Hap3
FB-0083xFB-0055	FB-0083xFB-0055-12.3	GG	GA	Het
FB-0083xFB-0055	FB-0083xFB-0055-12.4	GG	AA	Hap2
FB-0083xFB-0055	FB-0083xFB-0055-12.5	GG	AA	Hap2
FB-0116xFB-0055	FB-0116xFB-0055-62.1	GG	AA	Hap2
FB-0116xFB-0055	FB-0116xFB-0055-62.2	GA	GA	Het
FB-0116xFB-0055	FB-0116xFB-0055-62.3	GG	AA	Hap2
FB-0116xFB-0055	FB-0116xFB-0055-62.4	GA	GA	Het
FB-0116xFB-0055	FB-0116xFB-0055-62.5	GG	AA	Hap2
FB-0092xFB-0117	FB-0092xFB-0117-25.1	AA	GG	Hap1

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0092xFB-0117	FB-0092xFB-0117-25.2	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-25.3	AA	GG	Hap1
FB-0092xFB-0117	FB-0092xFB-0117-25.4	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-25.5	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-5.1	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-5.2	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-5.3	GG	AA	Hap2
FB-0092xFB-0113	FB-0092xFB-0113-5.4	AA	GG	Hap1
FB-0092xFB-0113	FB-0092xFB-0113-5.5	GG	AA	Hap2
FB-0070xFB-0055	FB-0070xFB-0055-2.1	GG	AA	Hap2
FB-0070xFB-0055	FB-0070xFB-0055-2.2	GA	GA	Het
FB-0070xFB-0055	FB-0070xFB-0055-2.3	AA	GG	Hap1
FB-0070xFB-0055	FB-0070xFB-0055-2.4	AA	GG	Hap1
FB-0070xFB-0055	FB-0070xFB-0055-2.5	GA	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-57.1	GA	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-57.2	GA	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-57.3	GG	AA	Hap2
FB-0083xFB-0092	FB-0083xFB-0092-57.4	AA	GG	Hap1
FB-0083xFB-0092	FB-0083xFB-0092-57.5	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-48.1	AA	GG	Hap1
FB-0092xFB-0113	FB-0092xFB-0113-48.2	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-48.3	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-48.4	GG	AA	Hap2
FB-0092xFB-0113	FB-0092xFB-0113-48.5	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-24.1	GG	AA	Hap2
FB-0092xFB-0117	FB-0092xFB-0117-24.2	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-24.3	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-24.4	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-24.5	GG	AA	Hap2
FB-0055xFB-0117	FB-0055xFB-0117-17.1	GA	GA	Het

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0055xFB-0117	FB-0055xFB-0117-17.2	GA	GA	Het
FB-0055xFB-0117	FB-0055xFB-0117-17.3	AA	GG	Hap1
FB-0055xFB-0117	FB-0055xFB-0117-17.4	GG	AA	Hap2
FB-0055xFB-0117	FB-0055xFB-0117-17.5	AA	GG	Hap1
FB-0070xFB-0055	FB-0070xFB-0055-3.1	GA	GA	Het
FB-0070xFB-0055	FB-0070xFB-0055-3.2	GA	GA	Het
FB-0070xFB-0055	FB-0070xFB-0055-3.3	GG	AA	Hap2
FB-0070xFB-0055	FB-0070xFB-0055-3.4	AA	GG	Hap1
FB-0070xFB-0055	FB-0070xFB-0055-3.5	GA	GA	Het
FB-0055xFB-0070	FB-0055xFB-0070-8.1	GA	GA	Het
FB-0055xFB-0070	FB-0055xFB-0070-8.2	AA	GG	Hap1
FB-0055xFB-0070	FB-0055xFB-0070-8.3	AA	GG	Hap1
FB-0055xFB-0070	FB-0055xFB-0070-8.4	AA	GG	Hap1
FB-0055xFB-0070	FB-0055xFB-0070-8.5	GA	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-56.1	GG	AA	Hap2
FB-0083xFB-0092	FB-0083xFB-0092-56.2	GA	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-56.3	GA	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-56.4	AA	GG	Hap1
FB-0083xFB-0092	FB-0083xFB-0092-56.5	GA	GA	Het
FB-0116xFB-0092	FB-0116xFB-0092-51.1	GG	AA	Hap2
FB-0116xFB-0092	FB-0116xFB-0092-51.2	GA	GA	Het
FB-0116xFB-0092	FB-0116xFB-0092-51.3	GA	GA	Het
FB-0116xFB-0092	FB-0116xFB-0092-51.4	AA	GG	Hap1
FB-0116xFB-0092	FB-0116xFB-0092-51.5	GA	GA	Het
FB-0113xFB-0055	FB-0113xFB-0055-67.1	AA	GG	Hap1
FB-0113xFB-0055	FB-0113xFB-0055-67.2	GG	AA	Hap2
FB-0113xFB-0055	FB-0113xFB-0055-67.3	GA	GA	Het
FB-0113xFB-0055	FB-0113xFB-0055-67.4	GG	AA	Hap2
FB-0113xFB-0055	FB-0113xFB-0055-67.5	GA	GA	Het
FB-0117xFB-0092	FB-0117xFB-0092-10.1	GA	GA	Het

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0117xFB-0092	FB-0117xFB-0092-10.2	GA	GA	Het
FB-0117xFB-0092	FB-0117xFB-0092-10.3	AA	GG	Hap1
FB-0117xFB-0092	FB-0117xFB-0092-10.4	GA	GA	Het
FB-0117xFB-0092	FB-0117xFB-0092-10.5	GG	AA	Hap2
FB-0113xFB-0092	FB-0113xFB-0092-70.1	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-70.2	AA	GG	Hap1
FB-0113xFB-0092	FB-0113xFB-0092-70.3	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-70.4	AA	GG	Hap1
FB-0113xFB-0092	FB-0113xFB-0092-70.5	GA	GA	Het
FB-0092xFB-0116	FB-0092xFB-0116-18.1	AA	GG	Hap1
FB-0092xFB-0116	FB-0092xFB-0116-18.2	GA	GA	Het
FB-0092xFB-0116	FB-0092xFB-0116-18.3	GG	AA	Hap2
FB-0092xFB-0116	FB-0092xFB-0116-18.4	GG	AA	Hap2
FB-0092xFB-0116	FB-0092xFB-0116-18.5	GG	AA	Hap2
FB-0092xFB-0116	FB-0092xFB-0116-19.1	GA	GA	Het
FB-0092xFB-0116	FB-0092xFB-0116-19.2	GG	AA	Hap2
FB-0092xFB-0116	FB-0092xFB-0116-19.3	GA	GA	Het
FB-0092xFB-0116	FB-0092xFB-0116-19.4	GG	AA	Hap2
FB-0092xFB-0116	FB-0092xFB-0116-19.5	AA	GG	Hap1
FB-0092xFB-0070	FB-0092xFB-0070-28.1	GA	GA	Het
FB-0092xFB-0070	FB-0092xFB-0070-28.2	GA	GA	Het
FB-0092xFB-0070	FB-0092xFB-0070-28.3	GA	GA	Het
FB-0092xFB-0070	FB-0092xFB-0070-28.4	GG	AA	Hap2
FB-0092xFB-0070	FB-0092xFB-0070-28.5	GA	GA	Het
FB-0116xFB-0092	FB-0116xFB-0092-52.1	AA	GG	Hap1
FB-0116xFB-0092	FB-0116xFB-0092-52.2	AA	GG	Hap1
FB-0116xFB-0092	FB-0116xFB-0092-52.3	GG	AA	Hap2
FB-0116xFB-0092	FB-0116xFB-0092-52.4	GG	AA	Hap2
FB-0116xFB-0092	FB-0116xFB-0092-52.5	GG	AA	Hap2
FB-0070xFB-0092	FB-0070xFB-0092-64.1	GG	AA	Hap2

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0070xFB-0092	FB-0070xFB-0092-64.2	GA	GA	Het
FB-0070xFB-0092	FB-0070xFB-0092-64.3	AA	GG	Hap1
FB-0070xFB-0092	FB-0070xFB-0092-64.4	GG	AA	Hap2
FB-0070xFB-0092	FB-0070xFB-0092-64.5	GA	GA	Het
FB-0092xFB-0070	FB-0092xFB-0070-30.3	AA	GG	Hap1
FB-0092xFB-0070	FB-0092xFB-0070-30.4	AA	GG	Hap1
FB-0092xFB-0070	FB-0092xFB-0070-30.5	GA	GA	Het
FB-0055xFB-0083	FB-0055xFB-0083-41.2	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-41.3	GG	GG	Hap3
FB-0055xFB-0083	FB-0055xFB-0083-41.4	GG	GA	Het
FB-0055xFB-0083	FB-0055xFB-0083-41.5	GG	GA	Het
FB-0083xFB-0055	FB-0083xFB-0055-13.1	GG	GA	Het
FB-0083xFB-0055	FB-0083xFB-0055-13.2	GG	GA	Het
FB-0083xFB-0055	FB-0083xFB-0055-13.3	GG	GA	Het
FB-0083xFB-0055	FB-0083xFB-0055-13.4	GG	GG	Hap3
FB-0083xFB-0055	FB-0083xFB-0055-13.5	GG	GA	Het
FB-0092xFB-0083	FB-0092xFB-0083-23.1	GG	GA	Het
FB-0092xFB-0083	FB-0092xFB-0083-23.2	GG	GA	Het
FB-0092xFB-0083	FB-0092xFB-0083-23.3	GG	GA	Het
FB-0092xFB-0083	FB-0092xFB-0083-23.4	GG	GG	Hap3
FB-0092xFB-0083	FB-0092xFB-0083-23.5	GG	AA	Hap2
FB-0113xFB-0055	FB-0113xFB-0055-66.1	GA	GA	Het
FB-0113xFB-0055	FB-0113xFB-0055-66.3	GA	GA	Het
FB-0113xFB-0055	FB-0113xFB-0055-66.5	AA	GG	Hap1
FB-0055xFB-0083	FB-0055xFB-0083-44.1	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-44.2	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-44.3	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-44.4	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-44.5	GG	AA	Hap2
FB-0117xFB-0055	FB-0117xFB-0055-37.1	GA	GA	Het

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0117xFB-0055	FB-0117xFB-0055-37.2	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-37.3	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-37.4	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-37.5	GG	AA	Hap2
FB-0070xFB-0055	FB-0070xFB-0055-1.1	AA	GG	Hap1
FB-0070xFB-0055	FB-0070xFB-0055-1.2	AA	GG	Hap1
FB-0070xFB-0055	FB-0070xFB-0055-1.3	AA	GG	Hap1
FB-0070xFB-0055	FB-0070xFB-0055-1.4	GG	AA	Hap2
FB-0070xFB-0055	FB-0070xFB-0055-1.5	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-4.1	AA	GG	Hap1
FB-0092xFB-0113	FB-0092xFB-0113-4.2	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-4.3	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-4.4	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-4.5	GG	AA	Hap2
FB-0116xFB-0092	FB-0116xFB-0092-49.1	GA	GA	Het
FB-0116xFB-0092	FB-0116xFB-0092-49.2	AA	GG	Hap1
FB-0116xFB-0092	FB-0116xFB-0092-49.3	AA	GG	Hap1
FB-0116xFB-0092	FB-0116xFB-0092-49.4	GA	GA	Het
FB-0116xFB-0092	FB-0116xFB-0092-49.5	GG	AA	Hap2
FB-0092xFB-0070	FB-0092xFB-0070-33.1	AA	GG	Hap1
FB-0092xFB-0070	FB-0092xFB-0070-33.2	GG	AA	Hap2
FB-0092xFB-0070	FB-0092xFB-0070-33.3	AA	GG	Hap1
FB-0092xFB-0070	FB-0092xFB-0070-33.4	GA	GA	Het
FB-0092xFB-0070	FB-0092xFB-0070-33.5	GA	GA	Het
FB-0055xFB-0070	FB-0055xFB-0070-9.1	GA	GA	Het
FB-0055xFB-0070	FB-0055xFB-0070-9.2	GA	GA	Het
FB-0055xFB-0070	FB-0055xFB-0070-9.3	GA	GA	Het
FB-0055xFB-0070	FB-0055xFB-0070-9.4	GG	AA	Hap2
FB-0055xFB-0070	FB-0055xFB-0070-9.5	GG	AA	Hap2
FB-0055xFB-0117	FB-0055xFB-0117-16.1	AA	GG	Hap1

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0055xFB-0117	FB-0055xFB-0117-16.2	GA	GA	Het
FB-0055xFB-0117	FB-0055xFB-0117-16.3	GA	GA	Het
FB-0055xFB-0117	FB-0055xFB-0117-16.4	GG	AA	Hap2
FB-0055xFB-0117	FB-0055xFB-0117-16.5	GG	AA	Hap2
FB-0092xFB-0117	FB-0092xFB-0117-55.1	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-55.2	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-55.3	AA	GG	Hap1
FB-0092xFB-0117	FB-0092xFB-0117-55.4	GG	AA	Hap2
FB-0092xFB-0117	FB-0092xFB-0117-55.5	GA	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-58.1	GG	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-58.2	GG	AA	Hap2
FB-0083xFB-0092	FB-0083xFB-0092-58.3	GG	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-58.4	GG	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-58.5	GG	AA	Hap2
FB-0092xFB-0070	FB-0092xFB-0070-32.1	AA	GG	Hap1
FB-0092xFB-0070	FB-0092xFB-0070-32.2	GA	GA	Het
FB-0092xFB-0070	FB-0092xFB-0070-32.3	AA	GG	Hap1
FB-0092xFB-0070	FB-0092xFB-0070-32.4	GG	AA	Hap2
FB-0092xFB-0070	FB-0092xFB-0070-32.5	AA	GG	Hap1
FB-0117xFB-0092	FB-0117xFB-0092-11.1	GA	GA	Het
FB-0117xFB-0092	FB-0117xFB-0092-11.2	GA	GA	Het
FB-0117xFB-0092	FB-0117xFB-0092-11.3	GA	GA	Het
FB-0117xFB-0092	FB-0117xFB-0092-11.4	GA	GA	Het
FB-0117xFB-0092	FB-0117xFB-0092-11.5	AA	GG	Hap1
FB-0092xFB-0117	FB-0092xFB-0117-27.1	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-27.2	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-27.3	AA	GG	Hap1
FB-0092xFB-0117	FB-0092xFB-0117-27.4	GG	AA	Hap2
FB-0092xFB-0117	FB-0092xFB-0117-27.5	GA	GA	Het
FB-0070xFB-0055	FB-0070xFB-0055-3.7	GG	AA	Hap2

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0070xFB-0055	FB-0070xFB-0055-3.8	GA	GA	Het
FB-0070xFB-0055	FB-0070xFB-0055-3.9	GA	GA	Het
FB-0070xFB-0055	FB-0070xFB-0055-3.1	GA	GA	Het
FB-0092xFB-0070	FB-0092xFB-0070-31.1	GG	AA	Hap2
FB-0092xFB-0070	FB-0092xFB-0070-31.2	GG	AA	Hap2
FB-0092xFB-0070	FB-0092xFB-0070-31.3	GG	AA	Hap2
FB-0092xFB-0070	FB-0092xFB-0070-31.4	GG	AA	Hap2
FB-0092xFB-0070	FB-0092xFB-0070-31.5	GG	AA	Hap2
FB-0083xFB-0092	FB-0083xFB-0092-59.1	GG	GG	Hap3
FB-0083xFB-0092	FB-0083xFB-0092-59.2	GG	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-59.3	GG	GG	Hap3
FB-0083xFB-0092	FB-0083xFB-0092-59.4	GG	GA	Het
FB-0083xFB-0092	FB-0083xFB-0092-59.5	AA	GG	Hap1
FB-0117xFB-0055	FB-0117xFB-0055-38.1	AA	GG	Hap1
FB-0117xFB-0055	FB-0117xFB-0055-38.2	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-38.3	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-38.4	AA	GG	Hap1
FB-0117xFB-0055	FB-0117xFB-0055-38.5	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-54.1	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-54.2	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-54.3	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-54.4	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-54.5	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-43.1	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-43.2	GA	GA	Het
FB-0055xFB-0083	FB-0055xFB-0083-43.4	GA	GA	Het
FB-0055xFB-0083	FB-0055xFB-0083-43.5	GA	GA	Het
FB-0113xFB-0055	FB-0113xFB-0055-69.1	GA	GA	Het
FB-0113xFB-0055	FB-0113xFB-0055-69.2	AA	GG	Hap1
FB-0113xFB-0055	FB-0113xFB-0055-69.3	GA	GA	Het

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0113xFB-0055	FB-0113xFB-0055-69.4	GG	AA	Hap2
FB-0113xFB-0055	FB-0113xFB-0055-69.5	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-26.1	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-26.2	AA	GG	Hap1
FB-0092xFB-0117	FB-0092xFB-0117-26.3	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-26.4	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-26.5	GA	GA	Het
FB-0055xFB-0113	FB-0055xFB-0113-47.6	AA	GG	Hap1
FB-0055xFB-0113	FB-0055xFB-0113-47.7	GA	GA	Het
FB-0055xFB-0113	FB-0055xFB-0113-47.8	GG	AA	Hap2
FB-0055xFB-0113	FB-0055xFB-0113-47.9	GG	AA	Hap2
FB-0055xFB-0113	FB-0055xFB-0113-47.1	GA	GA	Het
FB-0083xFB-0055	FB-0083xFB-0055-13.7	GG	GA	Het
FB-0083xFB-0055	FB-0083xFB-0055-13.1	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-42.1	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-42.2	GA	GA	Het
FB-0055xFB-0083	FB-0055xFB-0083-42.3	GA	GA	Het
FB-0055xFB-0083	FB-0055xFB-0083-42.4	GG	AA	Hap2
FB-0055xFB-0083	FB-0055xFB-0083-42.5	GA	GA	Het
FB-0116xFB-0055	FB-0116xFB-0055-60.1	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-53.1	GG	AA	Hap2
FB-0092xFB-0117	FB-0092xFB-0117-53.2	GG	AA	Hap2
FB-0092xFB-0117	FB-0092xFB-0117-53.3	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-53.4	AA	GG	Hap1
FB-0092xFB-0117	FB-0092xFB-0117-53.5	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-71.1	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-71.2	GA	GA	Het
FB-0113xFB-0092	FB-0113xFB-0092-71.3	GG	AA	Hap2
FB-0113xFB-0092	FB-0113xFB-0092-71.4	AA	GG	Hap1
FB-0113xFB-0092	FB-0113xFB-0092-71.5	GA	GA	Het

Cross	Genotype	SNP_b1	SNP08	Haplotype
FB-0117xFB-0055	FB-0117xFB-0055-39.1	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-39.2	GG	AA	Hap2
FB-0117xFB-0055	FB-0117xFB-0055-39.3	GG	AA	Hap2
FB-0117xFB-0055	FB-0117xFB-0055-39.4	GA	GA	Het
FB-0117xFB-0055	FB-0117xFB-0055-39.5	GG	AA	Hap2
FB-0092xFB-0113	FB-0092xFB-0113-6.6	AA	GG	Hap1
FB-0092xFB-0113	FB-0092xFB-0113-6.7	GA	GA	Het
FB-0092xFB-0113	FB-0092xFB-0113-6.8	AA	GG	Hap1
FB-0092xFB-0113	FB-0092xFB-0113-6.9	GG	AA	Hap2
FB-0092xFB-0113	FB-0092xFB-0113-6.1	GG	AA	Hap2
FB-0092xFB-0117	FB-0092xFB-0117-53.6	AA	GG	Hap1
FB-0092xFB-0117	FB-0092xFB-0117-53.7	GA	GA	Het
FB-0092xFB-0117	FB-0092xFB-0117-53.8	AA	GG	Hap1
FB-0092xFB-0117	FB-0092xFB-0117-53.9	AA	GG	Hap1
FB-0092xFB-0117	FB-0092xFB-0117-53.1	GA	GA	Het

List of abbreviations

AFLP	Amplified Fragment Length Polymorphism
ANFs	Antinutritional Factors
ARPDP	5-Amino-6-Ribosylamino-2,3(1H,3H)-Pyrimidinedione 5′-Phosphate
BCF2	Backcross Filial Generation 2
BLAST	Basic Local Alignment Search Tool
CAPs	Cleaved Amplified Polymorphic Sequence
cDNA	Complementary DNA
CIA	Chymotrypsin Inhibitory Activity
CNV	Copy number variation
CTs	Condensed Tannins
DARPP	2,5-Diamino-6-Ribosylamino-4(3H)-Pyrimidinone 5′-Phosphate
DHBPS	3,4-Dihydroxy-2-Butanone-4-Phosphate Synthase
DNA	Deoxyribonucleic Acid
ESF	Early Seed Filling
FAD	Flavin Adenine Dinucleotide
FAOSTAT	Food and Agriculture Organization Corporate Statistical Database
FMN	Flavin Mononucleotide
G6PD	Glucose-6-Phosphate Dehydrogenase
GCHII	Guanosine Triphosphate Cyclohydrolase II
gDNA	Genomic DNA
Gbp	Giga base pairs
GTP	Guanosine Triphosphate
HVC	High Vicine-Convicine
IPCC	Intergovernmental Panel on Climate Change

KASP	Kompetitive Allele-Specific PCR
LSF	Late Seed Filling
LVC	Low Vicine-Convicine
miRNA	MicroRNA
MAS	Marker-Assisted Selection
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NCBI	National Center for Biotechnology Information
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
QTL	Quantitative Trait Locus
RAPD	Random Amplified Polymorphic DNA
RFOs	Raffinose Family Oligosaccharides
RIBA	Riboflavin Biosynthesis Protein
RIL	Recombinant Inbred Line
RNA	Ribonucleic Acid
RT-qPCR	Reverse Transcription Quantitative Polymerase Chain Reaction
SNP	Single Nucleotide Polymorphism
TIA	Trypsin Inhibitory Activity
TPA	Target Pre-Amplification
TPC	Total Phenolic Content
v-c	Vicine and Convicine
VC1	*VICINE-CONVICINE 1*

Declaration

I declare that the dissertation here submitted is entirely my own work, written without any illegitimate help by any third party and solely with materials as indicated in the dissertation. I have indicated in the text where I have used texts from already published sources, either word for word or in substance, and where I have made statements based on oral information given to me. At all times during the investigations carried out by me and described in the dissertation, I have followed the principles of good scientific practice as defined in the 'Statutes of Justus Liebig University Giessen for the Safeguarding of Good Scientific Practice'

Giessen, 03/12/2025 Samson Ugwuanyi

Acknowledgements

I would like to acknowledge the incredible support of all those who were instrumental in the completion of this thesis. My heartfelt appreciation goes to everyone who, directly or indirectly, contributed to this work. Words are not enough to express my gratitude, but I would like to mention some individuals whose contributions were truly integral.

First, my deepest gratitude goes to my doctoral supervisor, Prof. Dr. Rod Snowdon, for giving me this wonderful opportunity to pursue my doctorate at the Institute. Since my very first visit to the institute as a free mover student to conduct my master's thesis, you have always been supportive. There was never a time I needed help that you did not provide. Your door was always open; your reassuring words and constant encouragement were a continual source of motivation. You were always ready to listen and advise. I sincerely appreciate your guidance and financial support, which made this work possible. I also gratefully acknowledge the financial support of the German Research Society (DFG), without which this thesis would not have been possible.

I would like to thank Dr. Christian Obermeier for his immense support throughout this work. Whenever I had questions or challenges, you always had ideas to help navigate them. I am thankful for the countless times we sat in your office to discuss and troubleshoot together. My gratitude also goes to Dr. Benjamin Wittkop for his support and funding, and to Prof. Dr. Agnieszka Golicz and Prof. Dr. Sarah Schiessl-Weidenweber for their immense support.

I am especially grateful to Stavros Tzigos for his invaluable technical support in the laboratory and greenhouse. You were always available to assist with technical work, and your advice was especially helpful when things did not go as expected. I will never forget how you encouraged me when I was struggling with an experiment, reminding me: "Samson, it's an experiment; sometimes it works, sometimes it doesn't." These words still encourage me today, and I even use them to motivate colleagues facing similar challenges.

I also wish to thank Annette Plank, Roland Kuerschner, Sabine Frei, Regina Illgner, and all the technicians at the Institute for their valuable assistance. I deeply appreciate Dr. Manar Makhoul, Dr. Sven Weber, Dr. Sruthy Augustine, Lennart Scheer, Erick Mikwa, Ashwini Mudke, Lennard Roscher-Ehrig, Rica-Hanna Schlichtermann, Jose Montero Tena, Kübra Arslan, Mamunur Rashid, Andris Finkbeiner, and all my colleagues at the Institute for their incredible support. My thanks also go to the secretaries, Mrs. Sabine Schomber and Mrs. Rebecca Feldbusch, for

their continuous help, and to Erika Schick and Liane Renno for their valuable technical support. I am also grateful to the team at Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ, Hohenlieth, Germany), especially Hanna Tietgen and Amine Abbadi, for their kind assistance.

I am profoundly indebted to my beautiful wife, Amarachi Queen, my most precious gift from God, for her unwavering love, sacrifices, support and prayers. Your advice, encouragement and constant presence as my support system have been priceless. Knowing you are always there for me has been my greatest motivation. I am also grateful to my adorable daughters, Chimbusomma and Urenna, for their love, encouragement and prayers. You are proof of God's faithfulness to me and a constant source of inspiration and strength.

I would also like to thank Dr. Elvis Katche and Dr. Elizabeth Katche for their advice and support. My heartfelt appreciation goes to my brother, Prof. Christian Ugwuanyi, and his family, for their love, prayers and support. I am deeply thankful to my parents, Nze Nicholas Ugwuanyi and Lolo Theresa Ugwuanyi, for their unconditional love, sacrifices, prayers and support. I also extend my gratitude to my siblings and their families, Kenneth, Ngozi, Juliet, Sarah, Ogochukwu, Chinedu, and Chiemezie, for their prayers and encouragement. I am grateful to my father-in-law for his advice and encouragement; you always reminded me to keep working hard on my thesis, and I do not take this for granted.

The list goes on and on, and while I cannot mention everyone by name, please know that your support has not gone unnoticed. I am deeply grateful to every one of you.

Finally, and above all, I give thanks to Almighty God for His loving kindness and faithfulness, for granting me strength and wisdom, for bringing me this far, and for making this work a reality.